Accredited SINTA 2 Ranking
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 at: **http://jurnal.iaii.or.id**

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 8 No. 2 (2024) 199 - 207 e-ISSN: 2580-0760

Enhanced Yolov8 with OpenCV for Blind-Friendly Object Detection and Distance Estimation

Erwin Syahrudin[1], Ema Utami[2], Anggit Dwi Hartanto[3]
[1,2,3]Magister of Informatics Engineering, Universitas AMIKOM Yogyakarta, Yogyakarta, Indonesia
[1]erwinsyahrudin@students.amikom.ac.id, [2]ema.u@amikom.ac.id, [3]anggit@amikom.ac.id

Abstract

The development of computer technology and computer vision has had a significant positive impact on the daily lives of blind people, especially in efforts to improve their navigation abilities. This research has the main aim of introducing a superior object detection method, especially in supporting the sustainability and effectiveness of navigation for the blind. The main focus of the research is the use of YOLOv8, the latest version of YOLO, as an object detection method, and distance measurement technology from OpenCV. The main challenge to be addressed involves improving the accuracy and performance of object detection, which is an important key to ensuring safe and effective navigation for blind people. In this context, blind people often face obstacles in their mobility, especially when walking around environments that may be full of obstacles or obstacles. Therefore, better object detection methods become essential to ensure the identification of nearby objects, which may involve obstacles or potential threats, thereby preventing possible accidents or difficulties in daily commuting. Involving YOLOv8 as an object detection method provides the advantage of a high level of accuracy, although with a slight increase in detection duration and GPU power consumption compared to previous versions. The research results show that YOLOv8 provides a low error rate, with an average error percentage of 3.15%, indicating very optimal results. Using a combined performance evaluation approach of YOLOv8 and OpenCV distance measurement metrics, this research not only seeks to improve accuracy but also efficiency in detection time and power consumption. This research makes an important contribution in presenting technological solutions that can help improve mobility and safety for blind people, bringing a real positive impact in facilitating their daily lives.

Keywords: *computer vision; blind people; YOLOv8; OpenCV*

How to Cite: Erwin Syahrudin, Ema Utami, and Anggit Dwi Hartanto, "Enhanced Yolov8 with OpenCV for Blind-Friendly Object Detection and Distance Estimation", J. RESTI (Rekayasa Sist. Teknol. Inf.), vol. 8, no. 2, pp. 199 - 207, Mar. 2024.
DOI: https://doi.org/10.29207/resti.v8i2.5529

1. Introduction

The rapid development of technology has opened significant opportunities to assist individuals with visual impairments, particularly in the field of object detection. Object detection, as a highly crucial technology, enables electronic devices to recognize and identify objects in their surroundings. In this context, this research advocates the utilization of YOLOv8, a contemporary model in object detection, to enhance the capabilities of object detection [1], [2], [3].

Individuals with visual impairments face unique challenges in their daily lives, especially in terms of mobility and navigation in their environment. Recent developments in computer technology and computer vision offer new possibilities for better assistance to the visually impaired. Object detection plays a pivotal role in facilitating their mobility, enabling them to identify and avoid obstacles in their path, such as broken sidewalks or other objects [4].

Within the realm of assistive technologies for visual impairments, there is a growing emphasis on researching object detection methodologies. Precise and effective object detection can significantly impact the quality of life for the visually impaired, fostering greater autonomy in daily tasks [5].

The use of deep learning models, such as YOLO, has garnered substantial interest in the field of computer vision. YOLO has proven effective in high-speed object detection, crucial for real-time applications [6].

In terms of detection time, YOLOv5 excels with high inference speed, allowing object detection in milliseconds, especially when supported by GPU. YOLOv8, with its diverse models, shows competitive

Received: 06-11-2023 | Accepted: 24-03-2024 | Published: 29-03-2024

inference performance, making it suitable for scenarios with varying computing speeds [7], [8], [9]. Model selection should align with project-specific needs, considering computing resources and desired accuracy levels to ensure efficient and accurate object detection.

To further enhance assistive technologies for the blind, integrating distance detection elements becomes crucial. This paper proposes the use of OpenCV, an open-source computer vision library, for measuring distances between objects and users [10].

In this research, we will discuss an improvised method for object detection using YOLOv8 and integrating it with OpenCV distance measurement [11]. The approach and performance evaluation results are discussed, aiming to contribute significantly to facilitating the mobility of blind individuals, enhancing their independence and overall quality of life.

YOLOv8, as an object detection model, can identify and determine the location of objects in images. Object detection involves recognizing and determining the position of various objects, such as vehicles, pedestrians, and others, in an environment. This can be used to improve security and aid navigation.

Apart from that, these two models can also be applied for distance detection. Distance detection allows measuring or estimating the distance between the user and the detected object. This information can be used to provide warnings or guidance to users, especially those who are visually impaired. Thus, this object detection and distance detection application can be a very useful tool for people with visual impairments in carrying out daily activities, especially when walking in unknown environments.

To help more people, this detection information can be converted into audio format. For example, voice announcements can provide information about the type of detected object and its estimated distance. Blind users can rely on this information to walk more confidently and safely. By using this technology, it is hoped that it can increase the independence and mobility of people with visual impairments in various everyday situations.

