INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2023, VOL. 69, NO. 4, PP. 691-696
Manuscript received March 28, 2023; revised October, 2023. DOI: 10.24425/ijet.2023.147689

The Role of Faster R-CNN Algorithm in the Internet of Things to Detect Mask Wearing: The Endemic Preparations

Marah Doly Nasution, Al-Khowarizmi, Romi Fadillah Rahmat*, Arif Ridho Lubis, and Muharman Lubis

Abstract—**Faster R-CNN is an algorithm development that continuously starts from CNN then R-CNN and Faster R-CNN. The development of the algorithm is needed to test whether the heuristic algorithm has optimal provisions. Broadly speaking, faster R-CNN is included in algorithms that are able to solve neural network and machine learning problems to detect a moving object. One of the moving objects in the current phenomenon is the use of masks. Where various countries in the world have issued endemic orations after the Covid 19 pandemic occurred. Detection tool has been prepared that has been tested at the mandatory mask door, namely for mask users. In this paper, the role of the Faster R-CNN algorithm has been carried out to detect masks poured on Internet of Thinks (IoT) devices to automatically open doors for standard mask users. From the results received that testing on the detection of moving mask objects when used reaches 100% optimal at a distance of 0.5 to 1 meter and 95% at a distance of 1.5 to 2 meters so that the process of sending detection signals to IoT devices can be carried out at a distance of 1 meter at the position mask users to automatic doors.**

Keywords— **Faster R-CNN; IoT; Mask; Endemic**

I. Introduction

MASK detection is a fundamental and important problem in computer vision today due to the Covid-19 virus pandemic where mask detection is an important step in suppressing the spread of the Covid-19 virus [1]. The purpose of mask detection is to determine the existence of the mask position in the video and if the mask is found, the mask position is marked with a bounding box [2]. There are many factors that affect mask detection, namely video quality, the position of the mask user on various face scales and occlusions[3]. Faster R-CNN is a new region-based object detection algorithm which shows excellent results on a wide range of object detection[4]. Currently the use of masks in public places is increasing. The use of masks is believed by experts to prevent the spread of Covid.

Regulations on the use of masks were enforced by the government in line with the Covid-19 virus that hit Indonesia [5]. Masks are a mandatory item during the adaptation period to new habits due to the virus pandemic [6]. Disciplined use of masks is believed to be able to reduce the level of spread of the Covid-19 virus [7]. During a pandemic, many types of masks are sold freely on the market, ranging from masks that have standards or do not have standards, making it difficult for the government to reduce the spread of Covid-19 [8]. Therefore the need for a prototype in detecting masks that have standards or non-standards with the intention of minimizing the transmission of Covid-19 and supporting the implementation of a new normal life which is said to be Endemic will run while reminding the importance of using masks [9].

In doing the identification of the mask used, of course, it involves several techniques such as image processing in a Neural Network. Image processing is a technique for analyzing data in the form of images that can be processed according to needs and provide accurate information for humans to receive [10][11][12]. Where the image processing process forms the function and intensity of light with representation in 2 dimensions and 3 dimensions [13]. From various images have the characteristics of information that is extracted into knowledge [14]. Image classification is really needed in terms of classifying, this classification aims to recognize and understand its use in computers quickly, this is included in the field of Neural Networks [15][16]. Methods in image processing certainly explain from an image with the application of computer vision which can be utilized in various fields such as what is trending in agriculture with smart farming, health with digital health, manufacturing with e-industry in order to extract information into knowledge in the field of computing [17].

The application of detection has been carried out in many studies such as those carried out by References [18] conducted research by modifying the Faster R-CNN for detection of PPI radars and obtained research results that the modified Faster R-CNN is capable of detecting PPI radars on the seabed compared to unmodified Faster R-CNN. On reference [19] classifying an object with Faster R-CNN also on the GPR system which has the best from before that in classifying R-CNN it is more accurate than Faster RCNN on the GPR system as an object classification technique. On reference [20] A detection technique has been carried out on book objects with a faster R-

Marah Doly Nasution is with Department of Mathematics Education, Universitas Muhammadiyah Sumatera Utara, Indonesia (e-mail: marahdoly@umsu.ac.id).

