



# Glove Detection System on Laboratory Members Using Yolov4

Abdul Khaliq Al Bari, Ema Rachmawati\*, Gamma Kosala

School of Computing, Telkom University, Bandung

Jl. Telekomunikasi. 1, Terusan Buahbatu - Bojongsoang, Telkom University, Sukapura, Kec. Dayeuhkolot, Kabupaten Bandung, Jawa Barat, Indonesia

Email: <sup>1</sup>abdulkhaliqalbari@student.telkomuniversity.ac.id, <sup>2,\*</sup>emarachmawati@telkomuniversity.ac.id,

<sup>3</sup>gammakosala@telkomuniversity.ac.id

Correspondence Author Email: emarachmawati@telkomuniversity.ac.id

Submitted: 07/07/2023; Accepted: 28/07/2023; Published: 31/07/2023

**Abstract**—The use of gloves by laboratory workers has become mandatory in laboratory work intending to maintain the safety of workers from the spread or side effects carried out in the laboratory, but there are still workers who violate the rules by not using gloves when workers are in the laboratory room. This study aims to detect the use of gloves by laboratory workers. The method used in this research is You Only Look Once (YOLO) version 4. YOLOv4 has a system that can complete computer visual tasks in detecting and detecting objects quickly in real time. Based on the results of experiments and testing conducted, the model obtain an Average IoU of 55.56%.

**Keywords:** Detection; Glove; Yolov4; Laboratory; Hyperparameter

## 1. INTRODUCTION

The laboratory is a place of work with a high risk by requiring workers to wear safety attributes. Laboratory work must follow procedures according to general standards because there are many risks for workers in the laboratory, such as the spread of viruses and exposure to the effects of chemicals [1]. One of the body parts that must be protected is the hand. Hands can be considered the most frequently used body part by laboratory workers.

Laboratories have general standards related to safety equipment on the attributes used by laboratory workers [2]. Safety equipment for laboratory workers can reduce risk, including gloves. Gloves have become a mandatory thing to use, which protects hands from various risks, such as exposure to chemicals or sharp objects so that workers are at risk of experiencing side effects from these chemicals [3]. The effects of the spread of chemicals can develop quickly. Negligence still occurs a lot among laboratory workers because workers violate health protocols that have become rules in the laboratory [4].

Research related to glove detection has been studied [5]. A safety glove detection system was built for workers who operate power grids. The glove detection system was built using the VGG-16 network-based method. The dataset used is self-built. Another study also examined glove detection to reduce defects in glove production [6]. The method used in their research is YOLOv5 which which produced the highest mAP accuracy of 0.9951.

Detected violations of helmets and masks and identified license plates using the YoloV4 method that utilizes TesseractOCR [7]. Tesseract model training is used to improve license plate character identification. The COCO dataset containing 600 images is used to build a system that obtains mAP results of 93.38% and an F1-Score of 0.77. Their system's shortcomings are the need to add geometric transformations to detect license plates using YoloV4.

Related researched a ball detection system using You Only Look Once (YOLO). The dataset used there are 202 photos divided into 70% training data, 20% validation, and 10% test data [8]. YOLOV4 performance evaluation uses a confusion matrix that calculates the results of accuracy, recall, and precision by testing under different conditions to detect an object. They conducted the test in real time with the level of the obstructed thing. The system can detect the ball with objects blocking the ball object by 50%, 60%, and 70%. Testing balls obstructed by other objects by 80%, 90%, and 100%, the system cannot detect the ball. There are areas for improvement in this detection system for testing objects with a distance above 200cm, and the system cannot detect the ball.

The glove detection system on the Personal Protective Equipment (PPE) attribute was also studied by Nurfirmansyah et al. [9]. The detection system using You Only Look Once (YOLO) was developed to reduce the negligence of building construction workers by wearing PPE. The system detects PPE objects in real time using a Webcam or CCTV. The dataset used in the training stage consists of 500 images containing five types of PPE and ten classes. They conducted real-time testing using a webcam and CCTV to detect PPE. The results of the tests carried out obtained an accuracy of 80%.