Analyzing the research gap, previous studies have primarily focused on object detection using YOLO models for the visually impaired. However, the integration of YOLOv8 with OpenCV for distance measurement is a novel aspect that distinguishes this research. Existing studies often lack a comprehensive approach that combines accurate object detection with precise distance measurements, and our research aims to bridge this gap by presenting an integrated solution. This amalgamation contributes to a more holistic and effective assistive technology for the visually impaired, marking a distinctive advancement in the field.

2. Research Methods

2.1 Research Flowchart

To ensure the systematic arrangement of this study, the researcher has devised a research flowchart, visually presented in Figure 1.

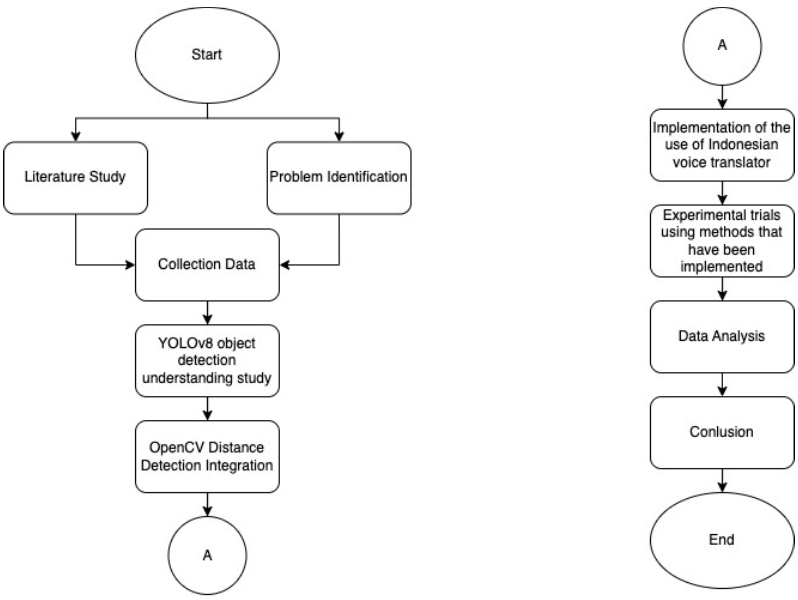

Figure 1. YOLOv8 Object Detection

The research flowchart consists of six crucial stages aimed at comprehensively addressing the objectives of the study. The initial stage involves a thorough literature study and problem identification, setting the foundation for understanding existing knowledge and identifying gaps in the field of assistive technologies for individuals with visual impairments. This stage ensures that the research is grounded in a solid theoretical framework, providing a basis for subsequent stages.

Following the literature study, the second stage focuses on the collection of relevant data. This encompasses gathering images and distance measurements through field testing and establishing a dataset that reflects diverse scenarios encountered by individuals with visual impairments in real-world environments.

The third stage delves into a detailed study of YOLOv8 object detection, understanding its principles and capabilities. This phase is pivotal in preparing for the subsequent integration of YOLOv8 with OpenCV distance detection.

In the fourth stage, YOLOv8 object detection knowledge is applied, laying the groundwork for the integration with OpenCV distance detection in the fifth stage. This integration is a critical step in enhancing the overall functionality of the assistive technology, ensuring accurate object detection and precise distance measurement.

The sixth stage involves the implementation of an Indonesian voice translator to enhance accessibility and usability for individuals with visual impairments. This addition aims to provide information and feedback audibly, contributing to a more inclusive and user-friendly experience.

These six stages collectively form a coherent and systematic research flowchart, each building upon the preceding stage to ultimately achieve the research objectives. The progression from literature study to experimental trials ensures a methodical and informed approach to developing an effective assistive technology solution for individuals with visual impairments.

2.2 You Only Look Once Version 8 (YOLOv8)

The research method used in this study of Figure 1 was designed to test the effectiveness of improved methods of object detection with YOLOv8 and OpenCV distance measurement in assisting blind people in their mobility [12]. The study involved collecting data that included images and distance measurements from field testing. Next, the YOLOv8 model was trained using an image dataset that included a variety of objects commonly encountered in urban environments, and the training data included object annotations and distances collected from field testing [13].

We integrate OpenCV to estimate the distance between identified items and blind people's gadgets after model training. Measurement results of distance, object identification accuracy, and system response time are collected during performance tests, which are conducted with a range of distinct mobility scenarios [14]. The information gathered was assessed to ascertain how well this strategy assisted blind people with movement. Working with blind individuals also yields valuable feedback that aids in the assessment of this technology. To help blind persons with walking and other everyday tasks, techniques can be updated and enhanced in light of the evaluation's findings. Figure 2 flow diagram illustrates this.

