Comparison of Faster R-CNN ResNet-50 and ResNet-101 Methods for Recycling Waste Detection

Puteri Nurul Ma'rifah Depart. of Electrical Engineering State Polytechnic of Malang Malang, Indonesia Moechammad Sarosa Depart. of Electrical Engineering State Polytechnic of Malang Malang, Indonesia Erfan Rohadi Depart. of Information Technology State Polytechnic of Malang Malang, Indonesia

Abstract: Inefficiencies in waste management contribute to the increasing amount of pollution in society, leading to public demands for better waste management and classification. Waste sorting is the beginning of the waste recycling process, which can help reduce the amount of waste in the environment. However, coupled with a lack of awareness of waste sorting due to minimal public education about waste management, the waste sorting system is still carried out manually using human power. Therefore, it is necessary to have a waste classification system to encourage people to manage their waste well. This research aims to design a tool to detect types of waste and classify them into three categories; metal, paper and plastic waste. The system can recognize the shape of trash images using a deep learning method developed using Faster R-CNN with ResNet-50 and ResNet-101 as the network architecture. This research began by collecting 250 datasets of metal, paper and plastic waste which were used as training data and test data in the testing process. The training data for the training process are 80, 120 and 200 datasets respectively. Test data for each training experiment, 30 and 50. Where 20 datasets in 50 test data are taken from the dataset for the training process. In each training process, the number of steps is carried out up to 3000, 4000 and 5000 steps, each of which has a total loss parameter. Based on the test results applied to the Faster R-CNN ResNet-50 and ResNet-101 methods, it produces an average F1 Score of 63% and 77% respectively. The best F1 Score is Faster R-CNN ResNet-101.

Keywords: Deep Learning, Object Detection, Waste Classification, Recycled Waste, Resnet50, Resnet10

1. INTRODUCTION

Waste is leftover goods/materials that are no longer used. Waste is produced from several processes, namely from human activities or natural ecosystems. There are many classifications of waste, one of which is classified into organic and non-organic waste. Based on research, Indonesia is the second largest waste producer after China. President Jokowi revealed that the floods experienced at the beginning of the new year 2020 were caused by damage to the ecosystem and ecology and because many people were still throwing rubbish carelessly. The Ministry of Environment and Forestry (KLHK), also said that the amount of waste generated nationally is 175,000 tons per day. It's pretty clear that waste is a big problem in Indonesia **[1]**

Object detection is concerned with detecting visual objects (such as people, animals, or cars) in digital images [2]. Object detection aims to detect target objects with theories and methods of image processing and pattern recognition, determine the semantic category of objects, and mark the specific position of target objects in the image [3].

Deep learning is part of Artificial Intelligence (AI) [4]. Deep learning allows computational models consisting of several processing layers to learn data representations with various levels of abstraction [5].

Convolutional neural networks (CNN) is one of the most powerful deep learning algorithms that has many applications in image classification, segmentation, and detection. [6]. R-CNN is a convolution-based algorithm or Region-based CNN. R-CNN is a combination of Region Proposal Network (RPN) and CNN proposed by Girshick et al in 2015.

2. METHOD



Figure 1. Research Stages

Literature studies are intended to study various reference sources such as journals, articles, papers on related research topics and themes. The topics or themes studied in this research are CNN, ResNet -50, and ResNet -101. Dataset is a collection of data needed in this research. The data needed in this research is plastic, paper and metal waste.

At the dataset collection stage, the dataset collected is images of plastic, paper and metal waste. Dataset collection was carried out by conducting independent data collection. This research carried out several stages of dataset training. The dataset used is 200 images. The data used for the training process includes 80 images for the first process, 120 images for the second process and 200 images for the third process. In each data training or training process is carried out using the ResNet-50 and ResNet-101 architectural models. Training is carried out up to step 5000. This step is carried out to test the algorithm of the system being designed. In the testing stage, the dataset used for test data is 30 images and 50 images for each training process. The 20 images in the test data of 50 images are obtained from the trained images. This step was taken to determine the reliability of the system in detecting recyclable waste, namely plastic, paper and metal waste. The steps taken are by entering test data into the system, then the system will detect and classify the waste contained in the image.