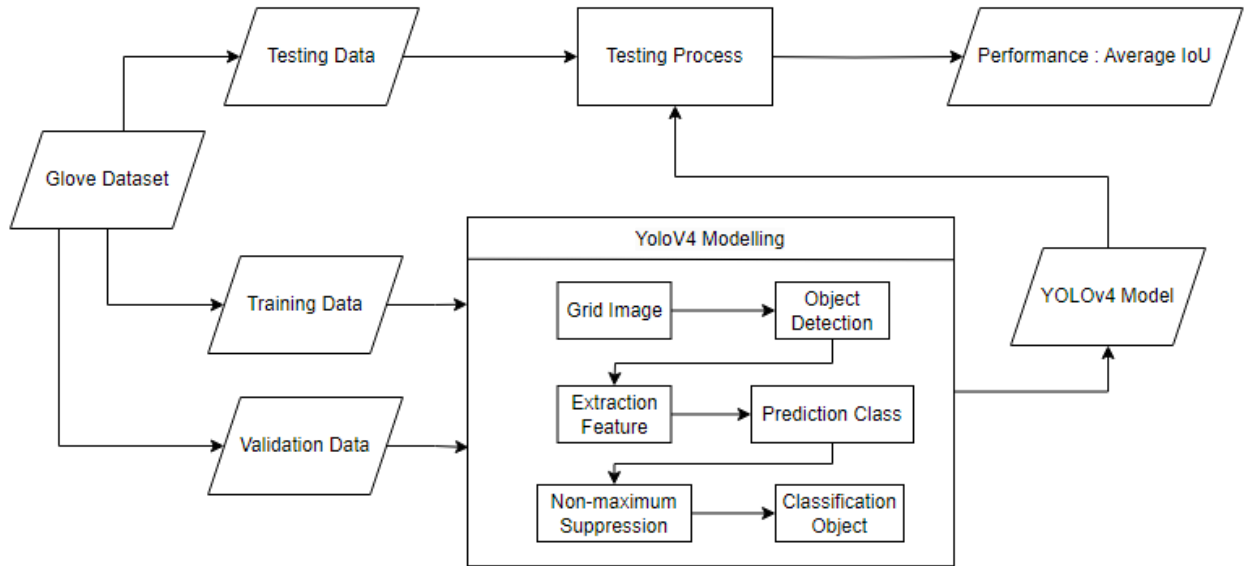
In this study, we built a system that wants to detect objects in the form of gloves as one of the health protocols that must be used during a pandemic. This research has built a detection system using the You Only Look Once (YOLO) version 4 method to detect gloves on laboratory workers. This method is used to obtain optimal accuracy and speed because the YOLO method is an algorithm that can detect objects in real-time [10]. The results of this study are expected to help officers to monitor laboratory workers who comply with one of the health protocols by using gloves or not and are expected to be one way to increase laboratory workers' awareness of the importance of using safety protocols at work in the laboratory [11].

This article is organized as follows. Section 1 discusses the background of the problem that led to this research. There are related previous studies, especially in glove detection research. Section 2 discusses Research Methodology. This chapter explains the stages carried out in this research. Section 3 discusses the experiments that have been carried out and the analysis of the experiments carried out. Section 4 discusses the conclusions of this research.

## 2. RESEARCH METHODOLOGY

### 2.1 Overview

The system built in this study is a detection system for the use of gloves on laboratory workers using YOLOv4. The illustration is shown in Figure 1.



