ISSN No:-2456-2165

Systematic Literature Review of Pedestrian Detection using the YOLO Algorithm

Lamsadi[1 *], Arief Setyanto[2], Tonny Hidayat[3]
[1]Informatics Engineering Study, Amikom University of Yogyakarta
[2]Informatics Engineering Study, Amikom University of Yogyakarta
[3]Informatics Engineering Study, Amikom University of Yogyakarta

Abstract:- Technology is developing so rapidly at this time. Every time various latest and cutting-edge technologies in various fields transmit life. One of them is in the field of object detection. As technology develops, the need for object detection systems becomes very strong. Object detection or object detection is the lifeblood of Computer Vision and Image Processing. There are 4 main focuses in Computer Vision, namely Recognition, Visual Tracking (visual tracking), Semantic Segmentation (semantic segmentation) and Image Restoration (image restoration). To be able to do these four things, we need an algorithm that can effectively be applied to detect objects, especially pedestrians, so YOLO was chosen as the answer. YOLO is one of several algorithms that are often used in Machine Learning. You Only Live Once or better known as YOLO is a very well-known and widely used algorithm. YOLO is a specific algorithm for object detection. In recent years, the YOLO Algorithm has shown interesting results in various areas of object detection, both large-scale and special, has solved many problems in the field of object detection in general, the detection of license plates of vehicles, pedestrians, etc. Through this systematic literature review, it is hoped that it will be able to provide enlightenment for the development of Object Detection science.

Keywords:- *Object Detection; Image Processing; YOLO; Pedestrian; Machine Learning; Systematic Literature Review.*

I. INTRODUCTION

Technology is developing so rapidly at this time. Every time, various new and cutting-edge technologies emerge in various fields and aspects of life. One of them is in the field of object detection. Object detection is the lifeblood of Computer Vision and Image Processing [1]. There are 4 main focuses in Computer Vision, namely recognition, visual tracking, semantic segmentation and image restoration [2]. To be able to count objects automatically, the first two things that must be done are to detect and classify objects (movable or immovable), for example vehicles, pedestrians and others [20]. You Only Live Once or better known as YOLO is a very well-known and widely used algorithm. YOLO is a specific algorithm for detecting objects [10]. In recent years, the YOLO Algorithm has shown interesting results in various areas in object detection[12], both large-scale and special, has solved many problems in the field of object detection in general, vehicle registration plate detection, pedestrians and others [9]. Even now YOLO (Object detection) can be used to control production processes in factories based on video data in real time [16]. In several studies YOLO is not only used to detect objects in the form of humans. Yolo can also be used to detect fish movements in the water [18].

II. RESEARCH METHODS

SLR is a method or approach in research, namely by reviewing some of the literature by reviewing a particular topic and emphasizing a focused question. Then the questions are selected, identified, assessed and concluded with predetermined criteria. In addition, SLR also aims to find research gaps so that new research areas emerge that have the opportunity to be studied.

A. Research Questions

To conduct SLR research, several steps or criteria are needed, called PICOC. PICOC stands for Population, Intervention, Comparison, Outcomes and Context. The following is a PICOC summary table.

Population	Image processing, machine learning
Intervention	Datasets, Models and Methods
Comparasion	Accuracy, precision, recall, performance and speed
Outcomes	Model detection accuracy and performance efficiency
Context	Pedestrian and YOLO

Table 1:- PICOC Summary

To complete research using the SLR method, several questions or Research Questions (RQ) are needed as below so that the research is more focused, directed and conical.

Table 2:- Research questions

ID	Research Questions	Purpose
RQ1	What is the research on object detection especially on pedestrians like in the last 6 years?	Identification of research developments on object detection, especially on pedestrians.
RQ2	Where does the research source used as a reference for object detection especially pedestrians using YOLO come from?	Identification of research sources as references.
RQ3	From which countries did the researchers raise research on object detection, especially on pedestrians?	Identify countries that have done a lot of research on object detection, especially on pedestrians.
RQ4	What and how is the object called the pedestrian?	Identify the pedestrian concept
RQ5	What and how do pedestrian objects becomes vital objects?	Identify pedestrian safety systems
RQ6	What and how does the YOLO Algorithm work?	Identify the concepts and how the Yolo Algorithm works.

