JMES The International Journal of
Mechanical Engineering and Sciences
https://iptek.its.ac.id/index.php/jmes
e-ISSN 2580-7471

Automated Corrosion Detection on Steel Structures Using Convolutional Neural Network

Mohammad Khoirul Effendi[1*], Bara Atmaja[1], Arif Wahjudi[1], Dedi Budi Purwanto[2]

[1]Department of Mechanical Engineering, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia
[2]Department of Naval Architecture, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia

Received: 29 January 2023, Revised: 28 February 2023, Accepted: 2 March 2023

Abstract

Steel is a widely used material in industry and construction. The tensile and compressive strengths of steel are relatively high compared to other materials. On the other hand, low corrosion resistance is the main weakness of steel, which can encourage steel deterioration and fatal accidents for the user. Furthermore, regular visual inspection by a human should be performed to prevent catastrophic incidents. However, human visual inspection increases the risk of work accidents and reduces work effectiveness. Therefore, a drone with a camera is one solution to increase efficiency, increase security levels, and minimize difficulties or risks during corrosion detection. In this research, the drone was used to capture corroded video of a construction structure. The convolutional neural network (CNN) method was used to detect the location of the corroded images. This study was conducted on Surabaya's Petekan bridge using the Mobilenet V1 SSD pre-training model. In this study, the distance between a drone and the detected object varied between 1 and 2 m. Next, the drone speed was varied into 0.6 m/s, 0.9 m/s, and 1.3 m/s. As a result, CNN could detect corrosion on the surface of steel materials. The best accuracy was 84.66% with a minimum total loss value of 1.673 by applying 200 images, 200000 epochs, batch size at 4, learning rate at 0.001 and 0.1, the distance at 1 m, and drone speed at 0.6 m/s.

Keywords: Corrosion detection, steel, convolutional neural network, Mobilenets V1 SSD

1. Introduction

Steel is an alloy with iron as the primary element and carbon as the main alloy. The advantages of using steel include relatively higher tensile and compressive strengths, easy production processes, and easy installation in various environments. Although there are deficiencies in corrosion resistance, steel materials are still used in the construction sector around the world [1].

Corrosion is the process of material damage due to chemical or electrochemical reactions with the environment. All country in the world has faced the corrosion problem. In 1998, The National Association of Corrosion Engineers (NACE) noted that the total annual cost of corrosion in the United States was $276 billion or about 3.1% of GNI [2]. The corrosion degrades the durability of the steel material, causing accidents and significant economic losses. For example, on 14 August 2018, an accident occurred due to corrosion at the Morandi Bridge in Genoa, Italy. A 100-meter-long section of the flyover in Genoa collapsed, resulting in 39 deaths. Engineers revealed corrosion on the metal cables, which reduced the bridge's strength by 20% [3]. Steel construction must be maintained periodically; otherwise, it can reduce the structure's durability. Therefore, periodic inspection is part of the maintenance process to ensure the construction is safe and functioning correctly. Recently, periodic inspection has been conducted conventionally by human vision. However, this method has several drawbacks, namely: the reduced time efficiency for identifying problems that occur, increased risk of accidents, respiratory disorders due to dust, electric shock, extreme ambient temperatures, and noise that can reduce hearing ability. These drawbacks limit the human ability to inspect difficult and far areas [4].

Drones are unmanned aerial vehicles (UAVs) that function to increase efficiency, increase security, and minimize difficulties and risks during the inspection process. Drones and camera combinations help record images/videos. The results of these images can be processed using computer vision to read the information from the camera's images/videos. One application of computer vision is object detection, inspired by the human ability to see and understand objects visible to the sense of sight. The method that is widely used in object detection is the Convolutional Neural Network (CNN). Forkan et al. [5] observed how to detect structural corrosion with the drone's captured images using the CorrDetector framework.The

*Corresponding author. Email: khoirul_effendi@me.its.ac.id.

Figure 1. Uniform corrosion.