**Figure 1.** Glove Detection System Design

In this stage, training is performed on the training data using YOLOv4. The YOLOv4 model recognizes objects based on existing classes, namely the glove class. YOLOv4 performs image recognition with the number of images determined by batch size at each iteration. After training, a weight file containing the updated network weights is obtained. The data used for training is divided into training data and validation data. The image is divided into a grid or small cells of a specific size. Each cell represents a part of the image to be analyzed. After the image is divided into grids, the object detection process begins. YOLOv4 performs object detection using a deep convolutional neural network (CNN)[12]. This network has a convolutional layer that contains a feature extraction layer and an object detection layer. The first convolution layer in the YOLOv4 network is responsible for feature extraction from the image [13]. It extracts essential features that help in recognizing objects. After feature extraction, the object detection layer predicts the location and class of the object. Each cell in the grid predicts multiple bounding boxes containing the object. The prediction of the object class and confidence score is also performed for each bounding box. After location prediction and object classification are performed, the non-maximum suppression (NMS) method eliminates overlapping bounding boxes and retains the bounding box with the highest confidence score. After the detection stage, the detected objects are classified based on their class [14].

YOLOv4 performs object detection on the validation data and calculates evaluation metrics to measure model performance. The validation process calculates the resulting loss from the model by comparing the model predictions with the actual annotations on the validation images. The Validation stage calculates the CIoU loss to measure the extent to which the model successfully detects the image and identifies the correct bounding box for each object. CIoU loss calculation can be formulated as follows:

$$L_{CIoU} = 1 - IoU + \frac{p^2(b, b^{gt})}{c^2} + \alpha v \quad (1)$$

The matrices denoted by  $b$  and  $b^{gt}$  correspond to the center points, respectively. The variable  $p^2$  is the Euclidean distance, the variable  $c$  is the diagonal length of the smallest enclosing box that covers both boxes,  $\alpha$  is a positive trade-off parameter, and  $v$  measures the consistency of the aspect ratio [15]. The validation stage also measures performance using the Intersection over Union (IoU) metric. IoU measures the extent to which the prediction bounding box overlaps with the ground truth bounding box. If the IoU between the prediction and ground truth exceeds the threshold of 0.5, the bounding box is considered a positive bounding box. If the IoU is

below the threshold, the bounding box is considered a negative bounding box or there are no objects in it. This stage can evaluate whether the model is adequate or whether further tuning of the model architecture and hyperparameters is required.

Testing is done to determine the detection results by testing the test data image. The testing process loads the YOLOv4 model that has been trained. This test displays the image detected by the system in the form of a prediction bounding box for the detected glove image. Performance measurement of the results of testing glove detection is carried out by calculating IoU.

## 2.2 Evaluation

The evaluation carried out in this study to measure the performance of the method used is that there are several performance parameters used. from the performance parameters can be used as an analysis in the performance of the method used.

### a) Average IoU

This performance measurement compares the image's ground truth with the model's predicted bounding box. If the bounding box is further from the ground truth, it results in a smaller IoU value [16], [17]. The IoU equation can be seen in Equation 2.

