Journal of Electrical Engineering and Computer (JEECOM)
Vol. 7, No. 1 (2025), DOI: 10.33650/jeecom.v4i2
p-ISSN: 2715-0410; e-ISSN: 2715-6427

Implementation of YOLOv7 Model for Human Detection in Difficult Conditions

Arijal B [1], Andi Sunyoto [2], M. Hanafi[3]
[1, 2] Magister Informatika, Universitas Amikom Yogyakarta, Sleman, Indonesia
[3] Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta, Sleman, Indonesia

Article Info

Article history:

Received January 20, 2025
Revision February 9, 2025
Published April 10, 2025

Keywords:

Objek Detection
Yolov7
Human Detection
Systematic Literature Review
Confusion Matrix

ABSTRACT

The rapid development of artificial intelligence technology in recent decades has led to the development of highly efficient object detection algorithms, including human detection under difficult conditions. Human detection is one of the major challenges in computer vision as it involves various complex factors such as obstructed human objects, pose variations, small low-resolution human objects, as well as the presence of fake human objects such as statues or images. The purpose of this study is to determine the detection accuracy of three types of YOLOv7 algorithm models in detecting humans in difficult conditions. This research uses the SLR (Systematic Literature Review) method to determine the algorithm used, namely YOLOv7 and selects three types of YOLOv7 models namely YOLOv7x.pt, YOLOv7-w6-person.pt, and YOLOv7-w6-pose.pt. These models are selected based on their advantages. These models were selected based on their excellence in detecting human objects and their relevance for complex scenarios. Tests were conducted using 100 images obtained from the internet and divided into four categories of human objects under difficult conditions, which represent various challenges in human detection. Analysis was performed using convusion matrix to evaluate performance metrics such as accuracy, precision, recall, and F1-score. Based on the test results, the YOLOv7-w6-person.pt model showed the best overall performance, especially in detecting humans in obstructed conditions and complex lighting with a precision of 90.4%, Recall 88.7%, and F1-Score 89.5%. This model has higher accuracy, precision, and F1-score than the other models, making it a reliable choice for human detection in difficult conditions. These findings not only demonstrate the relevance of YOLOv7 as a reliable human detection algorithm, but also provide a basis for further optimization of YOLOv7-based human detection systems, both through model architecture enhancement and more specific dataset adaptation. This research makes an important contribution to the development of human detection technologies for real-world applications, such as surveillance, crowd analysis, and automated safety systems.

Corresponding Author:
Arijal B
Universitas Amikom Yogyakarta, Jl. Ring Road Utara, Sleman and 55281, Indonesia
Email: arijalputrabading@gmail.com

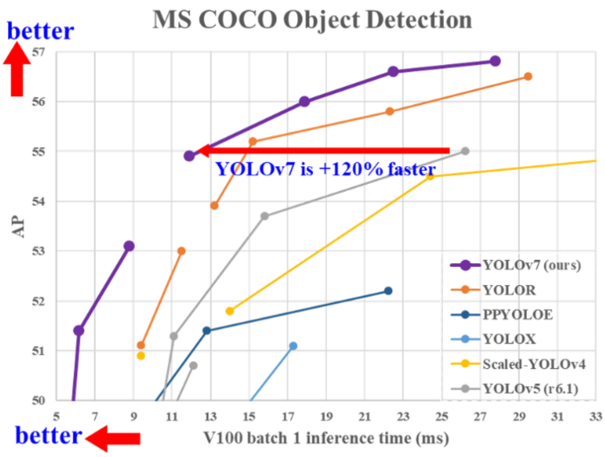
1. INTRODUCTION

The rapid advancement of technology in recent decades has encouraged the development of more complex technologies. One of the fields that has followed this stream of development is Artificial Intelligence technology [1]. With the advent of Artificial Intelligence, computer vision has become an integral component in AI-related research. And with the technological advancements in this field, we can create powerful artificial neural networks for various computer vision tasks, such as identification & localization, classification, tracking, pose estimation, motion analysis, etc. Today, it has been applied in various fields, ranging from security & surveillance to autonomous cars, from biometrics to forensics, from human detection to augmented reality [2].

Among the most interesting applications of computer vision is human detection. Human detection is an important topic to study because there are challenges that need to be solved, such as the challenge of detecting obstructed human objects, variations in human poses, light conditions, cluttered backgrounds, variations in viewing angles, and low resolution [3].