Figure 2. Drone DJI Mavic Air 2.

results of accuracy achieved 92% by evaluating the CorrDetector through empirical evaluation for structural corrosion analysis. Makantasis et al. [6] observed considerably different learning models, such as Artificial neural networks (ANNs), Support Vector Machines (SVMs), k-nearest neighbors (kNNs), and Classification trees (Ctree), for tunnel inspection. Using Quantitative Performance Metrics (QPM) as a data test set, CNN performed better than other learning methods. Ammar et al. [7] utilized YOLOv3 (You Only Look Once) to improve the performance of CNN for object detection.

CNN is part of a deep neural network with network depth more often implemented in images. The reasons for choosing CNN are the relatively high level of accuracy and produce significant output data in image processing [8]. CNN has been implemented in several cases, such as detecting rust on pipes [9], detecting defects on hull surfaces [10], and detecting biofouling inspection on Berth [11]. Wang et al. [12] conducted a corrosion detection study using CNN and used two methods: offline training and online training. Offline training used advanced modeling data to synchronize detection signals and speed maps. Meanwhile, online training used trained data from speed maps predicted in real-time with detection signals. The remaining thickness of the corroded structure was predicted using the dispersion curve. The success obtained from this study reached 82.73% and stated that the imaging performance was very good.

Based on the literature review, this study aims to implement CNN in detecting corrosion locations on steel materials using drones. This study determines the best variation of distance, drone speed, and CNN parameter settings for detecting corrosion with the highest accuracy. In this research, the DJI Mavic Air 2 was utilized for inspecting the corroded image with distance variations of 1 and 2 m. Next, the drone speed varied into 0.6 m/s, 0.9 m/s, and 1.3 m/s. The data used for training were corrosion images downloaded from Google Images and personal documents obtained at the Institut Teknologi Sepuluh Nopember Surabaya (ITS). The Mobilenet V1 SSD pre-training model functioned as CNN models with variations in hyperparameter settings, namely batch sizes of 4 and 8, as well as learning rates of 0.001 and 0.01.

1.1. Corrosion

The corrosion process on the metal surface occurs electrochemically, influenced by fluid velocity, temperature, and composition [13]. Uniform corrosion, as shown in Figure 1, appears on the entire surface of the metal with reduced size per unit of time. Uniform corrosion is the most destructive type since it has the potential to lose construction materials, reduce work safety, and cause environmental pollution [13].

1.2. Drone DJI Mavic Air 2

UAV is an uncrewed aircraft with several controllers: remotely controlled, semi-autonomous, autonomous, and combinations. Drones are an alternative to rotary and fixed-wing UAVs [14]. Mavic Air 2, as shown in Figure 2, is a drone with better camera, transmission, and flight performance capabilities than the previous version. The Mavic Air 2 features flight control, vision systems, infrared sensing, and an intelligent flight battery. In addition, the Mavic Air 2 has three flight modes: standard, sport, and tripod [15].

1.3. Convolutional Neural Network

Convolutional Neural Network (CNN) consists of neurons that have weight, bias, and activation functions. As the name implies, CNN uses a convolution process by moving a filter (kernel) on an input image. The computer processes new information from a collection of the input image and the selected filter [16].

1.4. MobileNet V1 SSD

Mobilenet V1 SSD is utilized to determine the location of the corroded image. MobileNet SSD is an object detection model consisting of SSD as the base model and MobileNet as the network model. SSD manages object detection by making bonding boxes and MobileNet as a feature extraction layer [17]. In MobileNet V1, the convolution box consists of depth-wise and point-wise convolutions to reduce the computation and size of the model. In addition, the performance of this model is faster seven times compared with standard convolutions [17].

Table 1. Confusion matrix.

		Actual values	
		TRUE	FALSE
Prediction Values	TRUE	TP (True Positive) Correct Result	FP (False Positive) Unexpected Result
	FALSE	FN (False Negative) Missing Result	TN (True Negative) Correct Absence of Result

Table 2. DJI Mavic Air 2 Drone specification.