$$IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}} \quad (2)$$

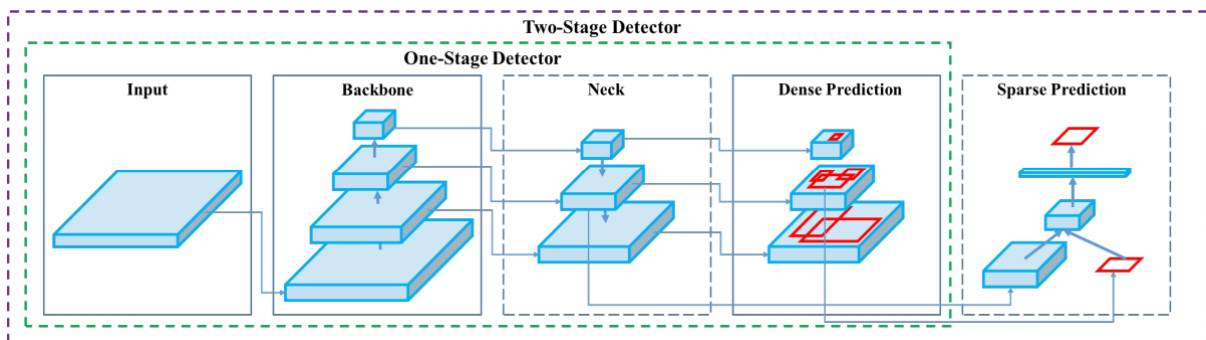
### b) mAP

Average Precision (AP) is calculated as the weighted mean of precisions at each threshold, the weight is the increase in recall from the prior threshold. Mean Average Precision(mAP) is the average of AP of each class [17]. The mAP Equation can be seen in Equation 3. There are a total of n object classes in the dataset, so the variable i goes from 1 to n to describe each object class individually. In this context, i represents the index number of the object class being evaluated.

$$mAP = \frac{1}{n} \sum_{i=1}^N AP_i \quad (3)$$

## 2.3 You Only Look Once (YOLO) Version 4

YOLOv4 (You Only Look Once version 4) is an object recognition model designed to identify objects in images or videos quickly and accurately [18]. It further develops the "one-stage object detection" approach, prioritizing speed and efficiency. YOLOv4 includes several improvements and innovations compared to the previous version (YOLOv3), including advanced techniques such as model optimization, accuracy improvement, and noise reduction. This method is an improvement over the previous version, YOLOv3. YOLOv4 has a 10% higher AP improvement and 12% higher FPS than the previous version [15]. YOLOv4 uses the CSPDarknet 53 Backbone architecture of CNN. The working system of YOLOv4 is to perform convolution and then the max pooling process to reduce the image dimension. The confidence value and class labelling are given to the bounding box in the following process. The last process is Non-Max Suppression which eliminates all bounding boxes with low confidence values and leaves 1 with the highest confidence value [17]. This method has an architecture that can be seen in Figure 2.



**Figure 2.** Architecture YOLOv4

YOLO4 mainly consists of 3 parts, namely Backbone, Neck and Head. For Backbone and Neck itself has a function for feature extraction and aggregation, and for Head has a function to detect or predict. In the YOLOv4 algorithm, the modules used are CSPDarknet53 (Backbone), PANet (Neck), and YOLOv3 (Head) [15]. Darknet is a powerful and efficient convolutional neural network architecture. In YOLOv4, the Darknet backbone is enhanced by using residual blocks, which allows the model to learn more robust feature representations. YOLOv4 uses bagging (ensembling) techniques to improve its detection accuracy. In YOLOv4, several different YOLO models are trained independently, combining their predictions to produce the final prediction. Combining predictions from multiple models, YOLOv4 can reduce detection errors and improve model reliability.

### 3. RESULT AND DISCUSSION

In this section, this research conducted data training using YOLOv4. The training process is carried out three times, where the best training results can be used to test with test data.

#### 3.1 Dataset

The dataset used in this study is a dataset taken from Kaggle's "Medical Personal Protective Equipment Dataset" [19]. The glove dataset consists of training, validation, and testing data. The dataset is in the form of images with jpg data type. Table 1 shows the number of dataset image divisions used in the system development. Figure 3 shows an example of a glove image dataset.

**Table 1.** Dataset

Data	Images
Training	721
Validation	228
Testing	64
Total	1013



**Figure 3.** Sample of Glove Image Dataset

#### 3.2 Training Yolov4

In this section, training is one of the important processes in making an object detection system. To achieve maximum training results, several parameters must be set to obtain maximum performance. The result of the training outputs the value which can produce output with IoU and mAP score. Hardware usage is one of the components that affect the duration of the training process. The hardware used in the training process is :

- Processor AMD Ryzen 5 3550H @ 2.1 GHz.
- RAM 12 GB DDR4.
- Graphics Card NVIDIA GeForce GTX 1050.
- Operating Windows 11 Home Single Language 64-bit.