Al-Khowarizmi is with Department of Information Technology, Universitas Muhammadiyah Sumatera Utara, Indonesia (e-mail: alkhowarizmi@umsu.ac.id).

Romi Fadillah Rahmat with Department of Information Technology, Universitas Sumatera Utara, Indonesia (e-mail:romi.fadillah@usu.ac.id).

Arif Ridho Lubis with Department of Management Informatics, Politeknik Negeri Medan, Indonesia (e-mail: arifridho@polmed.ac.id) .

Muharman Lubis with Department of Information Systems, Telkom University, Indonesia (e-mail: muharmanlubis@telkomuniversity.ac.id).

CNN which achieves accurate and fast results in object detection.

The prototype built in this paper will use the Faster Regional Convolutional Neural Network (Faster R-CNN) algorithm. Where Faster R-CNN which is a method of object detection that is able to detect an object captured by the camera [21][22][23]. prototype work begins with the detection process on mask detection or not using and whether the mask is used is standard or not so as to support prototype data stored in the cloud network [24]. After the prototype was developed in this paper, the application that was carried out was to provide a detection signal to IoT devices in carrying out automatic door opening for standard mask and mask users.

II. Material and Method

A. Dataset

This paper will use the Faster R-CNN method as a role in detecting standard and non-standard masters. Before entering deep learning on Faster RCNN, we first collect data sets so that they can be tested on IoT devices for mandatory masks. As a method for detecting standard and non-standard masks, the first thing to do is to take a dataset that begins with camera preparations from a mobile device. To start with the face facing the camera using a standard mask and a non-standard mask then the system will later detect whether the mask is standard or not with a distance of 0.5 metre, 1 metre, 1.5 metre and 2 metre. The dataset in this study amounted to 100 people with. The shooting is summarized in Figure 1 below.

Fig. 1. Summary of the mask user dataset

From Figure 1, it can be seen that taking pictures of mask users will then be taught in the Faster R-CNN method.

B. General and Focus Research

In the process of carrying out the research, of course, it is described in a flow map so that it is easy to understand both the process and the results to be aimed at and can be seen in Fig. 2.

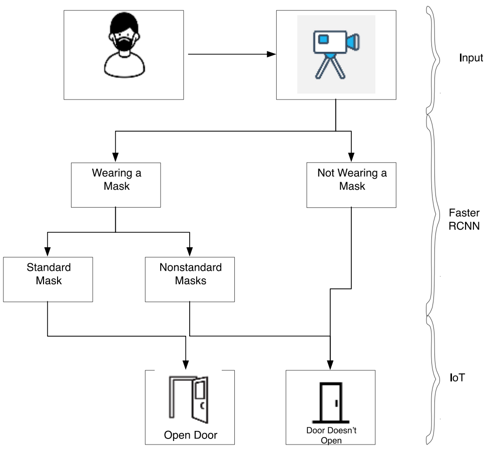


Fig. 2. Research Flow Map

From Figure 2 it can be seen that the core process consists of 3 parts, namely input which is part of the data for mask users and not wearing masks, then Faster RCNN to detect masks used standard or nonstandard and finally IoT devices to be used at mandatory mask doors. The prototype built in this study will use the Faster R-CNN algorithm. Where Faster R-CNN which is a method of object detection that is able to detect an object captured by the camera [21][22][23]. The workings of the prototype start with the process of detecting whether the object is wearing a mask or not, then if it is wearing a mask, whether it is using a standard mask or not, then the data is sent to the cloud to be sent to the IoT device [24]. So, the door is only open for those who wear masks and use standard masks. However, in optimizing the Faster RCNN method for detecting masks, of course a general architecture is designed so that the method in computing does not widen and expand as unwanted. The general architecture in the faster RCCN method is shown in Figure 3 as follows:

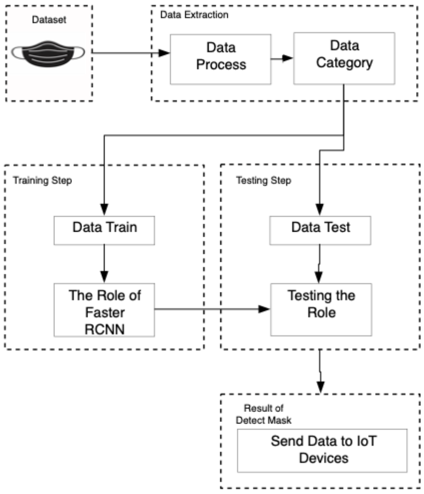


Fig. 3. General Architecture Faster RCNN

In Figure 3 it can be seen that the detection by optimizing the faster RCNN method is carried out based on the following steps:

1. Take a face with a mask
2. Perform image extraction with:
   a. Doing data processing, namely resizing the image to a size of 300 pixels x 300 pixels.
   b. Categorizing the data, namely dividing it into training data and test data. The data sample divides the training test data and test data with a statistical approach, namely the slovin approach which is calculated based on equation (1).

$$n = \frac{N}{1+Ne^2} \quad (1)$$

   Where:

   n is the number of datasets to be searched for

   N is dataset size

   e is margin of error value of the dataset size
3. Perform mask detection with Faster RCNN on training data
   a. Input neuron
   b. Doing a convolution layer by combining two number series to produce a third number series

c. Measuring the activation function with ReLu $f(x) = \max(x, 0)$.

4. Perform mask detection with faster RCNN on testing data
   a. Input neuron
   b. Doing a convolution layer by combining two number series to produce a third number series
   c. Measuring the activation function with ReLu $f(x) = \max(x, 0)$.
5. After getting detection on standard masks and non-standard masks, the results are sent to the IoT device for the door status to open automatically or not to open.

III. Result and Discussion

A. *The Role of Faster R-CNN Role for Mask Detection*

Around 2010 detection of deep learning models was indispensable at the mundane level[25]. It is better as a feature as deep convolutional which is an artificial conditional network that is best at learning strong and high-level features. This form of object recognition is better known as the CNN region (RCNN). [26]. The era of deep learning introduced two types of two-stage detection and single-stage detection models. However, in Faster R-CNN the data mining process is carried out completely based on the features of objects in the convolutional network and enters the deepest layer and is interconnected in the learning process. The Faster RCNN model can be seen in Figure 4.

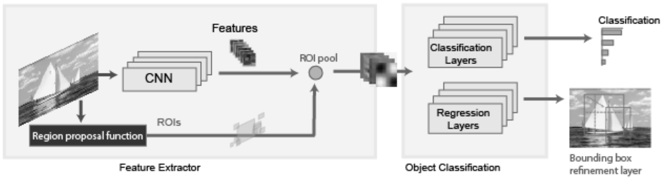


Fig. 4. Faster RCNN Model

Figure 4 shows a more complete faster R-CNN model by applying deep learning as a result of the development of R-CNN capable of producing accuracy in real time. Where, the previous process implemented a search algorithm by utilizing ROI. Meanwhile, in the faster R-CNN search algorithm process is eliminated and the artificial neural network process is applied to features to detect an object. Meanwhile, in detecting objects in this paper, an RPN process is developed for optimal results as depicted in Figure 5.

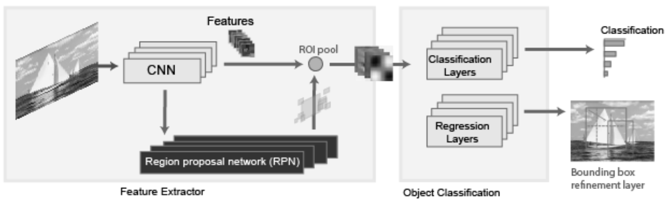


Fig. 5. Faster RCNN for object detection

In Figure 5 it can be seen that there is a development model for detecting an object, namely a moving image to detect standard mask users. The addition of this feature speeds up transmission into the neural network at the CNN layer. To implement the Faster R-CNN technique with the data set that has been obtained, it can then be calculated using the formula in equation (2) as follows [27]:

$$L(P_i) = \frac{1}{N_{cls}} \sum L_{cls}(p_i, P_i^*) \qquad (2)$$

Where p_i is a probability prediction P_i^* is the anchor value and L_{cls}= is the anchor number of several stages. The results of the calculations obtained from the training results are depicted on a chart shown in Figure 6 below.