Components	Specifications
Take-off weight	570 gr
Dimension	183 × 253 × 77 mm
Upward Velocity (max)	4 m/s (Sport Mode and Normal)
Downward Velocity (max)	3 m/s (Sport Mode and Normal)
Maximum Velocity	19 m/s (Sport Mode)
Flight Time (max)	34 mins (measured while flying at 18 kph in windless conditions)
Flight Distance (max)	18.5 km

Table 3. DJI Mavic Air 2 Drone camera specification.

Components	Specifications
Sensor	CMOS $\frac{1}{2}$
Lens	Effective Pixel: 12/48 MP
ISO	FOV: 84°
Video resolution	35 mm format equals 24 mm
Bitrate video max	Open: f/2.8

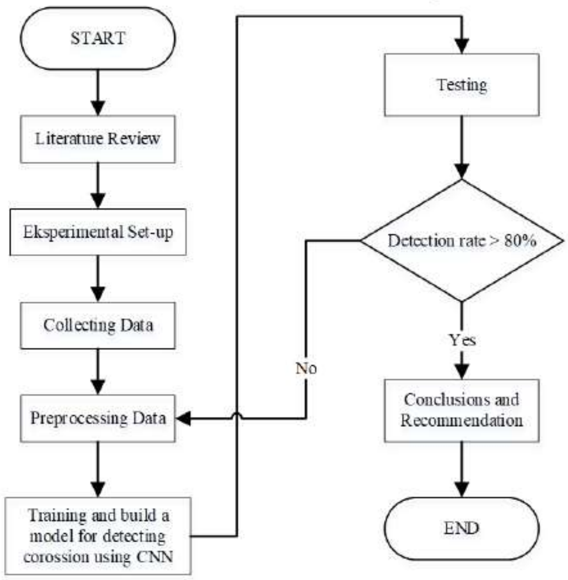


Figure 3. Experiment flowchart.

1.5. Confusion matrix

The confusion matrix is a table of classification results in actual and predicted values. The confusion matrix is a method for analyzing model performance in identifying data and objects [18]. The information used in the confusion matrix is as follows:

1. TP (True Positive): output where the model correctly predicts the positive class.

2. TN (True Negative): output where the model correctly predicts the negative class.

3. FP (False Positive): output where the model cannot predict the positive class.

4. FN (False Negative): output where the model cannot predict the negative class.

Table 1 shows the accuracy value of a model. Accuracy represents the value of a model in classifying data correctly. Equation (1) calculates the accuracy value.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (1)$$

2. Method

Figure 3 showed the research flowchart used for detecting corrosion areas using CNN.

2.1. Literature Review

The literature study aimed to explore the theoretical basis used in this study. The literature study began by searching for previous research journals that discuss similar topics, namely deep learning, object recognition, and convolutional neural networks (CNN). The results were collected and used as a reference in the writing and design of the experiments.

2.2. Experimental Set Up

The tools used in this study were Acer Aspire A315-41 Laptop and DJI Mavic Air 2 Drone with specifications in Tables 2 and 3.

Figure 4. Petekan Bridge, Surabaya.

2.3. Collecting Data

Two hundred images of uniform corrosion images were collected from the internet (Google Images) and Institut Teknologi Sepuluh Nopember (ITS) using a GoPro camera. These sources produced various types and positions of uniform corrosion. The image data set was divided into 170 for training data and 30 for testing data.

2.4. Preprocessing Data

Data preprocessing was divided into data augmentation and labeling to simplify the training and validation process. Data augmentation was needed to increase the number of images by performing various photo transformations. Furthermore, data labeling aimed to reduce the difficulty of the training process, so CNN could easily recognize corrosion objects.

2.5. Training Data

Data training aimed to find each image's features or characteristics and use it to update neurons' properties, such as bias and weight, during classification. The data training process carried out the following stages:

1. Labeling TFRecord images and using them as input for the training process.
2. Configuring the pipeline for setting the Convolutional Neural Network (CNN), determining the number of objects to be detected, assessing the training location, and image labeling. Next, the Mobilenet V1 SSD model configuration was employed to set CNN hyperparameters, such as learning rate, batch size, and steps.
3. Training for updating the CNN model parameters. This process generated automatic checkpoints created by Tensorflow in the form of graph tensors that stored information during the training process. The process was carried out until the best results with a slight loss were achieved. The process of database training could be explained as follows:
   (a) In the convolution stage, the input image passed through the convolution layer, and the convolution process was carried into a matrix with a specific size.
   (b) After that, the ReLU activation function aimed to change a negative value into a zero.
   (c) Reducing the dimensions of the feature map using the max pooling method.
   (d) Converting into a one-dimensional vector for input in the fully connected layer stage.
4. The updated CNN model was then used for the testing process.