B. Study Selection

In conducting study selection, the research included in this SLR is research published within the last 6 years in the form of journals or conferences. There are 3 main keywords used in this study, namely object detection, pedestrian and YOLO. In general, there are two categories of research taken as literature, namely experimental and survey research. The stages of the literature search process to find the right one are shown in Figure 1.

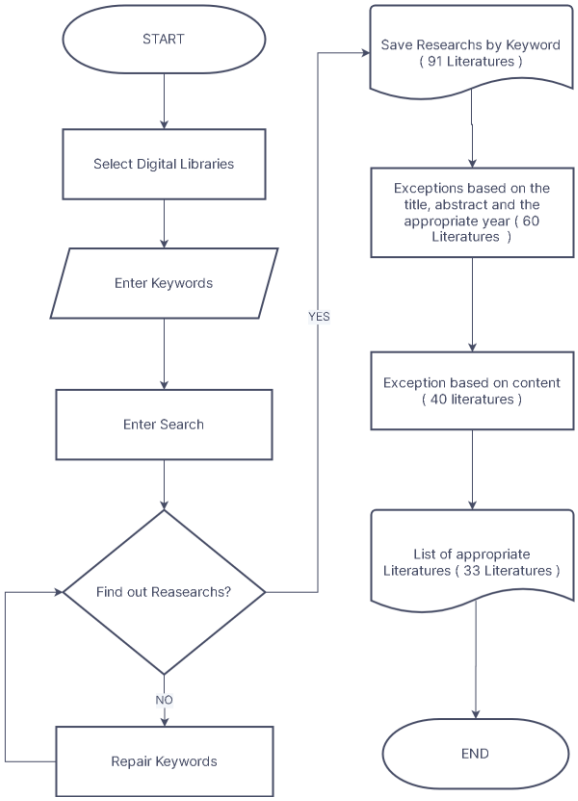

Fig 1:- Research Search Flowchart

III. RESULTS AND DISCUSSION

The following is a discussion of the literature that was collected.

A. Research Years

In the last 6 years, research in the field of object detection, especially those using the YOLO algorithm, seems to have progressed quite rapidly, although there has been a slight decline in 2021, this does not mean that this research is less interesting. The most research is in 2022 and this indicates that as more years are added it is likely to increase. The following is a graphical image of the distribution with the percentage of research years.

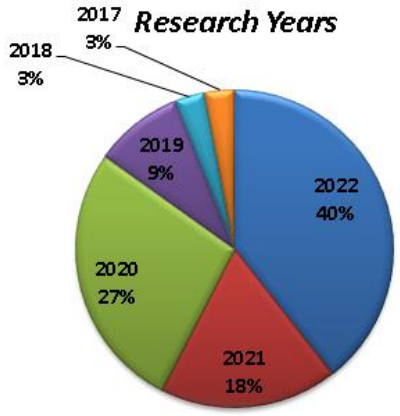

Fig 2:- Research Years

B. Research Sources

The databases used for research were ScienceDirect (sciencedirect.com), IEEE Xplore (ieeexplore.ieee.org), Springer (springerlink.com), IET Search (ietresearch.onlinelibrary.wiley.com), SPIE (https://www.spiedigitallibrary.org), MDPI (https://www.mdpi.com) and Software Impact (www.softwareimpacts.com). Details of the distribution of sources are explained in Figure 3 below:

ISSN No:-2456-2165

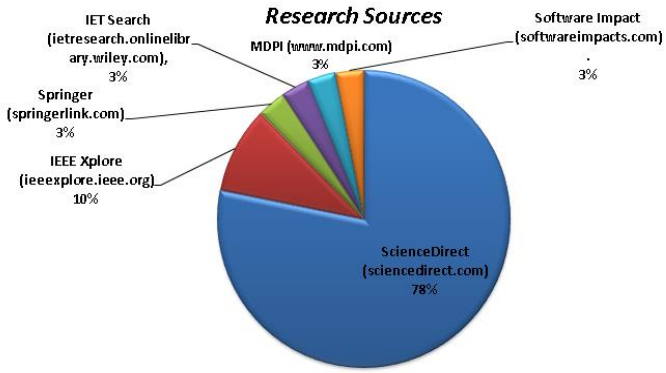

Fig 3:- Research Sources

As can be seen in the ScienceDirect graph, it occupies the majority or the main choice as a research source. This is due to ScienceDirect's excellent reputation as a provider of world-class journals.