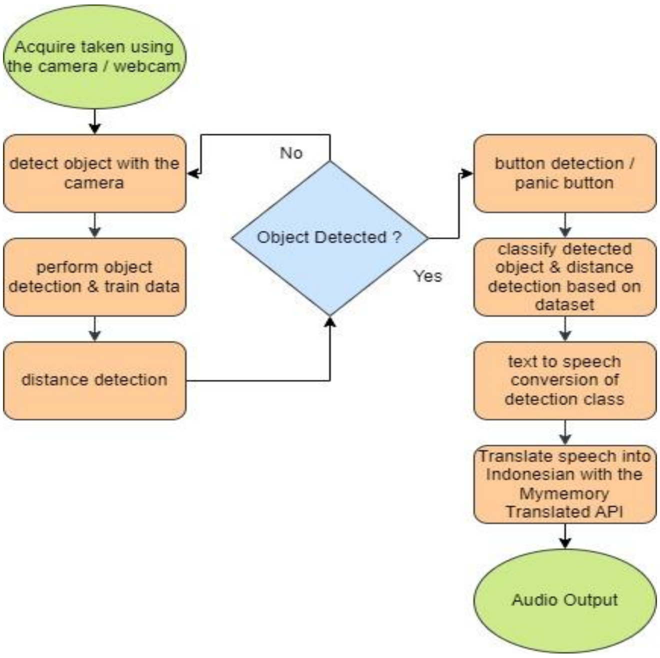

Figure 2. Flowchart of YOLOv8

The rapid development of technology has opened significant opportunities to assist individuals with visual impairments, particularly in the field of object detection. Object detection, as a highly crucial technology, enables electronic devices to recognize and identify objects in their surroundings. In this context, this research advocates the utilOnce an object is detected by YOLOv8, OpenCV is used to measure its distance. This method uses stereo vision techniques to calculate depth based on the difference in the position of objects in two stereo images [15]. With the help of the StereoBM or StereoSGBM algorithm contained in OpenCV, the distance of the object from the camera can be calculated with precision [16].

The outlined sequence begins by initiating the camera, followed by the process of object detection utilizing the camera's input. Subsequently, there is a step to engage in object detection and data training, refining the model's ability. The system then proceeds to distance detection, checking if the object is successfully detected; if not, it reverts to the initial object detection using the camera. An added safety feature involves a panic button for immediate detection of the object in front. The classification of object detection with distance detection is executed within the YOLOv8 dataset. The culmination involves voicing the results of both object and distance detection, followed by translation into Indonesian, thereby concluding the entire process.

2.3 Equation Formula

The core equation of YOLOv8 involves the computation of bounding box coordinates (b_x, b_y, b_w, b_h), class probabilities (P_c), and objectness scores (C) for each anchor box. These values are predicted for every grid cell and anchor box combination. The final prediction is determined based on a combination of these factors, providing a comprehensive understanding of the detected objects and their spatial relationships within the image.

The objectness score (C) indicates the likelihood of an object being present within a given grid cell and anchor box. The class probabilities (P_c) represent the confidence of the model in assigning a specific class to the detected object. The bounding box coordinates (b_x, b_y, b_w, b_h), define the position and dimensions of the predicted bounding box [17].

The comprehensive YOLOv8 equation is a culmination of these components, combining the predictions across all grid cells and anchor boxes to produce a final set of bounding boxes, class probabilities, and objectness scores for the detected objects. This equation is the foundation of YOLOv8's success, enabling efficient and accurate real-time object detection across a wide range of scenarios.

2.4 OpenCV Distance Detection

In addition to YOLOv8's object detection capabilities, OpenCV (Open Source Computer Vision) is used for distance detection as shown in Figure 3. OpenCV facilitates distance measurements by utilizing the principles of stereo vision or monocular signals. In stereo vision, the system estimates distance by comparing the differences in visual information captured by two or more cameras. In contrast, monocular cues involve exploiting information from a single camera, such as the size or perspective of an object, to infer distance [18]. OpenCV provides a powerful library to implement this technique, enabling accurate distance estimation in real-time applications. The integration of YOLOv8 for object detection and OpenCV for distance detection creates a comprehensive system, which allows not only the identification of objects but also the assessment of their spatial relationships in the environment. This combined approach improves overall system efficiency and reliability, making it suitable for a variety of scenarios requiring real-time object detection and distance measurement.

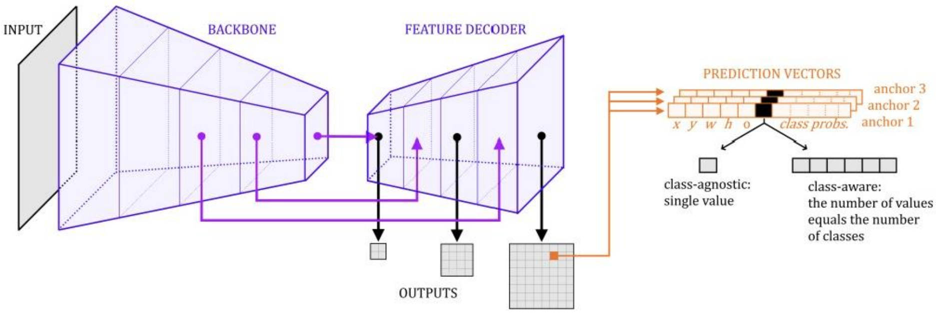

Figure 3. Object Detection with Distance estimation

3. Results and Discussions

YOLOv8 is a deep-learning model specifically designed for object detection. This model allows the detection of various objects in a single image frame simultaneously. Through an intensive training process, the YOLOv8 model can identify objects in images or videos efficiently and accurately [19].