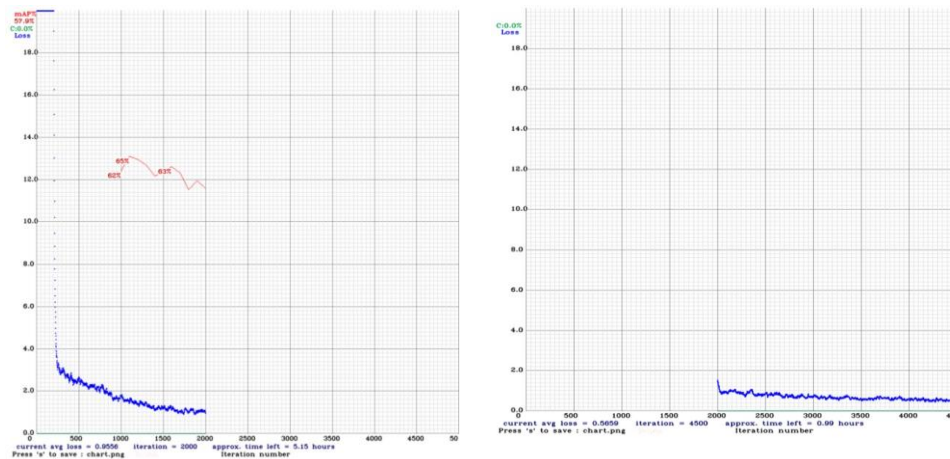
Experiments were conducted by training three times based on different hyperparameters. Batch size is one of the most important hyper parameters, because batch size is a variable that contains the number of training samples used in one batch for one iteration [20]. This experiment was conducted to find the best hyperparameter to obtain the best model results. Table 2 shows the results of the experiments using different hyperparameters.

**Table 2.** Result of Training

Metric	Precision	Recall	F1-Score	Average IoU	mAP	Time
Batch Size = 32						
Max Batches = 6000	53	53	53	32.25	46.23	4 hours
Image Resize = 224x224						
Random = 0						
Batch Size = 64						
Max Batches = 4000	62	51	56	44.75	46.5	5 hours
Image Resize = 416x416						
Random = 0						
Batch Size = 64						
Max Batches = 6000	77	80	78	58.25	76.54	10 hours
Image Resize = 416x416						
Random = 1						

The hyperparameters used in the first training are batch 32, maximum batch 6000, and random value 0. The first experiment was done with an image that was resized to 224x224. In the first experiment the results obtained from the first experiment were less than optimal due to improper hyperparameter configuration settings to obtain maximum results from the training conducted. The training process takes little time with this configuration, but with the hyperparameter configuration, it still needs to obtain good results. In addition, the random value also affects the results and duration of training that has been done. Using a random value of 0, the duration of the training process is fast, but this dramatically affects the results obtained, so the training results from the first scenario do not perform well.

The second experiment was conducted using a different hyperparameter configuration with a batch size of 64 and max batches 4000. There is an adjustment uses different size settings as well, namely by resizing the image to a 416x416 size. With this hyper-parameter configuration the performance obtain from the training process in this second experiment increased in terms of IoU.



**Figure 4.** Graph of Training Process

The third experiment using a batch size of 64, max batches at 6000, and a random value of 1. Figure 4 shows a decrease in training loss at each iteration which can produce good results. This third experiment was conducted using a dataset that was resized to 416x416. With the hyperparameter configuration used and the size of the trained image, the results of the training in this experiment found that a performance improvement was obtained. The best performance is obtain from training process at the 2000th iteration. However, although the performance obtained is better than before, the training duration is much longer, but the resulting training obtain better performance.

There are differences in the results of the training process performed. Hyperparameters and image size of the training process determine the results of the performance obtained. In Table 3, we can see the difference from each configuration. The improvement of the training results can be seen from the first experiment to the last. The best training results is show the third experiment using a batch size of 64 and a random value of 1 where mAP value of 76.54% and an Average IoU value of 58.25%.

### 3.3 Testing

The system that has been trained using the Yolov4 method is tested to find out how well the performance is obtained when detecting gloves. Tests were conducted on 64 images from the test data. Sample of test results can be seen in Figure 5 and Figure 6.