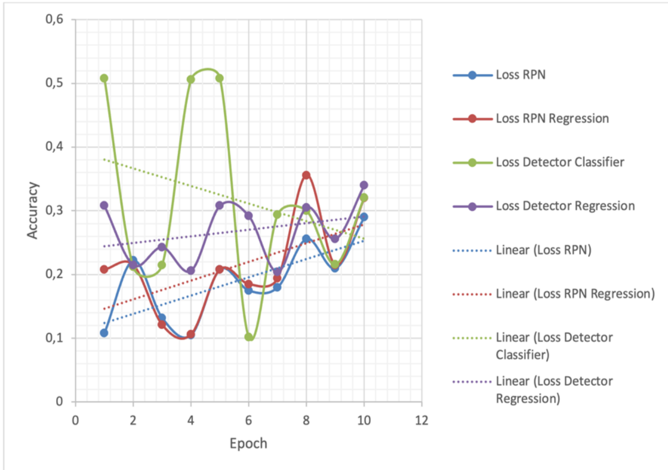


Fig. 6. Training Process

From Figure 6. The result of doing the data set in the training process is a model consisting of 3 stages, namely the Convolutional layer, Pooling layer, and fully connected layer. The results of the data set using the faster R-CNN are the formation of a model and data record that is ready for further detection. In the data training process, loss is interpreted as a penalty for the system in predicting wrong results. Where loss is said to be a number that can show an indication of how bad the predictions of the sample data set model are being tested. If the data set being tested is good, then the loss will be zero. If the opposite happens, the amount of loss you get will definitely be more. The purpose of training a model is to find the load and data bias that has a low loss on average as much as the amount of data in the dataset.

The test was carried out to determine the performance of the Faster R-CNN algorithm by looking for accuracy, precision and recall using the K-Fold Cross Validation method. In this test the dataset is divided into 100 dataset images which are divided into train data and test data randomly. The train data for each is 50 images, while the test data is 50 images. At each fold, five trials were carried out and the accuracy was calculated. in Table 1 with the K-Fold Cross Validation the results obtained from the test display the accuracy, precision, and recall that has been calculated for each fold by calculating using the Confussion Matrix. The confusion matrix is a method used to calculate accuracy values in data mining concepts where the results are shown in table I.

TABLE I
TESTING THE RESULTS

Fold	Accuracy %	Precision %	Recall %
1	80%	80%	80%
2	75%	75%	75%
3	95%	95%	95%
4	79%	79%	79%
5	90%	90%	90%

Accuracy is defined as the amount of data that is predicted correctly, from the entire dataset. The accuracy calculation is the number of true positives and true negatives of the expression given by the formula divided by the total number of true positives, true negatives, false positives, and false negatives based on equation (3) below.

$$Acc = \frac{TP+TN}{TP+TN+FF+FN} \quad (3)$$

Precision comes from the fraction of relevant data for all test data that is true. A perfect classifier has a precision of 1, or 100%. Precision is calculated as the number of true positives divided by the sum of true positives and false positives, as shown in the formula based on Equation (4) below.

$$Prec = \frac{TP}{TP+TF} \quad (4)$$

Recall, sometimes called sensitivity, is part of the relevant data and has the true value of all relevant data. A perfect classifier has a recall of 1 or 100%. Recall is determined by dividing the number of true-positive data by the sum of true-positive and false-negative data, as in Equation (5) below.

$$Recc = \frac{TP}{TP+TN} \quad (5)$$

B. *Mask Distance Accuracy Testing*

After successfully detecting the accuracy in reading the mask, then determine the accuracy of the dataset that has been tested. Testing is carried out on the system by preparing the data set first. The dataset that has recorded 100 shots using both standard masks and non-standard masks with a distance of 0.5 meters, 1 meter, 1.5 meters and 2 meters and the measurement results are shown in table II as follows.

TABLE II
STANDARD MASK DISTANCE ACCURACY TEST

Testing Distance	Accuracy	Giving Signals
0.5 Metre	100 %	Open
1 Metre	100%	Open
1.5 Metre	95%	Open
2 Metre	90%	Open

After that, testing the standard mask distance accuracy table using a dataset that has been recorded 100 shots using both standard masks and non-standard masks with a distance of 0.5 meters, 1 meter, 1.5 meters and 2 meters and the calculation results are shown in table III.