C. Countries

Countries that are so intensively conducting research in the field of object detection, especially with the Yolo Algorithm, are still dominated by two countries, namely China and Egypt, even though the percentage distance between these two countries is quite far. This is as shown in Figure 4 below:

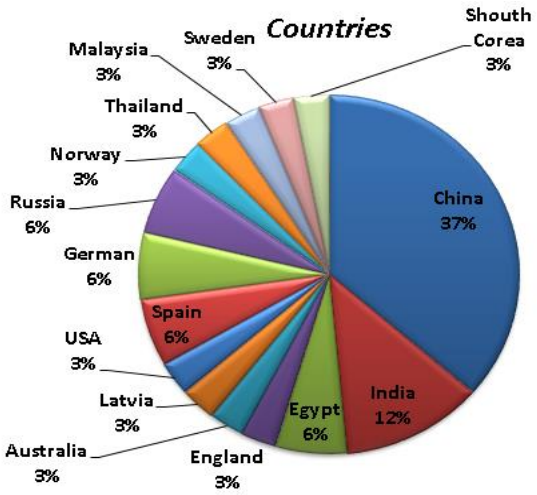

Fig 4:- Countries

D. Object Types

Research on object detection, especially pedestrians, is very important. This is due to the increasing number of devices or technologies that use this system. An example is autonomous vehicles, which are increasingly developing day by day because more and more are doing research [29]. Not only that, integration between object detection and other branches of science can support the health of pedestrians with air control systems on the road [30]. Pedestrian detection is the basis of many human-based tasks, including speed tracking, detection of pedestrian movement, automation of pedestrian recognition and appropriate response actions or reject detection of pseudo pedestrians [8].

In general, pedestrians are divided into 2 (two), namely real or real people (actual person) and pseudo people (the depiction of a person). Depictions of people can be in the form of pictures of people, statues or dolls and so on [3]. In other cases pedestrians are also called Passengers (Passengers). This can be seen from the location where they are located, for example at a bus terminal or train station (Sipetas et al., 2020). Detection of the movement of objects (people) or the Euclidian distance between objects and their surroundings can be important information in distinguishing between real and imaginary people [24].

Even though there are lots of traffic indicators, crossroads and pedestrian safety signs, the possibility of accidents between vehicles (cars, motorcycles, trains and others) and pedestrians is still very high. Therefore, the development of advanced cognitive systems (eg, pedestrian detection) is a promising step towards a rapid reduction in the number of traffic accidents. Recently, the development of a pedestrian safety system has received a lot of positive response and attention. However, pedestrian safety is quite a difficult job due to reasons such as illumination and appearance effects (texture, ratio, area)[8]. Therefore we need a system that can distinguish between pedestrians and the environment, such as ITMSs and SMRIK. ITMSs or Intelligent Traffic Management Systems is a system that not only uses object detection and image processing approaches but also Geometric Computing [21]. SMRIK is a Pedestrian Automatic Emergency Braking (PAEB)-based machine learning [28].

E. Yolo Algorithm

The YOLO algorithm is a method based on regression that predicts through bounding boxes and class objects to determine the location of an object in an image using a Single Neural Network. (Feng et al., 2019; Yu, J., & Choi, H. 2022; Saada et al., 2022). The YOLO algorithm works by dividing the image (image) into several parts (cells), each cell is used to predict a number of bounding boxes if there is more than one object in the image. (Feng et al., 2019; Xue et al., 2021; Han et al., 2021). Then the prediction results will be collected and the bounding box with the smallest probability will be removed. The bounding box with the largest predicted probability value will be the final result. (Han et al., 2021). The following is an example of the bounding box and class object shown in Figure .5:

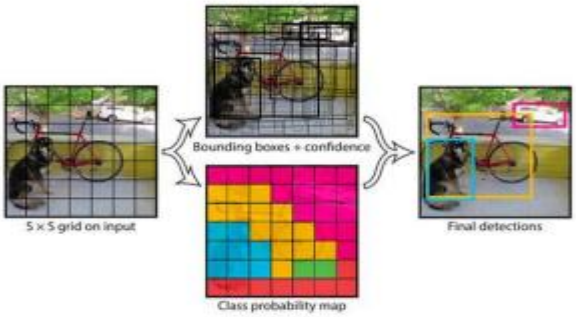

Fig 5. bounding box and class object
Source (Han et al., Procedia Computer Science, volume : 183, page : 63 https://doi.org/10.1016/j.procs.2021.02.031)

Yolo is capable of detecting both 2D and 3D objects in [27] the form of images (static images) or videos [19]. Yolo is able to detect objects in real time[5]. At a time when COVID-19 cases were still rampant, Yolo was used to detect social distancing and [22] the use of face masks and was found to have an accuracy rate of up to 90% in 150 facial samples tested [7]. Yolo can process images up to 45 frames per second (FPS) [4].