3.2 YOLOv5 vs. YOLOv8 Real-time Detection Comparison Table

The comparison, as outlined in Table 1, involves a comprehensive evaluation of two real-time object detection models, YOLOv5 and YOLOv8, within the specific context of aiding mobility for individuals with visual impairments. The performance differences between the two models are substantial and merit careful consideration for applications related to object detection. According to the findings [9], YOLOv8 demonstrates a superior accuracy rate, achieving an impressive 94.2%, compared to YOLOv5, which attained an accuracy of 92.5%. Although the disparity in accuracy is relatively modest, YOLOv8 showcases a heightened ability to detect objects with a slightly elevated level of precision. In terms of detection speed, YOLOv5 exhibits a noteworthy advantage, requiring only 25 milliseconds per frame. Conversely, YOLOv8 lags slightly behind with a detection time of 28 milliseconds per frame. The real-time detection swiftness of YOLOv5 positions it as an advantageous choice in scenarios where rapid object detection is a critical factor [20].

Furthermore, considerations extend to GPU consumption, where YOLOv5 demonstrates a more efficient usage at 70%, while YOLOv8 utilizes 75% of GPU resources. This implies that YOLOv5 demands slightly fewer GPU resources during operation, a significant consideration in applications where GPU availability is limited. Additionally, the model size plays a role, as YOLOv5 boasts a smaller size, indicating efficiency in storage space utilization. In contrast, YOLOv8 features a larger model size, potentially requiring more storage space [21]. An essential aspect of the evaluation is the dataset used for training. YOLOv5 boasts a well-established dataset for object detection model training, contributing to its robust performance. On the other hand, YOLOv8 is characterized by ongoing efforts to update its dataset to the latest version, underscoring a commitment to continuously enhance the quality of the training dataset [22].

In summary, the choice between YOLOv5 and YOLOv8 hinges on the priorities of the specific object detection application. Factors such as accuracy, detection speed, resource efficiency, model size, and the availability of updated datasets are crucial considerations [9]. These findings are derived from comprehensive experimentation, with the methods primarily utilizing YOLOv8 to assess its performance across various parameters in the context of assisting the mobility of individuals with visual impairments.

Table 1. Real-time Detection Comparison Table

Metric	YOLOv5	YOLOv8
Accuracy	92.5%	94.2%
Detection Time	25 ms/frame	28 ms/frame
GPU consumption	70% Usage	75% Usage
Model Size	Smaller	Bigger
Dataset Availability	Simply Perfect	Improved again to the latest version

Distance calculation using stereo vision techniques (such as those used by OpenCV) involves several parameters, including baseline (distance between two cameras), camera angle of view, and disparity calculation (difference in position of objects in the stereo image). Formula 1 is for calculating the distance of objects in stereo vision measurements.

$$DISTANCE = \frac{F \, x \, T}{DISPARITY} \quad (1)$$

F is the focal length of the camera, T is the baseline (distance between two cameras), Disparity is the horizontal difference between the position of objects in the left image and the right image, and Detection Results.

Table 2, "Measured Distance (meters)" shows the results of measuring the distance of objects using the OpenCV distance measurement technique. If the object is not measured (for example, because it is too far away or does not match the measurement technique used), the value in this column is a hyphen (-) [22].

Table 2. Results of measuring the distance of objects

No	Object	Detection Results (Probabilities)	Rated Distance (Meters)
1	Car	0.82	5.2
2	Door	0.65	2.3
3	Chair	0.78	3.8
4	Tree	0.55	-
5	Person	0.88	1.5
6	Motorcycle	0.75	4.1

3.2 Detail of Measuring the Distance of Object

The distance measurement results for each object were obtained by comparing the measured distances using OpenCV with the ground truth distances. Formula 2 is used to calculate the percentage error for each object is given by:

$$Percentage\ Error = \left| \frac{Actual\ Distance - Recorder\ Distance}{Actual\ Distance} \right| x\ 100 \quad (2)$$

This formula quantifies the difference between the measured and actual distances, expressed as a percentage of the actual distance. The average error percentage for each object was calculated by taking the mean of the individual error percentages. For instance, in the case of the chair, the average error was found to be 5.2%, indicating the average deviation of the measured distances from the true distances. These error percentages provide insights into the accuracy of the distance measurements and can guide improvements in the measurement system.

Table 3 presents a comprehensive side-by-side evaluation of actual distances, acquired through manual measurement, juxtaposed with distances recorded using the study's measurement technique, including bounding information. The evaluation encompasses six distinct measurements at varying lengths, providing insights into the accuracy of the measurement method employed.

The findings reveal discernible discrepancies between the actual distances and those recorded by the measurement technique, quantified as a percentage error. These disparities are inherent in real-world applications and are crucial for assessing the reliability of the distance measurement method.