From table II and table III, the dataset results were successfully tested and the system did not detect non-standard masks. Furthermore, testing the data set time aims to determine the time needed to detect standard and non-standard types of masks in taking action. The calculation results from the dataset that has been recorded are shown in table IV as follows:

TABLE III
NON-STANDARD MASK DISTANCE ACCURACY TEST

Testing Distance	Accuracy	Giving Signals
0.5 Metre	100 %	Not Open
1 Metre	100%	Not Open
1.5 Metre	95%	Not Open
2 Metre	90%	Not Open

TABLE IV
MASK TIME ACCURACY TEST

Brand	Accuracy	Detection
No Brand	20 Second	Not Detected
Duckbil	20 Second	Detected
KN - 45	20 Second	Detected
Surgical Mask 2 ply	30 Second	Detected
Bedah 3 ply	30 Second	Detected

From table IV it can be seen that various brands of masks have been observed that Duckbil, KN-45, Surgical Mask 2 ply, and Surgical 3 ply masks are included in the standard masks. Meanwhile, cloth masks, which are mostly produced by human hands, are not included in standard masks. It can also be seen from the results of the accuracy when detecting masks on standard masks with Duckbil and KN-45 masks, the fastest time is 20 seconds for mask detection, while cloth masks, which are masters, do not get 20 seconds for detection. The application of the Faster RCNN method applied to mobile devices can be seen in the summary of Figure 7 as follows:

Fig. 7. Role of Faster RCNN on Mask Detection

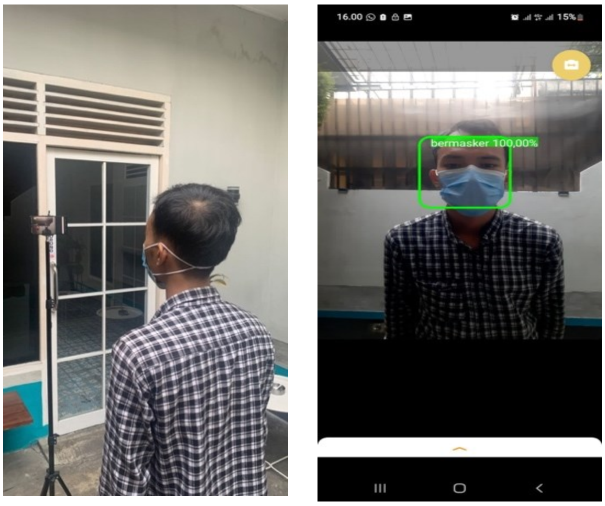
From Figure 7, the results of detection are summarized using a faster RCNN where the initial detection is if you don't wear a mask and a non-standard mask will not be detected. Whereas a standard mask will gain accuracy and give a signal to IoT devices to give orders to open or not open the door automatically.

C. *Signal Sending Process to IoT Devices*

With the optimization of the Faster RCNN method for detecting standard masks, the next stage is sending signals to IoT devices. The test was carried out by making a miniature or simulating an artificial door with a small size as shown in Figure 8 below.

Fig. 8. Door Automation Simulation Process from an IoT device

From Figure 8 it can be seen that the simulation process on IoT devices when the standard mask detection process is read by a mobile device then a signal is given to the IoT device. After succeeding in the automatic miniature door simulation, it is then carried out on the mandatory mask door to detect standard masks and non-standard masks, as shown in Figure 9 below.



(a) (b)

Fig. 9. Mask Detection on IoT Devices
(a). From Rear View and (b). from Mobile devices

From figure 9a. and 9b. it can be seen that the IoT device is placed at the door, but on detection using a mobile device so that it is easy to send a signal from the detection of standard masks and non-standard masks that the results of detection of standard masks have been successfully detected with 100% accuracy. So that when it is successfully detected, the mobile device will give signals and commands to the Rasberry Pi device to open the door automatically. If it is detected not using a mask and wearing a non-standard mask, the system will not be detected and will not open the door automatically.

IV. Conclusion

From the test results that have been outlined in this paper, it can be seen that all processes, both in detecting mask users and not wearing masks, are running well. In addition, the detection of standard masks and non-standard masks also works well. Where the detection model used uses the role of the faster R-CNN which is then poured into the IoT device for application to automatic doors. From the overall observation results in testing the detection of a mask object that moves when used it reaches 100% optimally at a distance of 0.5 to 1 meter and 95% at a distance of 1.5 to 2 meters so that the process of sending a detection signal to an IoT device can be carried out at a distance of 1 meter at the position mask users to automatic doors.