Several years ago, experts have published several YOLO versions such as YOLO V2, YOLO V3, YOLO V4, YOLO V5 and There are several limited-revision versions, such as YOLO-LITE [10]. Yolov3 is capable of accurately detecting up to 79% of pedestrians out of 20,000 detected objects [11]. The latest generation of the YOLO Algorithm is YOLOv5 and uses the Python programming language unlike its predecessors which still use the C [28][13]. However, in terms of access speed and detection with an accuracy that is not inferior to YOLO in general, Tiny Yolo still dominates [26].

IV. CONCLUSION

This SLR research aims to identify and analyze pedestrian types based on the nature of the object and its location. In general, pedestrians are divided into 2 (two), namely real or real people (actual person) and pseudo people (the depiction of a person). Pedestrians are also called passengers. By identifying and analyzing pedestrians, they will be able to maximize the performance of the YOLO Algorithm. The YOLO algorithm is a method based on regression that predicts through bounding boxes and class objects to determine the location of an object in an image using a Single Neural Network. Of the 33 literatures reviewed, 78% or the majority of research sources were from sciencedirect.com. China is the country with the highest number of studies, reaching 37%, followed by India, 12% and the others below 10%. The year of research from the literature reviewed came from 2022 as much as 40% and 2020 as much as 27%.

In the end object detection based on the YOLO algorithm becomes very important and needed when it comes to the safety of pedestrians, especially on roads or around public transportation roads such as trains. The implementation of an automatic emergency braking system such as SMRIK provides an additional sense of security for pedestrians [28].

REFERENCES

[1]. Arulprakash, E., & Aruldoss, M, "A Study On Generic Object Detection With Emphasis On Future Research Directions", *Journal of King Saud University – Computer and Information Sciences*, 34(9):7347–7365,https://doi.org/10.1016/j.jksuci.2021.08.001, 2022.

[2]. Chai, J., Zeng, H., Li, A., & Ngai, Eric, E.W.T, "Deep Learning in Computer Vision: A Critical Review of Emerging Techniques and Application Scenarios", *Machine Learning with Application*, 6, 10013. https://doi.org/10.1016/j.mlwa.2021.100134 , 2021.

[3]. Donoso, F.G., Amoros, J.C., Escalona, F., & Cazorla, M, "Three-dimensional reconstruction using SFM for actual pedestrian classification", *Expert Systems With Applications*, 213, 119006. https://doi.org/10.1016/j.eswa.2022.119006, 2022.

[4]. Feng, X., Jiang, Y., Yang, X., Du, M., & Xin Li, X, "Computer vision algorithms and hardware implementations: A survey", *Integration, the VLSI Journal*, 69: 309–320. https://doi.org/10.1016/j.vlsi.2019.07.005 China, 2019.

[5]. Hana, Z., Huanga, H., Fan, Q., Li, Y., Li, Y., & Chen, X, "SMD-YOLO: An Efficient And Light Weight Detection Method For Mask Wearing Status During The COVID-19 Pandemic", *Computer Methods and Programs in Biomedicine*, 221, 106888. https://doi.org/10.1016/j.cmpb.2022.106888, 2020.

[6]. Han, X., Chang, J., & Wang, K, "Real-Time Object Detection Based On YOLO-V2 For Tiny Vehicle Object", *10th International Conference of Information and Communication Technology (ICICT-2020), Procedia Computer Science*, 183: 61–72. https://doi.org/10.1016/j.procs.2021.02.031, 2022.

[7]. Hou, Y. C., Baharuddin, M. Z., Yussof, S. & Dzulkifly, S, "Social Distancing Detection with Deep Learning Model", *2020 8th International Conference on Information Technology and Multimedia (ICIMU):* 334-338. https://doi.org/10.1109/ICIMU49871.2020.9243478, 2020.