**Figure 5.** Sample of Correct Glove Detection Test

Figure 5 show examples of correct glove detection. With the model that has been trained, the glove detection process can be seen from the prediction of the bounding box against the glove.



**Figure 6.** Sample of Incorrect Glove Detection Test

Figure 6 shows the results of incorrect glove detection. There are detection defects in the bounding box that do not match the coordinates with the intended glove. Figure 6(a) shows an example of a mismatch in the detection result caused by a blurred image. The system is unable to detect the gloves correctly. Figure 6(b) shows an undetected object. The detected gloves have a similar colour, so the system cannot detect the object as a glove. Figure 6(c) shows that undetected gloves are also caused by the object being truncated so that it is not fully visible as a glove.

The results of each image test that has been carried out, there are gloves that are detected correctly and there are gloves that are detected incorrectly. There are three conditions of incorrect glove detection, when the image is blurred, the object is the same color, and the glove is truncated. There are also detection defects with bounding boxes that deviate from the gloves. The system test results obtain an average IoU of 55.56% from the whole test.

## 4. CONCLUSION

Glove Detection using the YOLOv4 method effectively ensures safety and quality in laboratory environments. The system can help prevent contamination risks and maintain high hygiene standards with fast and accurate glove detection. Using YOLOv4, the system received an average IoU score of 55.56% on the test. The glove detection system using the YOLOv4 method has the potential to be further developed. Model updates and larger data sets can optimise and improve the method. Performance development can also be done by training with more significant iterations using a larger image size.

## REFERENCES

- [1] A. Galasso Hong Luo Brooklynn Zhu *et al.*, “Laboratory Safety and Research Productivity,” *Research Policy*, vol. 52, no. 8, 2023, doi: 10.1016/j.respol.2023.104827.
- [2] A. D. Ménard and J. F. Trant, “A review and critique of academic lab safety research,” *Nature Chemistry*, vol. 12, no. 1. Nature Research, pp. 17–25, Jan. 01, 2020. doi: 10.1038/s41557-019-0375-x.
- [3] R. M. Broyer, K. Miller, S. Ramachandran, S. Fu, K. Howell, and S. Cutchin, “Using Virtual Reality to Demonstrate Glove Hygiene in Introductory Chemistry Laboratories,” *J Chem Educ*, vol. 98, no. 1, pp. 224–229, Jan. 2021, doi: 10.1021/acs.jchemed.0c00137.
- [4] Y. Yang, G. Reniers, G. Chen, and F. Goerlandt, “A bibliometric review of laboratory safety in universities,” *Safety Science*, vol. 120. Elsevier B.V., pp. 14–24, Dec. 01, 2019. doi: 10.1016/j.ssci.2019.06.022.
- [5] M. Jin *et al.*, “Glove detection system based on VGG-16 network,” in *Proceedings - 2020 13th International Symposium on Computational Intelligence and Design, ISCID 2020*, Institute of Electrical and Electronics Engineers Inc., Dec. 2020, pp. 172–175. doi: 10.1109/ISCID51228.2020.00045.
- [6] Y. C. How, A. F. Ab. Nasir, K. F. Muhammad, A. P. P. Abdul Majeed, M. A. Mohd Razman, and M. A. Zakaria, “Glove Defect Detection Via YOLO V5,” *MEKATRONIKA*, vol. 3, no. 2, pp. 25–30, Jan. 2022, doi: 10.15282/mekatronika.v3i2.7342.