Acknowledgements

Our thanks go to Prof. Dr. Agussani, M.AP. The Rector of the Universitas Muhammadiyah Sumatrera Utara and Prof. Dr. Muryanto Amin, M.Si. The Rector of the Universitas Sumatera Utara who have supported this research process in terms of fundingv

References

[1] S. Singh, U. Ahuja, M. Kumar, K. Kumar, and M. Sachdeva, "Face mask detection using YOLOv3 and faster R-CNN models: COVID-19 environment," Multimed. Tools Appl., vol. 80, no. 13, pp. 19753–19768, 2021, https://doi.org/10.1007/s11042-021-10711-8

[2] W. Fang et al., "A deep learning-based approach for mitigating falls from height with computer vision: Convolutional neural network," Adv. Eng. Informatics, vol. 39, pp. 170–177, 2019, https://doi.org/10.1016/j.aei.2018.12.005

[3] Z.-Q. Zhao, P. Zheng, S. Xu, and X. Wu, "Object detection with deep learning: A review," IEEE Trans. neural networks Learn. Syst., vol. 30, no. 11, pp. 3212–3232, 2019, https://doi.org/10.1109/TNNLS.2018.2876865

[4] C. Cao et al., "An improved faster R-CNN for small object detection," Ieee Access, vol. 7, pp. 106838–106846, 2019, https://doi.org/10.1109/ACCESS.2019.2932731

[5] S. Sulistyawati et al., "Knowledge, attitudes, practices and information needs during the covid-19 pandemic in indonesia," Risk Manag. Healthc. Policy, vol. 14, p. 163, 2021, https://doi.org/10.2147/RMHP.S288579

[6] F. Kahar, G. D. Dirawan, S. Samad, N. Qomariyah, and D. E. Purlinda, "The epidemiology of COVID-19, attitudes and behaviors of the community during the Covid pandemic in Indonesia," structure, vol. 10, p. 8, 2020, https://doi.org/10.38124/IJISRT20AUG670

[7] U. Anand et al., "Novel coronavirus disease 2019 (COVID-19) pandemic: from transmission to control with an interdisciplinary vision," Environ. Res., vol. 197, p. 111126, 2021, https://doi.org/10.1016/j.envres.2021.111126

[8] F. Nurahmadi, F. Lubis, and P. I. Nainggolan, "Analysis Of Deep Learning Architecture In Classifying SNI Masks," J. INFORMATICS Telecommun. Eng., vol. 5, no. 2, pp. 473–482, 2022, https://doi.org/10.31289/jite.v5i2.6341

[9] P. Forouzandeh, K. O'Dowd, and S. C. Pillai, "Face masks and respirators in the fight against the COVID-19 pandemic: An overview of the standards and testing methods," Saf. Sci., vol. 133, p. 104995, 2021, https://doi.org/10.1016/j.ssci.2020.104995

[10] Al-Khowarizmi and Suherman, "Classification of Skin Cancer Images by Applying Simple Evolving Connectionist System," IAES Int. J. Artif. Intell., vol. 10, no. 2, pp. 421–429, 2021, https://doi.org/10.11591/ijai.v10.i2.pp421-429

[11] J. Aaron and T.-L. Chew, "A guide to accurate reporting in digital image processing–can anyone reproduce your quantitative analysis?," J. Cell Sci., vol. 134, no. 6, p. jcs254151, 2021, https://doi.org/10.1242/jcs.254151

[12] R. Herrera-Pereda, A. T. Crispi, D. Babin, W. Philips, and M. H. Costa, "A Review On digital image processing techniques for in-Vivo confocal images of the cornea," Med. Image Anal., vol. 73, p. 102188, 2021, https://doi.org/10.1016/j.media.2021.102188

[13] R. Syah and A.-K. Al-Khowarizmi, "Optimization of Applied Detection Rate in the Simple Evolving Connectionist System Method for Classification of Images Containing Protein," J. Ilm. Tek. Elektro Komput. dan Inform., vol. 7, no. 1, p. 154, 2021, https://doi.org/10.26555/jiteki.v7i1.20508