Research in the field of human detection in difficult conditions has been carried out using various existing methods, such as [4] conducting research on human object detection from obstructed Infrared images using the YOLOv4 algorithm. [5] detects pedestrians who are obstructed by the DCPDN method. Research [6] detects small people in large-scale images. Research [7] overcomes the problem of human detection that is difficult to see. Research [8] provides a comprehensive analysis of the evolution of the YOLO architecture. Research [9] compared the accuracy of YOLOv5 and YOLOv7. Research [10] reviews and evaluates the Backbone Network.

Based on the description above, there has been a lot of research on human object detection in difficult conditions with different methods, both optimization of existing algorithms and, combining algorithms, and adjusting dataset alignment in certain cases and the results presented with good accuracy values and recommending further research to use the latest and more sophisticated technology for detection results with better accuracy. Currently in the field of object detection one of the new and advanced algorithms is YOLO, here is a comparison of several versions of YOLO, in Figure 1.



Comparison of YOLOv7 with other object detectors [11]

It can be seen that YOLOv7 is better than other YOLO versions in detecting objects in general consisting of many classes of objects, including human objects. Then from the results of the literature study conducted, it was found that the YOLOv7 algorithm has 22 different types of models.

From the results of the review conducted, there is no research that specifically analyzes the implementation of 22 types of YOLOv7 algorithm models to detect human objects in difficult conditions. So the researcher will implement and test the YOLOv7 model to detect humans in various difficult conditions. The tested model consists of three models selected according to the relevance to the case study studied, namely the YOLOv7x.pt model, the YOLOv7-w6-person.pt model, and the YOLOv7-w6-pose.pt model. The evaluation was conducted to determine the extent to which these models are able to overcome the challenges of human detection under difficult conditions.

2. METHODS

2.1. *Systematic Literature Review* (SLR) Method

This research was conducted using the systematic literature review or SLR method. SLR is a process for identifying, assessing, and interpreting facts and evidence from available research with the aim of finding answers to a specific research question [12]. The SLR is used to review previous research on human detection in difficult conditions :

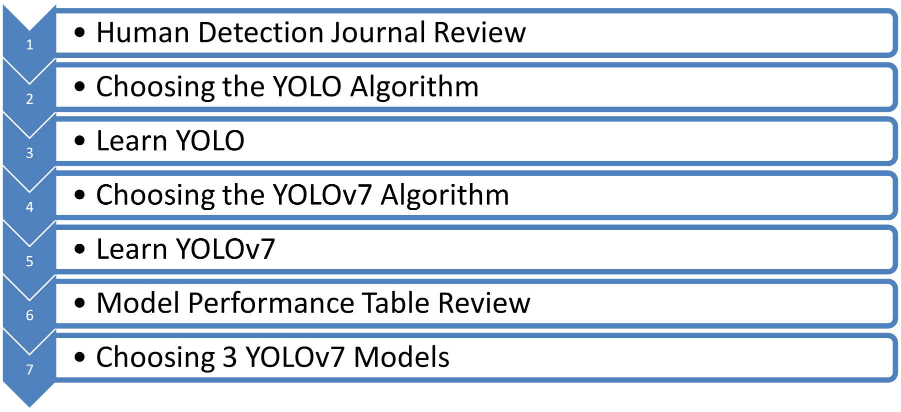


Diagram 1. Schematic of *Systematic Literature Review* (SLR) Method

2.2. Research Question

To conduct research using the *Systematic Literature Review* (SLR) method, the first step is to formulate a set of research questions (RQs). The questions listed in Table 1 are crucial in establishing a clear, structured and focused research framework. This systematic approach is designed to improve the quality and efficiency of the research process, allowing researchers to work with better focus and more optimized results.

Table 1. *Research Question*

TABLE I. PROBLEM FORMULATION

ID	Problem Formulation
RQ1	What are the challenges in human object detection?
RQ2	What is the ability of the YOLOv7 algorithm for human object detection in difficult conditions?
RQ3	What are the findings resulting from testing multiple YOLOv7 Models?

2.3. Research Stages

The following stages are carried out in testing the YOLOv7 algorithm for human object detection in difficult conditions.