2.6. Corrosion Detection

The Petekan Bridge, as shown in Figure 4, located on Jakarta Street, Surabaya City, East Java, was used as an object model for testing. At this stage, data collection was in the form of video using the DJI Mavic Air 2 drone, which flew horizontally for 5 meters to capture bridge videos using variations drone's speed and distance, as shown in Table 4.

Table 4. Drone parameter.

Test Variation		
Combination	Drone Velocity (m/s)	Drone Distance (m)
1	0.6	1
2	0.6	2
3	0.9	1
4	0.9	2
5	1.3	1
6	1.3	2

Table 5. Hyperparameter variation.

Hyperparameter variation		
Model	Batch Size	Learning Rate
1	4	0.001
2	4	0.01
3	8	0.001
4	8	0.01

Table 6. Hyperparameter results.

Hyperparameter results			
Batch Size	Learning Rate	Total Training Time	Minimum total Loss
4	0.001	25 hours 50 minutes17 seconds	1.647
4	0.01	25 hours 1 minute 58 seconds	1.389
8	0.001	50 hours 52 minutes 41 seconds	1.673
8	0.01	49 hours 1 minute 38 seconds	1.561

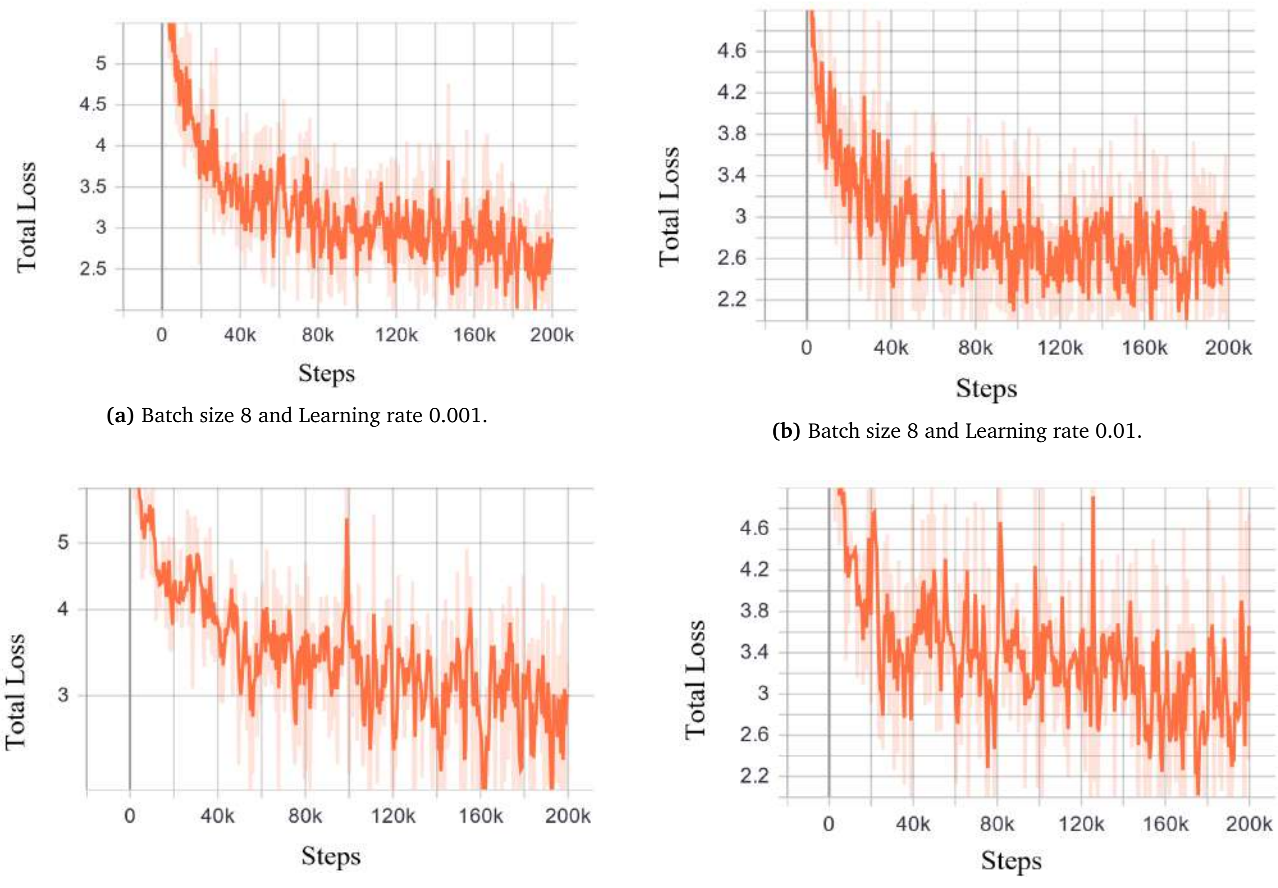
2.7. Setting Hyperparameter Convolutional Neural Network using the Mobilenet V1 SSD