[8]. Iftikhar, S., Asim, M., Zhang, Z., & El-Latif, A. A. A, "Advance Generalization Technique Through 3D CNN To Overcome The False Positives Pedestrian In Autonomous Vehicles", *Telecommunication Systems*, volume 80: pages 545–557. https://doi.org/10.1007/s11235-022-00930-1, 2020.

[9]. Jamtsho, Y., Riyamongkol, P., & Waranusast, R, "Real-Time Bhutanese License Plate Localization Using YOLO", *The Korean Institute of Communications and Information Sciences (KICS), ICT Express*, 6: 121–124. https://doi.org/10.1016/j.icte.2019.11.001, 2020.

[10]. Jiang, P., Ergu, D., Liu, F., Cai, Y., & Ma, B, "A Review of Yolo Algorithm Developments", *The 8th International Conference on Information Technology and Quantitative Managemen, Procedia Computer Science*, 199: 1066–1073. https://doi.org/10.1016/j.procs.2022.01.135, 2022.

[11]. Kataev, G., Varkentin, V., & Nikolskaia, K, "Method To Estimate Pedestrian Traffic Using Convolutional Neural Network", XIV International Conference 2020 SPbGASU "Organization and safety of traffic in large cities", *Transportation Research Procedia*, 50: 234–241. https://doi.org/10.1016/j.trpro.2020.10.029, 2020.

[12]. Kumar, A., Kalia, A., & Kalia, A, "ETL-YOLO v4: A Face Mask Detection Algorithm In Era Of COVID-19 Pandemic", *Optik - International Journal for Light*