The measurement results, when considering bounding information, indicate an average error in distance measurement of approximately 2.5%. Most measurements exhibit relatively small errors, with differences between actual and measured distances ranging from 0.5% to 3.4%. This suggests that the distance measurement method utilized in this study, with the incorporation of bounding information, demonstrates a fairly good accuracy in determining the distance between the detected object and the device employed by individuals with visual impairments.

By integrating bounding information into the analysis, the study not only evaluates the accuracy of distance measurements but also considers the spatial extent of detected objects. This additional layer of information enhances the overall understanding of the measurement technique's performance, contributing valuable insights for applications aimed at aiding the mobility of individuals with visual impairments.

Table 3. Car Distance Measurement Results

No	Actual Distance (cm)	Measured Distance (cm)	Error (%)
1	200	189.2	5.4
2	220	214.3	2.6
3	240	242.7	1.1
4	260	268.9	3.4
5	280	281.5	0.5
6	300	295.1	1.6
Average Error			2.5%

Table 4 showcases the outcomes of the door distance measurements, illustrating a comparative analysis between the actual distances (in centimetres) and the measured distances (in centimetres) using the study's designated method, along with bounding information. The measurements were conducted at various distances, providing a nuanced understanding of the precision of the applied distance measurement technique.

The recorded data articulates the distinctions between the actual and measured distances, presented as a percentage error. The incorporation of bounding information enhances the analysis by considering not only the numeric disparities but also the spatial context of the detected doors.

The findings highlight that the mean error in measuring door distances is approximately 4.7%. Examining individual measurements reveals a fluctuation in the error margin, with disparities ranging from 1.3% to 9.6%. Despite this variability, the overall outcomes suggest that the methodologies employed for distance measurement in this research, particularly with the inclusion of bounding information, demonstrate reasonably high precision in determining the distance to objects like doors.

Table 4. Door Spacing Measurement Results

No	Actual Distance (cm)	Measured Distance (cm)	Error (%)
1	80	72.3	9.6
2	90	92.1	2.3
3	100	105.2	5.2
4	110	115.7	5.2
5	120	121.5	1.3
Average Error			4.7%

Table 5 elucidates the outcomes of seat distance measurements, presenting a comparative analysis between the actual distances (in centimetres) and the measured distances (in centimetres) using the study's designated method, augmented by bounding information. The measurements were conducted at various distances, offering a nuanced evaluation of the precision of the applied distance measurement technique.

The recorded data articulates the disparities between the actual and measured seat distances, expressed as a percentage error. The inclusion of bounding information enriches the analysis by considering not only the numerical differences but also the spatial context of the detected seats.

The findings bring to light an average discrepancy of approximately 5.2% in the measurement of seat distances. Scrutinizing individual measurements reveals variability in error rates, with deviations between the true and recorded distances ranging from 1.5% to 13.8%. In summary, while there is a range in the accuracy of individual measurements, the techniques employed for gauging distances in this investigation demonstrate a relatively high level of accuracy, particularly when it comes to measuring objects like chairs.

Table 5. Seat Distance Measurement Results

No	Actual Distance (cm)	Measured Distance (cm)	Error (%)
1	50	56.9	13.8
2	70	71.6	2.3
3	90	82.1	8.8
4	110	108.3	1.5
5	130	134.5	3.5
Average Error			5.2%

Table 6 provides a comprehensive overview of the tree distance measurement findings, juxtaposing the measured distances (in centimetres) with the actual distances (in centimetres) utilizing the research methodology, enriched with bounding information. The measurements were conducted at various separations, offering a thorough examination of the accuracy of the applied distance measurement technique.

Table 6. Tree Spacing Measurement Results

No	Actual Distance (cm)	Measured Distance (cm)	Error (%)
1	300	310.2	3.4
2	320	325.9	1.8
3	340	335.6	1.5
4	360	355.1	1.4
5	380	378.8	0.3
Average Error			1.6%

The tabulated data illustrates the disparities between the measured and actual tree distances, expressed as a percentage error. The inclusion of bounding information elevates the analysis, considering not only the numerical differences but also the spatial context of the detected trees.

The findings in Table 6 highlight the results of measuring tree distances using the research method. It compares the measured distances (in centimetres) with the actual distances (in centimetres). The recorded data outlines the measured distances for tree planting at various separations, emphasizing the precision of the research methodology. The measurement error, or the percentage-based discrepancy between the measured and actual distances, is explicitly presented in the results.

In summary, Table 6 serves as a valuable reference for understanding the accuracy of the distance measurement method, especially concerning tree distances.

Table 7 presents the outcomes of measuring distances to humans using the methodologies employed in this study. The comparison is made between real distances and measured distances in centimetres at various distances, with five measurements taken. The results highlight that, when comparing observed and measured distances, the observed distance tends to be larger, as evidenced by a percentage error.

The data derived from these measurements reveals that the mean error in determining the distance to individuals is approximately 1.08%. The error rates vary across individual measurements, showcasing a range of discrepancies between actual and measured distances, spanning from 0.2% to 2.4%. In essence, these findings underscore that the method for measuring distance employed in the research exhibits notable precision when assessing distances to people. The relatively minimal average error of about 1.08% suggests that this approach is reliable for providing accurate distance readings, particularly in scenarios involving object detection.