Corrosion detection testing on video data used batch size and learning rate as the hyperparameter setting, as shown in Table 5.

3. Results and Discussion

3.1. Training result using the Mobilenet V1 SSD

This study conducted a training model for corrosion detection. The training used hyperparameter variations, such as batch size and learning rate values, to produce the best CNN model with the smallest minimum loss value using the Mobilenet V1 SSD. The selected batch sizes were four and eight. The selected learning rate values were 0.001 and 0.01. Table 6 shows the minimum total loss in each variable. The total loss value represented the model's performance detecting objects with a complete iteration of 200,000 steps to achieve convergence. Figure 5 shows the correlation between the total loss and steps, where the loss decreased as the step number increased.



(a) Batch size 8 and Learning rate 0.001.

(b) Batch size 8 and Learning rate 0.01.

(c) Batch size 4 and Learning rate 0.001.

(d) Batch size 4 and Learning rate 0.001.

Figure 5. Total loss graphs with hyperparameter variation.

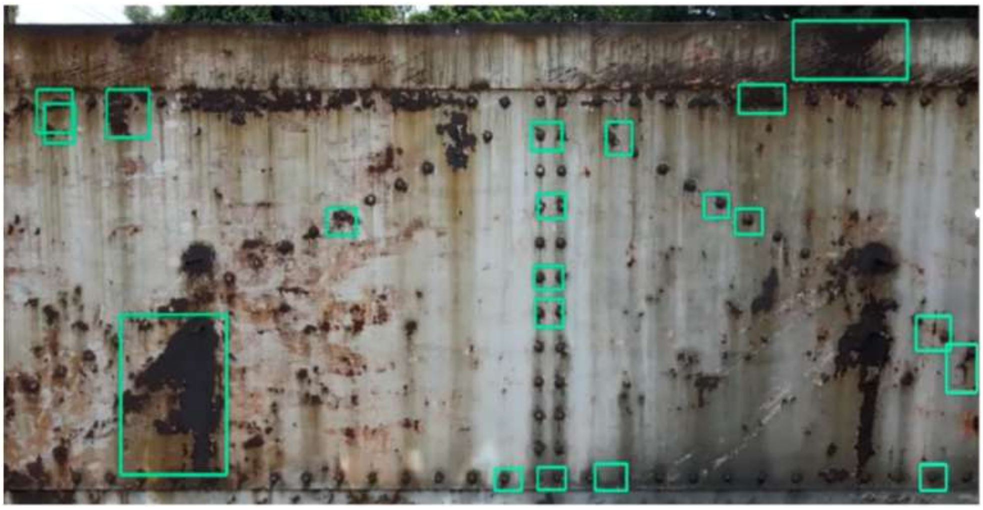


Figure 6. The result of corrosion detection using CNN.