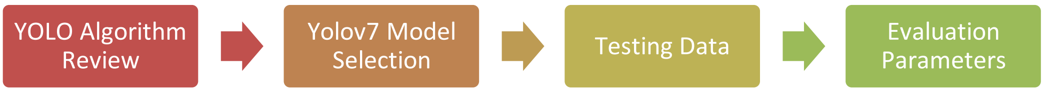


Diagram 2. Research Stages

The research stages are divided into four stages as in the diagram above. each stage can be seen the process as follows:

2.3.1. Algoritma YOLO Review

Reviewed the YOLOv7 journal written by Chien-Yao Wang, Alexey Bochkovskiy, and Hong-Yuan Mark Liao as the developers of YOLOv7 published on July 6, 2022 and introduced YOLOv7, which surpasses all known object detectors in terms of speed and accuracy in the range of 5 FPS to 160 FPS. YOLOv7 achieved the highest accuracy of 56.8% AP among all known real-time object detectors with 30 FPS or higher on the V100 GPU.

2.3.2. YOLOv7 Model Selection

YOLOv7 is a relatively advanced object detector at present and has shown excellent results. Based on the literature review, YOLOv7 was chosen because of its advantages over other YOLO versions. The following is the architecture of YOLOv7 [13].

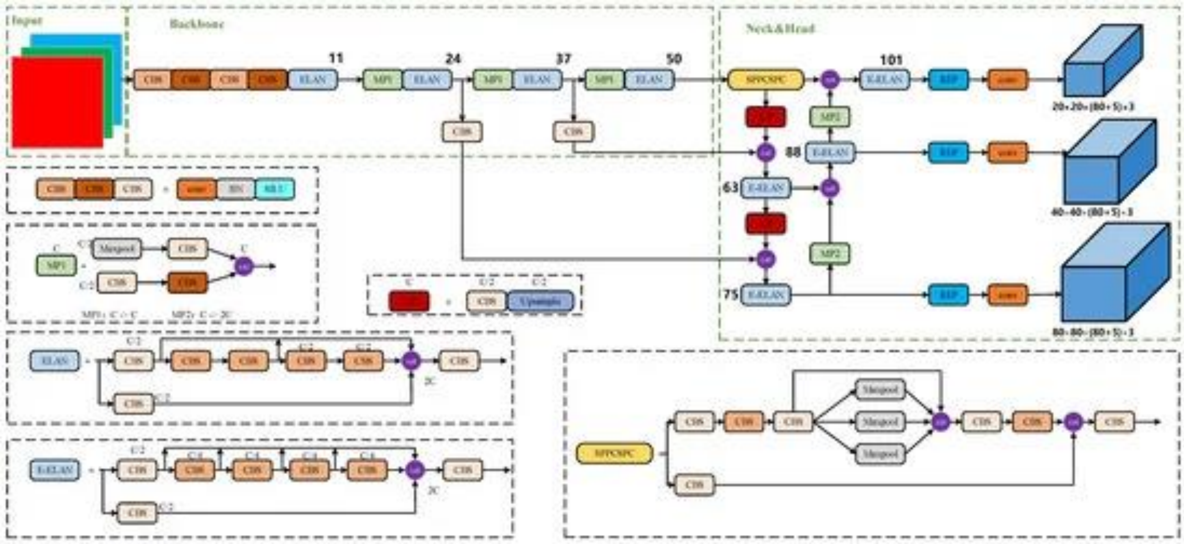


Figure 2. YOLOv7 algorithm architecture [13]

The YOLOv7 algorithm has several variants designed for specific needs such as model size, speed, and detection type. The following table shows the performance of YOLOv7 model variants.

Table 2. Model Variant Performance

Model	Test Size	AP^{test}	AP_{50}^{test}	AP_{75}^{test}	batch 1 fps	batch 32 average time
YOLOv7	640	**51.4%**	**69.7%**	**55.9%**	161 *fps*	2.8 *ms*
YOLOv7-X	640	**53.1%**	**71.2%**	**57.8%**	114 *fps*	4.3 *ms*
YOLOv7-W6	1280	**54.9%**	**72.6%**	**60.1%**	84 *fps*	7.6 *ms*
YOLOv7-E6	1280	**56.0%**	**73.5%**	**61.2%**	56 *fps*	12.3 *ms*
YOLOv7-D6	1280	**56.6%**	**74.0%**	**61.8%**	44 *fps*	15.0 *ms*
YOLOv7-E6E	1280	**56.8%**	**74.4%**	**62.1%**	36 *fps*	18.7 *ms*