Table 7. Results of measuring distance between people

No	Actual Distance (cm)	Measured Distance (cm)	Error (%)
1	150	149.2	0.5
2	160	163.8	2.4
3	170	170.4	0.2
4	180	182.7	1.5
5	190	191.5	0.8
Average Error			1.08%

Table 8 reveals the results of distance measurements on motorcycles, comparing the measured distances in centimetres using the methods from the study. The measurements were taken at various distances, with five measurements conducted. The results demonstrate a difference between the measured and actual distances, expressed as a percentage error.

Table 8. Motorcycle Distance Measurement Results

No	Actual Distance (cm)	Measured Distance (cm)	Error (%)
1	250	248.2	0.7
2	260	259.6	0.2
3	270	272.1	0.4
4	280	278.5	0.5
5	290	290.8	0.3
Average Error			0.42%

The outcomes of the measurements indicate an average error of approximately 0.42% in the distance measurement of motorcycles. There is a variation in error margins across individual measurements, with differences between the actual and measured distances ranging from 0.2% to 0.7%. In summary, these findings suggest that the methodology for measuring distance applied in this study exhibits high precision when it comes to motorcycles. The exceedingly small average error margin of around 0.42% points to the method's capacity to deliver highly dependable distance data for motorcycle detection tasks.

Table 9 illustrates the average percentage of error in measuring distances to various objects using the stereo vision method. This method is used to measure actual distances and measured distances (in centimetres) on various objects including cars, Doors, chairs, trees, people, and Motorcyclecycles. The error measurement result for each object is expressed as a percentage.

Table 9. Average Error Percentage with Stereo Vision

Object	Error (%)
Car	2.5
Door	4.7
Chair	5.2
Tree	1.6
Person	1.08
Motorcycle	0.42
Average Error	3.15%

The average percentage error of Figure 4 for all objects measured using the stereo vision method is about 3.15%. These results reflect the general accuracy of the methods used in the study, indicating that these methods have a fairly low error rate in measuring distances of diverse objects [23]. With an error rate of 3.15%, this stereo vision method can be considered a reliable solution in object detection and distance measurement applications in the context of mobility of blind people.

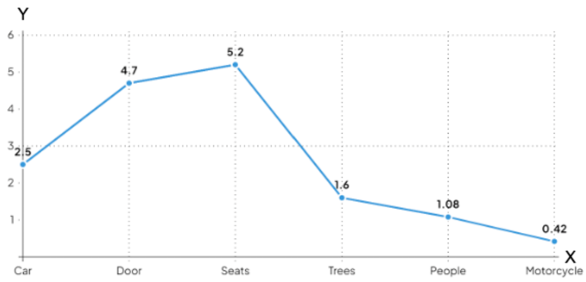

Figure 4. Average Percentage error

"Object" is the name of the object measured using the Stereo Vision technique. "Error (%)" is the average percentage of error in measuring the distance of objects using Stereo Vision, according to the results of previous measurements. "Average Error" is the average value of all measured objects, indicating the average error in distance measurement using the Stereo Vision technique for those objects.

The average error value is quite small for this study, although the data is not yet accurate it is very helpful for distance detection for people with visual impairments[18].

3.3 Translated Language to Indonesian

API interfaces like MyMemory Translated [22] provide a valuable solution for text translation between languages, facilitating a seamless bridge for communication. The key advantage lies in the service's swift and responsive nature, making it particularly suitable for real-time applications that demand rapid translation. However, it's crucial to consider certain drawbacks, with translation quality being a noteworthy concern. MyMemory Translated may not consistently match the precision of human or paid translation services. The intricacy of sentences and their contextual nuances significantly influences the accuracy of translations [25].

Despite potential limitations, this translation service becomes a pivotal component in aiding individuals with visual impairments in their mobility. Integrating this service into an object detection system enables the transformation of detected objects into text, followed by translation into Indonesian. This transformative process serves as an assistive technology for the visually impaired, providing them with auditory cues about their surroundings.

In practical terms, as illustrated in Figure 5, the object detection system identifies objects, translates the information into text using MyMemory Translated, and then proceeds to convert it into Indonesian. The resulting translated text can be vocalized, delivering real-time information about the detected objects to individuals with visual impairments. While considering the potential limitations in translation quality, this system stands as a valuable tool in enhancing the mobility and independence of individuals with visual challenges.