[14] A. Khowarizmi, Akhm, M. Lubis, and A. R. Lubis, "Classification of Tajweed Al-Qur'an on Images Applied Varying Normalized Distance Formulas," no. 3, pp. 21–25, 2020, https://doi.org/10.1145/3396730.3396739

[15] A. R. Lubis, S. Prayudani, Y. Y. Lase, and Y. Fatmi, "Similarity Normalized Euclidean Distance on KNN Method to Classify Image of Skin Cancer," in 2021 4th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), 2021, pp. 68–73, https://doi.org/10.1109/ISRITI54043.2021.9702826

[16] S. Suherman, F. Fahmi, Z. Herry, M. Al-Akaidi, and Al-Khowarizmi, "Sensor Based versus Server Based Image Detection Sensor using the

433 Mhz Radio Link," in 2020 4rd International Conference on Electrical, Telecommunication and Computer Engineering (ELTICOM), 2020, pp. 7–10, https://doi.org/10.1109/ELTICOM50775.2020.9230502
[17] I. H. Sarker, "Data science and analytics: an overview from data-driven smart computing, decision-making and applications perspective," SN Comput. Sci., vol. 2, no. 5, pp. 1–22, 2021, https://doi.org/https://doi.org/10.1007/s42979-021-00765-8
[18] X. Mou, X. Chen, J. Guan, B. Chen, and Y. Dong, "Marine target detection based on improved faster R-CNN for navigation radar PPI images," in 2019 International Conference on Control, Automation and Information Sciences (ICCAIS), 2019, pp. 1–5, https://doi.org/10.1109/ICCAIS46528.2019.9074588
[19] V. Kafedziski, S. Pecov, and D. Tanevski, "Detection and classification of land mines from ground penetrating radar data using faster R-CNN," in 2018 26th Telecommunications Forum (TELFOR), 2018, pp. 1–4, https://doi.org/10.1109/TELFOR.2018.8612117
[20] B. Zhu, X. Wu, L. Yang, Y. Shen, and L. Wu, "Automatic detection of books based on Faster R-CNN," in 2016 third international conference on digital information processing, data mining, and wireless communications (DIPDMWC), 2016, pp. 8–12, https://doi.org/10.1109/DIPDMWC.2016.7529355
[21] R. Gavrilescu, C. Zet, C. Foșalău, M. Skoczylas, and D. Cotovanu, "Faster R-CNN: an approach to real-time object detection," in 2018 International Conference and Exposition on Electrical And Power Engineering (EPE), 2018, pp. 165–168, https://doi.org/10.1109/ICEPE.2018.8559776
[22] Y. Chen, H. Wang, W. Li, C. Sakaridis, D. Dai, and L. Van Gool, "Scale-aware domain adaptive faster r-cnn," Int. J. Comput. Vis., vol. 129, no. 7, pp. 2223–2243, 2021, https://doi.org/10.1007/s11263-021-01447-x
[23] R. Meng, S. G. Rice, J. Wang, and X. Sun, "A fusion steganographic algorithm based on faster R-CNN," Comput. Mater. Contin., vol. 55, no. 1, pp. 1–16, 2018, https://doi.org/10.3970/cmc.2018.055.001
[24] J. Julham, M. Lubis, A. R. Lubis, A.-K. Al-Khowarizmi, and I. Kamil, "Automatic face recording system based on quick response code using multicam," IAES Int. J. Artif. Intell., vol. 11, no. 1, p. 327, 2022, http://doi.org/10.11591/ijai.v11.i1.pp327-335
[25] M. Meyer and G. Kuschk, "Automotive radar dataset for deep learning based 3D object detection," EuRAD 2019 - 2019 16th Eur. Radar Conf., no. January 2019, pp. 129–132, 2019.
[26] Y. Liu, P. Sun, N. Wergeles, and Y. Shang, "A survey and performance evaluation of deep learning methods for small object detection," Expert Syst. Appl., vol. 172, p. 114602, 2021, https://doi.org/10.1016/j.eswa.2021.114602
[27] B. Benjdira, T. Khursheed, A. Koubaa, A. Ammar, and K. Ouni, "Car detection using unmanned aerial vehicles: Comparison between faster r-cnn and yolov3," in 2019 1st International Conference on Unmanned Vehicle Systems-Oman (UVS), 2019, pp. 1–6, https://doi.org/10.48550/arXiv.1812.10968