3.2. Detection result using the Mobilenet V1 SSD

Corrosion detection used batch size four and height hyperparameter variations with learning rates of 0.001 and 0.01, variations of drone distance to objects of 1 and 2 m, and drone speed of 0.6 m/s, 0.9 m/s, and 1.3 m/s. Figure 6 shows the detection result using CNN, where green squares denote the corroded part.

3.2.1. Results of corrosion detection

Table 7 shows the test results with a batch size of four and learning rates of 0.001 and 0.01. Furthermore, Table 8 shows the test results with a batch size of eight and learning rates of 0.001 and 0.01. Based on this result, the accuracy decreased as distance and speed increased.

3.2.2. The comparison of batch size and learning rate

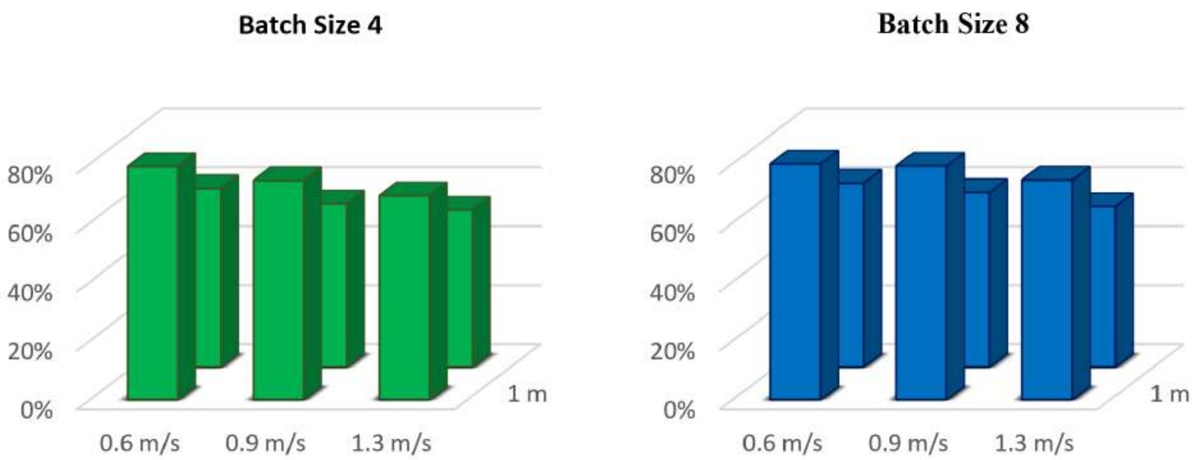
Figure 7 shows how the batch size value affected object detection accuracy. As the batch size value increased, the accuracy increased too. Testing with a learning rate of 0.001 on batch sizes eight and four had an accuracy difference of 1% to 6%. Meanwhile, testing with a learning rate of 0.01 on batch sizes eight and four had an accuracy difference of 4% to 7%. In addition, the learning rate value affected the object detection accuracy value. As the learning rate decreased, the accuracy increased. A comparison of testing on batch size four with a learning rate of 0.001 and 0.01 produced an accuracy difference of 9% to 17%. Whereas batch size eight, with a learning rate of 0.001 and 0.01, had an accuracy difference of 8% to 12%. Based on the results, the highest accuracy value of 84.66% was obtained on a batch size of eight and a learning rate of 0.001.

Table 7. Test result for batch size 4.