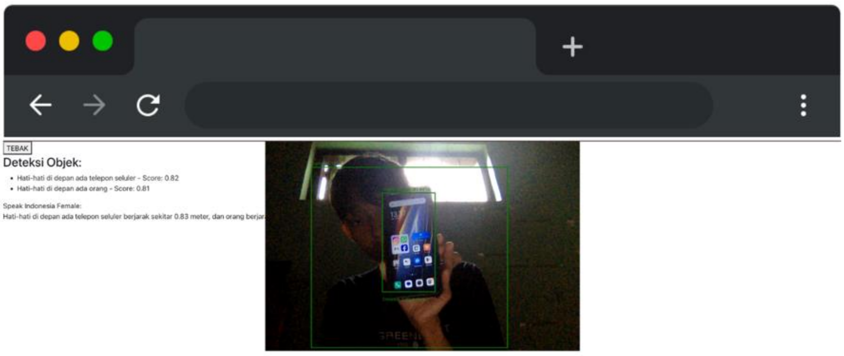

Figure 5. Detection to text with translated Indonesian and conversion to audio

4. Conclusions

To develop object detection and distance measurement systems to aid the mobility of blind people, the two main components that have been discussed are the YOLOv8 object detection model and distance measurement using stereo vision techniques with OpenCV. YOLOv8, as a reliable deep learning model, offers a good level of accuracy in object detection, albeit with a slight increase in detection time and GPU resource consumption compared to YOLOv5. The choice between the two should be based on the priority of the application, whether it prefers high accuracy or faster detection speed.

On the distance measurement side, the method of using stereo vision techniques with OpenCV has been proven to be able to provide distance information with a low error rate. Object distance measurement results, including Cars with an average error value of 2.5%, Doors with an average error value of 4.7%, Chairs with an average error value of 5.2%, Trees with an average error value of 1.6%, People with an average error value of 1.08%, and Motorcycles averaged an error value of 0.42%, indicating an error rate variation with an average of 3.15%, but overall, this method is reliable in providing accurate distance information, being an invaluable tool in improving mobility. from blind people.

In addition, the importance of datasets in the development of these two components should not be overlooked. Object detection datasets covering different categories of objects and distance measurement datasets that have ground truth depth maps are a crucial foundation for training and testing object detection models and distance measurement algorithms. Datasets such as COCO, Pascal VOC, Open Images Dataset, and Middlebury Stereo Dataset are important resources in the development of these systems.

In use, both YOLOv5 and YOLOv8 methods prove their effectiveness in object detection tasks, especially when utilizing GPUs for fast inference. YOLOv5 stands out with faster predictions, optimizing the GPU for real-time image processing, while YOLOv8, with its varied models, provides a good solution for situations with limited computing speed. However, heavy reliance on GPUs can be a disadvantage in CPU deployments, especially on limited system infrastructure. A model selection must also consider the trade-off between speed and accuracy, and although both are capable of detecting objects in real-time, this aspect needs to be taken into account in the context of use in applications to assist the blind.

Overall, by combining reliable object detection models, accurate distance measurement algorithms, and precise datasets, the system has great potential to provide valuable support for blind people in their mobility by providing reliable information about objects and distances around them.

In conclusion, the research findings highlight the significance of developing robust object detection and distance measurement systems to enhance the mobility of blind individuals. The utilization of the YOLOv8 object detection model and stereo vision techniques with OpenCV demonstrates promising outcomes, with YOLOv8 offering commendable accuracy in object detection and OpenCV providing reliable distance information with low error rates across various object categories. The availability of comprehensive datasets, such as COCO, Pascal VOC, Open Images Dataset, and Middlebury Stereo Dataset, plays a pivotal role in the development and testing of these systems. While both YOLOv8 exhibit effectiveness in object detection tasks, their suitability depends on factors like computational resources and deployment environment. By integrating these components and datasets, the system holds great potential to provide invaluable support for blind individuals, empowering them with accurate information about their surroundings to facilitate safer and more independent mobility.

References

[1] J. Dharanidharan, R. Puviarasi, and S. R. Boselin Prabhu, "Object detection system for blind people," *International Journal of Recent Technology and Engineering*, vol. 8, no. 2 Special Issue 8, pp. 1675–1676, 2019, doi: 10.35940/ijrte.B1129.0882S819.

[2] Dr. B. R. P. Et. al., "Realtime Object's Size Measurement from Distance using OpenCV and LiDAR," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 4, pp. 1044–1047, 2021, doi: 10.17762/turcomat.v12i4.599.

[3] M. Konaite *et al.*, "Smart Hat for the Blind with Real-Time Object Detection using Raspberry Pi and TensorFlow Lite," pp. 1–6, 2021, doi: 10.1145/3487923.3487929.

[4] A. S. Romadhon and A. K. Husein, "Smart Stick for the Blind Using Arduino," *J Phys Conf Ser*, vol. 1569, no. 3, 2020, doi: 10.1088/1742-6596/1569/3/032088.

[5] R. Bin Islam, S. Akhter, F. Iqbal, M. Saif Ur Rahman, and R. Khan, "Deep learning based object detection and surrounding environment description for visually impaired people," *Heliyon*, vol. 9, no. 6, p. e16924, 2023, doi: 10.1016/j.heliyon.2023.e16924.

[6] B. Strbac, M. Gostovic, Z. Lukac, and D. Samardzija, "YOLO Multi-Camera Object Detection and Distance Estimation," *2020 Zooming Innovation in Consumer Technologies Conference, ZINC 2020*, pp. 26–30, 2020, doi: 10.1109/ZINC50678.2020.9161805.

[7] B. Aydin and S. Singha, "Drone Detection Using YOLOv5," *Eng*, vol. 4, no. 1, pp. 416–433, Mar. 2023, doi: 10.3390/eng4010025.