Based on the results of the literature study YOLOv7 has 22 models and three models were selected for testing, namely:

a. YOLOv7x.pt (Large model with high accuracy)
b. YOLOv7-w6-person.pt (Special model for human detection)
c. YOLOv7-w6-pose.pt (Special model for detecting human poses)

2.3.3. Testing Data

The testing data used contains 100 images of human objects in difficult conditions taken from various sources on the internet that include:

a. Human Object Obstructed
b. Pose Variations
c. small human object
d. Fake Human Object (statue or picture)

2.3.4. Evaluation Parameters

The three selected models will be tested using a dataset with a difficult condition category with image size parameters of 640 x 640, Confidence threshold 0.4 and the test results are analyzed using confusion matrix namely accuracy, precision, recall and f1-score.

Figur 3. Yolov7x.pt Model Testing Result Sample

Figur 4. Sample Model Testing Results YOLOv7-w6-person.pt

From the sample above, in the image there are 11 human objects with various poses, obstructions, and small human objects. The resulting Yolov7x.pt model detects 7 human objects, while the YOLOv7-w6-person.pt model detects 8 human objects.

3. RESULT AND DISCUSSION

From the testing process carried out, the model performance is based on the following categories.

Table 3. Model Performance by Category

Kategori	YOLOv7x.pt	YOLOv7-w6-person.pt	YOLOv7-w6-pose.pt
Human object obstructed	85.3%	91.2%	87.5%
Variations of poses	82.1%	89.4%	86.0%
Low-resolution small human object	78.6%	87.8%	84.9%
False human object	84.2%	92.3%	89.1%

The following table shows the Confusion Matrix Comparison of precision, recall, and F1-score for each category of difficult conditions in the three models.

Table 4. Cofusion Matrix of Each Model

Model	TP	FP	FN	Precision	Recall	F1-Score
YOLOv7x.pt	210	40	50	84.0%	80.8%	82.3%
YOLOv7-w6-person.pt	235	25	30	90.4%	88.7%	89.5%
YOLOv7-w6-pose.pt	220	35	40	86.3%	84.6%	85.4%

4. CONCLUSIONS

This research shows that the YOLOv7-w6-person.pt model yields the best overall performance, especially in detecting humans in obstructed and complex lighting conditions with 90.4% precision, 88.7% Recall, and 89.5% F1-Score. This model has higher accuracy, precision, and F1-score than the other models, making it a reliable choice for human detection in difficult conditions. These findings not only demonstrate the relevance of YOLOv7 as a reliable human detection algorithm, but also provide a basis for further optimization of YOLOv7-based human detection systems, both through model architecture enhancement and more specific dataset adaptation. This research makes an important contribution to the development of human detection technologies for real-world applications, such as surveillance, crowd analysis, and automated safety systems.

ACKNOWLEDGMENTS

On behalf of the authors, we would like to thank our supervisors, Dr. Andi Sunyoto, M.Kom. and M. Hanafi, M.Eng., Ph.D. for their guidance during this research, providing advice, opinions, and motivating us throughout this study process.

REFERENCES

[1] Khairunnas, K., Yuniarno, E. M., & Zaini, A. (2021). Making Human Object Detection Module Using YOLO Method for Mobile Robot. ITS Engineering Journal, 10(1), A50-A55.

[2] Shrestha, A., Yeophantong, T., & Nongpong, K. (2019, April). An Approach for Handling Partially Visible Human Shapes in People Detection Systems. In 2019 6th International Conference on Control, Decision and Information Technologies (CoDIT) (pp. 1350-1354). IEEE.

[3] Gajjar, V., Khandhediya, Y., Gurnani, A., Mavani, V., & Raval, M. S. (2018). In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 1908-1916).

[4] Du, S., Zhang, B., Zhang, P., Xiang, P., & Xue, H. (2021). FA-YOLO: An improved YOLO model for infrared occlusion object detection under confusing background. Wireless Communications and Mobile Computing, 2021, 1-10.

[5] Chen, Y., Wang, X., & Ouyang, Y. (2023). DCPDN: High-Performance Pedestrian Detection Networks for Difficult Conditions. IEEE Access.

[6] Yu, X., Gong, Y., Jiang, N., Ye, Q., & Han, Z. (2020). Scale match for tiny person detection. In Proceedings of the IEEE/CVF winter conference on applications of computer vision (pp. 1257-1265).