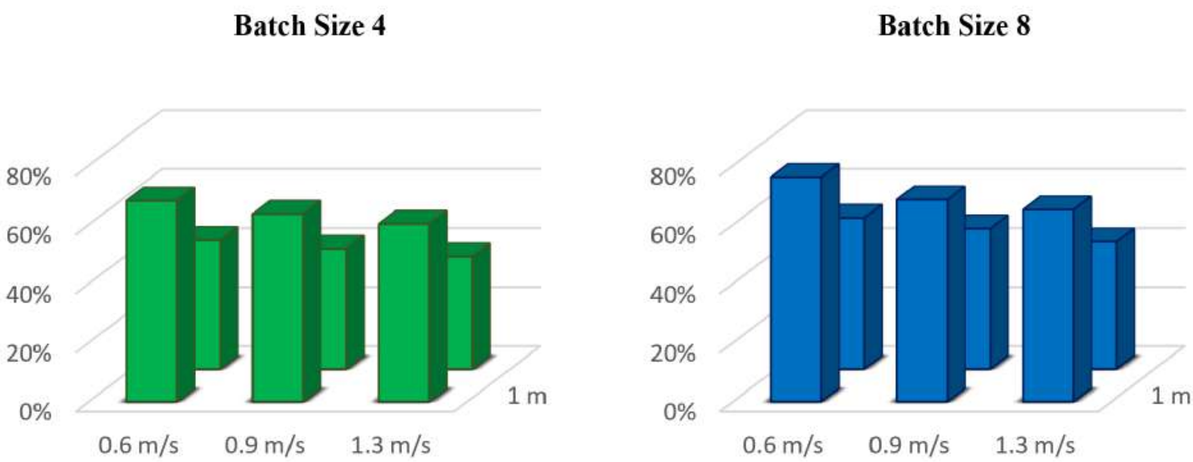
Accuracy value with batch size 4						
Learning Rate	Velocity (m/s)	Distance (m)	TP	FN	FP	Accuracy
0.001	0.6	1	129	32	2	79.14%
		2	100	61	4	60.61%
	0.9	1	122	39	3	74.39%
		2	92	69	5	55.42%
	1.3	1	115	52	5	69.28%
		2	89	72	6	53.29%
0.01	0.6	1	111	50	2	68.10%
		2	72	89	4	43.64%
	0.9	1	105	56	4	63.64%
		2	68	93	6	40.72%
	1.3	1	100	61	5	60.24%
		2	64	97	7	38.10%

Table 8. Test result for batch size 8.

Accuracy value with batch size 8						
Learning Rate	Velocity (m/s)	Distance (m)	TP	FN	FP	Accuracy
0.001	0.6	1	138	23	2	84.66%
		2	103	58	4	62.42%
	0.9	1	130	31	3	79.27%
		2	99	62	6	59.28%
	1.3	1	123	38	4	74.55%
		2	91	70	6	54.49%
0.01	0.6	1	124	37	2	76.07%
		2	85	76	5	51.20%
	0.9	1	113	48	4	68.48%
		2	79	82	5	47.59%
	1.3	1	108	53	5	65.06%
		2	72	89	6	43.11%



(a) Learning rate 0.001.



(b) Learning rate 0.01.

Figure 7. Learning rate graphs for batch size 4 and 8.

3.3. Analysis

This study aimed to detect corrosion areas using a drone camera and image processing that used the CNN method with the MobileNet V1 SSD model. This study intended to find the best hyperparameter variations, such as learning rate, batch size, distance, and drone speed, to obtain the highest accuracy and lowest total loss. The highest accuracy of 84.66% and lowest total loss of 0.057 were obtained on a batch size of four, a learning rate of 0.001, a distance between the drone and the detected object was 1 m, and a drone speed of 0.6 m/s. Decreasing batch size, reducing learning rate, reducing drone speed, and reducing distance should be considered significant parameters to increase corrosion detection accuracy.

4. Conclusion

The corrosion detection using the MobileNet V1 SSD with a variation of hyperparameter batch size, learning rate, drone distance, and speed was carried out. This research produces several conclusions, namely:

1. A Convolutional Neural Network (CNN) can be used to detect corrosion in steel materials with the Mobilenet V1 SSD as the pre-training model.
2. Decreasing batch size, reducing learning rate, reducing drone speed, and reducing the distance between the drone and the detected object should be considered to increase corrosion detection accuracy.
3. The best result, with an accuracy of 84.66% and a total loss of 1.673, was obtained at a distance of 1 m and a drone speed of 0.6 m/s.