and Electron Optics, 259, 169051. https://doi.org/10.1016/j.ijleo.2022.169051, 2022.
[13]. Li, P., & Zhao, W, "Image Fire Detection Algorithms Based On Convolutional Neural Networks", *Case Studies in Thermal Engineering*, 19, 100625. https://doi.org/10.1016/j.csite.2020.100625, 2020.
[14]. Li, S., Li, Y., Li, Y., Li, M., & Xu, X, "YOLO-FIRI : Improved YOLOv5 for Infrared Image Object Detection*", IEEE Access*, VOLUME 9: 141861-141875. https://doi.org/10.1109/ACCESS.2021.3120870, 2021.
[15]. Loey, M., Manogaran, G., Taha, M. H. N., & Khalifa, N. E. M, "Fighting against COVID-19: A Novel Deep Learning Model Based On Yolo-V2 With Resnet-50 For Medical Face Mask Detection", *Sustainable Cities and Society*, 65, 102600. https://doi.org/10.1016/j.scs.2020.102600, 2021.
[16]. Malburg, L., Rieder, M. P., Seiger, R., Klein, P., & Bergmann, R, "Object Detection for Smart Factory Processes by Machine Learning", *The 4th International Conference on Emerging Data and Industry 4.0 (EDI40) March 23 - 26, 2021, Warsaw, Poland, Procedia Computer Science*, 184: 581–588. https://doi.org/10.1016/j.procs.2021.04.009, 2021.
[17]. Marcos, A. N., Gorka Azkune, G., & Carreras, I. A, "Egocentric Vision-based Action Recognition: A survey*", Neurocomputing*, 472 (2022), 175–197. https://doi.org/10.1016/j.neucom.2021.11.081, 2022.
[18]. Mohamed, H. E., Fadl, A., Anas, O., Wageeh, Y., El Masry, N., Nabilm, A., & Atia, A, "MSR-YOLO: Method to Enhance Fish Detection and Tracking in Fish Farms", *The 11th International Conference on Ambient Systems, Networks and Technologies (ANT), Warsaw, Poland, Procedia Computer Science*, 170: 539–546. https://doi.org/10.1016/j.procs.2020.03.123, April 6 - 9, 2020
[19]. Molchanov, V. V., Vishnyakov, B. V., Vizilter, Y. V., Vishnyakova, O. V., & Knyaz, V. V, "Pedestrian Detection In Video Surveillance Using Fully Convolutional YOLO Neural Network*", Proceedings Volume 10334, Automated Visual Inspection and Machine Vision II* , 103340Q. https://doi.org/10.1117/12.2270326 , 2017.
[20]. Namatevs, I., Sudars, K., & Polaka, I, "Automatic data labeling by neural networks for the counting of objects in videos", *Procedia Computer Science*, 149: 151–158. https://doi.org/10.1016/j.procs.2019.01.118, 2019.
[21]. Namazi, E., Mester, R., Lu, C., & Li, J, "Geolocation Estimation Of Target Vehicles Using Image Processing and Geometric Computation", *Neurocomputing*, 499: 35–46. https://doi.org/10.1016/j.neucom.2021.10.127, 2022.
[22]. Prasad, J., Jain, A., Velho, D., & K S, S.K, "COVID Vision: An Integrated Face Mask Detector And Social Distancing Tracker", *KeAi Chinese Roots Global Impact International Journal of Cognitive Computing in Engineering*, 3, 106–113. https://doi.org/10.1016/j.ijcce.2022.05.001, 2022.
[23]. Saada, M., Kouppas, C., Li, B., & Meng, Q, "A Multi-Object Tracker Using Dynamic Bayesian Networks And A Residual Neural Network Based Similarity Estimator", *Computer Vision and Image Understanding*, 225, 103569. https://doi.org/10.1016/j.cviu.2022.103569, 2022.
[24]. Sipetas, C., Keklikoglou, A., & Gonzales, E. J, "Estimation Of Left Behind Subway Passengers Through Archived Dataand Video Image Processing", *Transportation Research Part C*, 118, 102727. https://doi.org/10.1016/j.trc.2020.102727, 2020.
[25]. Shinde, S., Kothari, A., & Gupta, V, "YOLO based Human Action Recognition and Localization", International Conference on Robotics and Smart Manufacturing (RoSMa2018*)", Procedia Computer Science*, 133, 831–838. https://doi.org/10.1016/j.procs.2018.07.112, 2018.
[26]. Shuo, Z., Yanxia, W., Chaoguang, M., & Xiaosong, L, "Tiny YOLO Optimization Oriented Bus Passenger Object Detection", *Chinese Journal of Electronic*, 29(1): 132-138. https://doi.org/10.1049/cje.2019.11.002 , 2020.
[27]. Simon, M., Amende, K., Kraus, A., Honer, J., Samann, T., Kaulbersch, T., & Milz, S, "Complexer-YOLO: Real-Time 3D Object Detection and Tracking on Semantic Point Clouds", *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops 2019*, pp. 0-0. https://openaccess.thecvf.com, 2019.
[28]. Socha, K., Borg, M., & Henriksson, J, "SMIRK: A Machine Learning-Based Pedestrian Automatic Emergency Braking System With A Complete Safety Case", *Software Impacts*, 13, 100352. https://doi.org/10.1016/j.simpa.2022.100352, 2022.
[29]. Song, X., Gao, S., & Chen, C, "A Multispectral Feature Fusion Network For Robust Pedestrian Detection*", Alexandria Engineering Journal*, 60: 73–85. https://doi.org/10.1016/j.aej.2020.05.035, 2021.
[30]. Verstaevel, N., Barthélemy, J., Forehead, H., Arshad, B., & Perez, P, "Assessing The Effects Of Mobility On Air Quality: The Liverpool Smart Pedestrian Project", *World Conference on Transport Research – WCTR* 2019, Mumbai, 26-30 May 2019, Transportation Research Procedia, 48: 2197–2206. https://doi.org/10.1016/j.trpro.2020.08.276, 2020.
[31]. Wu, P., Li, H., Zeng, N., & Li, F, "FMD-Yolo: An Efficient Face Mask Detection Method For COVID-19 Prevention And Control In Public*", Image and Vision Computing*, 117, 104341. https://doi.org/10.1016/j.imavis.2021.104341, 2022.
[32]. Xue, L., Yan, W., Luo, P., Zhang, X., Chaikovska, T., Liu, K., Gao, W., & Yan, K, "Detection and localization of hand fractures based on GA_Faster R-CNN", *Alexandria Engineering Journal*, 60: 4555–4562. https://doi.org/10.1016/j.aej.2021.03.005, 2021.
[33]. Yu, J., & Choi, H, "YOLO MDE: Object Detection with Monocular Depth Estimation", *MDPI Electronics*, *11*(1), 76. https://www.doi.org/10.3390/electronics11010076, 2022.