[7] Kataoka, H., Ohki, S., Iwata, K., & Satoh, Y. (2018, August). Occlusion Handling Human Detection with Refocused Images. In 2018 24th International Conference on Pattern Recognition (ICPR) (pp. 1701-1706). IEEE.

[8] Terven, J., & Cordova-Esparza, D. (2023). A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. arXiv preprint arXiv:2304.00501.

[9] Anyim, A. C. (2023). A Comparative Study of YOLOv5 and YOLOv7 Modifications for Face Detection on a Custom Dataset. Center for Postgraduate Studies, University of the Balearic Islands.

[10] Elharrouss, O., Akbari, Y., Almaadeed, N., & Al-Maadeed, S. (2022). Backbones-review: Feature extraction networks for deep learning and deep reinforcement learning approaches. arXiv preprint arXiv:2206.08016.

[11] Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 7464-7475).

[12] Diponegoro, M. H., Kusumawardani, S. S., & Hidayah, I. (2021). Systematic literature review: implementation of deep learning methods in student performance prediction. National Journal of Electrical Engineering and Information Technology, 10(2), 131-138. [23] S. Leonelli and N. Tempini, *Data Journeys in the Sciences*. 2020.

[13] Feng, G., Yang, Q., Tang, C., Liu, Y., Wu, X., & Wu, W. (2024). Mask-Wearing Detection in Complex Environments Based on Improved YOLOv7. Applied Sciences, 14(9), 3606. [18] C. Shang and F. You, "Data Analytics and Machine Learning for Smart Process Manufacturing: Recent Advances and Perspectives in the Big Data Era," *Engineering*, vol. 5, no. 6, pp. 1010-1016, 2019, doi: 10.1016/j.eng.2019.01.019.

[14] Mulyo, H., & Kusumodestoni, H. (2022). Object Detection on CCTV for Smart City Kendal Regency. AMRI National Journal of Information Engineering Method Analysis, 1(2), 121-124.

[15] Aljabri, M., AlAmir, M., AlGhamdi, M., Abdel-Mottaleb, M., & Collado-Mesa, F. (2022). Towards a better understanding of annotation tools for medical imaging: a survey. Multimedia tools and applications, 81(18), 25877-25911.

[16] Shrestha, A., Yeophantong, T., & Nongpong, K. (2019, April). An Approach for Handling Partially Visible Human Shapes in People Detection Systems. In 2019 6th International Conference on Control, Decision and Information Technologies (CoDIT) (pp. 1350-1354). IEEE.

[17] Pradana, R. L., Khumaidi, A., & Andiana, R. (2022). Identification of Defect Causes in Welding Results with Image Processing Using the YOLO Method. J. Tech. Electro and Comput. Triac, 9(3), 1-5.

[19] Zhou, J., Zhang, Y., & Wang, J. (2023). RDE-YOLOv7: an improved model based on YOLOv7 for better performance in detecting dragon fruits. Agronomy, 13(4), 1042.

[20] Wahyuni, S., & Sulaeman, M. (2022). Application of deep learning algorithm for face detection attendance system at PT Karya Komponen Presisi. SIMANTIK Informatics Journal, 7(1), 12-21.

[21] Zhang, Z., Lu, X., Cao, G., Yang, Y., Jiao, L., & Liu, F. (2021). ViT-YOLO: Transformer-based YOLO for object detection. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 2799-2808).

[22] Zahari, I., Pratama, Z., Mahmud, W., & Wibowo, D. A. (2022). Computer Answer Sheet Checking System with Optical Mark Recognition (OMR) Based on Open Computer Vision Python. In National Seminar on Innovation and Development of Applied Technology (SENOVTEK) (Vol. 1, pp. 9-16).

[24] Hero, M. R., Setyanto, A., & Arief, M. R. (2024). A Comprehensive Review of Clasifier used with Imbalanced Data in Machine Learning. Journal of Electrical Engineering and Computer (JEECOM), 6(1), 177-185. [25] Q. Song, H. Ge, J. Caverlee, and X. Hu, "Tensor completion algorithms in big data analytics," *arXiv*, vol. 13, no. 1, 2017.

[25] Cholissodin, I., Sutrisno, S., Soebroto, A. A., Hasanah, U., & Febiola, Y. I. (2020). AI, Machine Learning & Deep Learning. Faculty of Computer Science, Brawijaya University, Malan

[26] Ertel, W. (2018). Introduction to artificial intelligence. Springer.