- [7] R. Syamsul Huda, R. Wulanningrum, and D. Swanjaya, "Pemanfaatan YOLOv4 Untuk Deteksi Pelanggaran Helm dan Masker Serta Identifikasi Pelat Nomor Menunggunakan Tesseract-OCR," *Joutica: Journal of Informatic Unisla*, 7(2), 2022, pp. 596–602.
- [8] J. Nuralim, N. Fath, A. Musafa, and S. Broto, "Perancangan Sistem Pendeteksian Obyek Bola Dengan Metode Framework YOLOv4," *Jurnal Maestro*, vol. 5, no. 2, 2022, pp. 289–294.
- [9] A. Nurfirmansyah and R. Dijaya, "Deteksi Kelalaian Alat Pelindung Diri (APD) Pada Pekerja Konstruksi Bangunan," *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)*, 6(1), 058–063, Aug. 2022, Doi: 10.29407/inotek.v6i1.2452.
- [10] T. Diwan, G. Anirudh, and J. V. Tembhurne, "Object detection using YOLO: challenges, architectural successors, datasets and applications," *Multimed Tools Appl*, vol. 82, no. 6, pp. 9243–9275, Mar. 2023, doi: 10.1007/s11042-022-13644-y.
- [11] N. Balla, H. Kumar, and V. P. Mishra, "Health and safety Protocols in Chemistry laboratory and its Precursors Disposal with Chemical Incompatibility," in *Proceedings of 2nd IEEE International Conference on Computational Intelligence and Knowledge Economy, ICCIKE 2021*, Institute of Electrical and Electronics Engineers Inc., Mar. 2021, pp. 166–169. doi: 10.1109/ICCIKE51210.2021.9410763.
- [12] C. Dewi, R. C. Chen, Y. T. Liu, X. Jiang, and K. D. Hartomo, "Yolo V4 for Advanced Traffic Sign Recognition with Synthetic Training Data Generated by Various GAN," *IEEE Access*, vol. 9, pp. 97228–97242, 2021, doi: 10.1109/ACCESS.2021.3094201.
- [13] E. R. Setyaningsih and M. S. Edy, "YOLOv4 dan Mask R-CNN Untuk Deteksi Kerusakan Pada Karung Komoditi," *Teknika*, vol. 11, no. 1, pp. 45–52, Mar. 2022, doi: 10.34148/teknika.v11i1.419.
- [14] B. Chethan Kumar, R. Punitha, and Mohana, "YOLOv3 and YOLOv4: Multiple object detection for surveillance applications," in *Proceedings of the 3rd International Conference on Smart Systems and Inventive Technology, ICSSIT 2020*, Institute of Electrical and Electronics Engineers Inc., Aug. 2020, pp. 1316–1321. doi: 10.1109/ICSSIT48917.2020.9214094.
- [15] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," Apr. 2020, [Online]. Available: <http://arxiv.org/abs/2004.10934>
- [16] W. Rahmانيar and A. Hernawan, "Real-time human detection using deep learning on embedded platforms: A review," *Journal of Robotics and Control (JRC)*, vol. 2, no. 6. Department of Agribusiness, Universitas Muhammadiyah Yogyakarta, pp. 462–468Y, Nov. 01, 2021. doi: 10.18196/jrc.26123.
- [17] T. A. A. H. Kusuma, K. Usman, and S. Saidah, "People Counting For Public Transportations Using You Only Look Once Method," *Jurnal Teknik Informatika (Jutif)*, vol. 2, no. 1, pp. 57–66, Feb. 2021, doi: 10.20884/1.jutif.2021.2.2.77.
- [18] O. ÖNAL and E. DANDIL, "Object Detection for Safe Working Environments using YOLOv4 Deep Learning Model," *European Journal of Science and Technology*, Jun. 2021, doi: 10.31590/ejosat.951733.
- [19] R. Dagli and A. M. Shaikh, "CPPE-5: Medical Personal Protective Equipment Dataset," *SN Comput Sci*, vol. 4, no. 3, May 2023, doi: 10.1007/s42979-023-01748-7.
- [20] Y. Ji, H. Zhang, Z. Zhang, and M. Liu, "CNN-based encoder-decoder networks for salient object detection: A comprehensive review and recent advances," *Inf Sci (N Y)*, vol. 546, pp. 835–857, Feb. 2021, doi: 10.1016/j.ins.2020.09.003.