The results of corrosion detection are strongly influenced by the number and type of corrosion images used in the training process. Therefore, increasing the number of images with various colors, texture variations, wind speed, humidity, and temperature should be performed to improve the detection accuracy of the CNN model. Moreover, selecting different CNN architectures (e.g., MobileNet V2 SSD, Faster R-CNN, etc.) should be considered to evaluate the performance of MobileNet V1 SSD in detecting corroded steel.

References

[1] N. Aini, "Corrosion behavior of AISI 1021 and AISI 304 in various acidic media," 2016.

[2] C. Verma, E. E. Ebenso, and M. Quraishi, "Corrosion inhibitors for ferrous and non-ferrous metals and alloys in ionic sodium chloride solutions: A review," *Journal of Molecular Liquids*, vol. 248, pp. 927–942, 2017.

[3] E. Y. Saputra, "Sebelum roboh, ahli sebut kabel jembatan Genoa, Italia berkarat." https://dunia.tempo.co/read/1119184/sebelum-roboh-ahli-sebut-kabel-jembatan-genoa-italia-berkarat, (accessed 4 February 2022), 2022.

[4] M. I. Akbari, "Interface and crack detection subsystem design in visual investigation of bridge structure conditions using autonomic drone," 2021.

[5] A. R. M. Forkan, Y.-B. Kang, P. P. Jayaraman, K. Liao, R. Kaul, G. Morgan, R. Ranjan, and S. Sinha, "Corrdetector: A framework for structural corrosion detection from drone images using ensemble deep learning," *Expert Systems with Applications*, vol. 193, p. 116461, 2022.

[6] K. Makantasis, E. Protopapadakis, A. Doulamis, N. Doulamis, and C. Loupos, "Deep convolutional neural networks for efficient vision based tunnel inspection," in *2015 IEEE International Conference on Intelligent Computer Communication and Processing (ICCP)*, pp. 335–342, 2015.

[7] A. Ammar, A. Koubaa, M. Ahmed, and A. Saad, "Aerial images processing for car detection using convolutional neural networks: Comparison between faster R-CNN and YOLOv3," *CoRR*, vol. abs/1910.07234, 2019.

[8] S. R. Dewi *et al.*, "Deep learning object detection pada video menggunakan tensorflow dan convolutional neural network," 2018.

[9] N.-D. Hoang and V.-D. Tran, "Image processing-based detection of pipe corrosion using texture analysis and metaheuristic-optimized machine learning approach," *Computational intelligence and neuroscience*, vol. 2019, 2019.

[10] C. Fernández-Isla, P. J. Navarro, and P. M. Alcover, "Automated visual inspection of ship hull surfaces using the wavelet transform," *Mathematical Problems in Engineering*, vol. 2013, 2013.

[11] S. Azhary, D. B. Purwanto, H. Nurhadi, B. Pramujati, M. K. Effendi, S. Widjaja, A. Wahjudi, and I. S. Arief, "Design of remotely operated vehicle prototype for ship biofouling inspection on berth," in *2021 International Conference on Advanced Mechatronics, Intelligent Manufacture and Industrial Automation (ICAMIMIA)*, pp. 223–228, 2021.

[12] X. Wang, M. Lin, J. Li, J. Tong, X. Huang, L. Liang, Z. Fan, and Y. Liu, "Ultrasonic guided wave imaging with deep learning: Applications in corrosion mapping," *Mechanical Systems and Signal Processing*, vol. 169, p. 108761, 2022.

[13] M. G. Fontana, N. D. Greene, *et al.*, *Corrosion engineering*. McGraw-hill, 2018.

[14] H. Eisenbeiß, "UAV Photogrammetry," 2009.

[15] DJI, "Mavic Air 2." Online, 2022.

[16] A. Geitgey, "Machine learning is fun part 4: Modern face recognition with deep learning," 2016.

[17] B. Sanjana, “SSD MobileNetV1 architecture.” Online, 2020.

[18] M. D. Kartikasari *et al.*, “Implementasi deep learning object detection rambu K3 pada video menggunakan metode convolutional neural network (CNN) dengan tensorflow (studi kasus: Rambu kesehatan dan keselamatan kerja (K3) jalur evakuasi dan alat pemadam api pada gedung FMIPA UII),” 2020.