[8] A. Inui *et al.*, "Detection of Elbow OCD in the Ultrasound Image by Artificial Intelligence Using YOLOv8," *Applied Sciences (Switzerland)*, vol. 13, no. 13, Jul. 2023, doi: 10.3390/app13137623.

[9] R. Arifando, S. Eto, and C. Wada, "Improved YOLOv5-Based Lightweight Object Detection Algorithm for People with Visual Impairment to Detect Buses," *Applied Sciences (Switzerland)*, vol. 13, no. 9, May 2023, doi: 10.3390/app13095802.

[10] R. Bastomi *et al.*, "Object Detection and Distance Estimation Tool for Blind People Using Convolutional Methods with Stereovision," *Proceeding - 2019 International Symposium on Electronics and Smart Devices, ISESD 2019*, pp. 1–5, 2019, doi: 10.1109/ISESD.2019.8909515.

[11] Y. Lei, S. L. Phung, A. Bouzerdoum, H. Thanh Le, and K. Luu, "Pedestrian Lane Detection for Assistive Navigation of Vision-Impaired People: Survey and Experimental Evaluation," *IEEE Access*, vol. 10, pp. 101071–101089, 2022, doi: 10.1109/ACCESS.2022.3208128.

[12] R. A. Ramadani, I. K. G. D. Putra, M. Sudarma, and I. A. D. Giriantari, "A new technology on translating Indonesian spoken language into Indonesian sign language system," *International Journal of Electrical and Computer Engineering*, vol. 11, no. 4, pp. 3338–3346, 2021, doi: 10.11591/ijece.v11i4.pp3338-3346.

[13] D. Kumar and N. Muhammad, "Object Detection in Adverse Weather for Autonomous Driving through Data Merging and YOLOv8," *Sensors (Basel)*, vol. 23, no. 20, Oct. 2023, doi: 10.3390/s23208471.

[14] U. Fadlilah, A. K. Mahamad, and B. Handaga, "The Development of Android for Indonesian Sign Language Using Tensorflow Lite and CNN: An Initial Study," *J Phys Conf Ser*, vol. 1858, no. 1, 2021, doi: 10.1088/1742-6596/1858/1/012085.

[15] A. Aboah, B. Wang, U. Bagci, and Y. Adu-Gyamfi, "Real-time Multi-Class Helmet Violation Detection Using Few-Shot Data Sampling Technique and YOLOv8," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, vol. 2023-June, pp. 5350–5358, 2023, doi: 10.1109/CVPRW59228.2023.00564.

[16] Q. Zhou, J. Qin, X. Xiang, Y. Tan, and N. N. Xiong, "Algorithm of helmet wearing detection based on AT-YOLO deep mode," *Computers, Materials and Continua*, vol. 69, no. 1, pp. 159–174, 2021, doi: 10.32604/cmc.2021.017480.

[17] I. H. Hsieh, H. C. Cheng, H. H. Ke, H. C. Chen, and W. J. Wang, "A CNN-based wearable assistive system for visually impaired people walking outdoors," *Applied Sciences (Switzerland)*, vol. 11, no. 21, Nov. 2021, doi: 10.3390/app112110026.

[18] F. Wahab, I. Ullah, A. Shah, R. A. Khan, A. Choi, and M. S. Anwar, "Design and implementation of a real-time object detection system based on the single-shoot detector and OpenCV," *Front Psychol*, vol. 13, Nov. 2022, doi: 10.3389/fpsyg.2022.1039645.

[19] K. Liu, "STBi-YOLO: A Real-Time Object Detection Method for Lung Nodule Recognition," *IEEE Access*, vol. 10, pp. 75385–75394, 2022, doi: 10.1109/ACCESS.2022.3192034.

[20] N. E. Albayrak, "Object Recognition using TensorFlow," pp. 1–1, 2021, doi: 10.1109/isec49744.2020.9397835.

[21] C. J. Lin, S. Y. Jeng, and H. W. Lioa, "A Real-Time Vehicle Counting, Speed Estimation, and Classification System Based on Virtual Detection Zone and YOLO," *Math Probl Eng*, vol. 2021, 2021, doi: 10.1155/2021/1577614.

[22] D. S. Bacea and F. Oniga, "Single stage architecture for improved accuracy real-time object detection on mobile devices," *Image Vis Comput*, vol. 130, Feb. 2023, doi: 10.1016/j.imavis.2022.104613.

[23] D. Bhavesh, R. Patel, S. A. Goswami, P. S. Kapatel, and Y. M. Dhakad, "Realtime Object's Size Measurement from Distance using OpenCV and LiDAR," 2021.

[24] S. Gamal, "The Memory of Knowledge: An Analytical Study on Translators' Perceptions and Assessment of CAT Tools with Regard to Text Genre," *International Journal of Linguistics and Translation Studies*, vol. 1, no. (2), pp. 1–18.

[25] C. Vieira, L. N., O'Hagan, M., & O'Sullivan, "Understanding the societal impacts of machine translation: a critical review of the literature on medical and legal use cases," *Inf Commun Soc*, vol. 24, no. (11), pp. 1515–1532, 2021.