3. RESULT AND DISCUSSION

Computation time is the time required for the calculation process carried out by a computer when solving a problem using the algorithm used. From the experiments that have been carried out, there is computing time or time required during the training process and testing process for each image on the Faster R-CNN ResNet-50 and ResNet-101 network architectures..

3.1 Computation Time for Data Training Process (Data Training)

The computing time required during the training process for each architecture is shown in Table 1.

Network Architec-	Amount of Training	Average			
ture	Data	3000	4000	5000	
Faster	80	15m 24s	20m 26s	25m 19s	20m 23s
R-CNN	120	15m 15s	20m 41s	25m 03s	20m 26s
ResNet	200	15m 40s	20m 39s	26m 45s	21m 01s
50		20m 37s			
Faster	80	18m 52s	25m 18s	31m 55s	25m 22s
R-CNN	120	19m 39s	26m 20s	32m 42s	26m 14s
ResNet	200	19m 43s	26m 50s	32m 49s	26m 27s
101		26m 01s			

Table 1. Training Process Computation Time

Based on Table 1, it can be seen that the longest time required for the training process is the Faster R-CNN ResNet-101 architecture with an average time reaching 26 minutes 01 seconds, while the fastest is the Faster R-CNN ResNet network architecture. -50 with an average time of 20 minutes 37 seconds.

3.2 Computation Time Data Testing Process (Data Testing)

The computing time required during the testing process for each architecture is shown in Table 2.

Table 1	1. Test	Process	Comput	ation	Time
---------	---------	---------	--------	-------	------

Network		Test Co			
Architecture	I riai	3000	4000	5000	Average
	1	01m 47s	01m 43s	01m 41s	01m 44s
	2	01m 41s	01m 43s	01m 41s	01m 42s
Faster R-	3	01m 15s	01m 43s	01m 16s	01m 25s
CNN	4	01m 12s	01m 40s	01m 15s	01m 22s
ResNet 50	5	01m 18s	01m 18s	01m 44s	01m 27s
	6	01m 20s	01m 13s	01m 13s	01m 15s
		01m 29s			
	1	01m 42s	01m 18s	01m 22s	01m 27s
	2	01m 42s	01m 42s	01m 39d	01m 41s
Faster R-	3	01m 43s	01m 45s	01m 22s	01m 37s
CNN ResNet	4	01m 20s	01m 41s	01m 21s	01m 27s
101	5	01m 44s	01m 22s	01m 20s	01m 29s
	6	01m 41s	01m 40s	01m 42s	01m 41s
		01m 34s			

Based on Table 2, it can be seen that the longest time required for the testing process is the Faster R-CNN ResNet-101 architecture with an average time reaching 01 minutes 34 seconds, while the fastest is the Faster R-CNN ResNet network architecture. -50 with an average time of 01 minutes 29 seconds. ResNet-101 consists of 101 initial layers so it has a longer computing time than ResNet-50 which consists of 50 layers [7].

3.3 Recyclable Waste Detection Results

Recyclable waste detection results based on each network architecture model are shown in Figure 2 to Figure 9, In this table there are several images from the detection of recyclable waste based on the Faster R-CNN ResNet-50 and Faster R-CNN ResNet-101 architectural models.

3.3.1 Faster RCNN ResNet-50 Detection Results

The results of detecting recyclable waste on the Faster R-CNN ResNet-50 network architecture can be seen in the image below:



Figure 2. Faster RCNN ResNet-50 detection results for metal can waste



Figure 3. Faster RCNN ResNet-50 detection results for paper waste



Figure 4. Faster RCNN ResNet-50 Detection Results of plastic bottle waste



Figure 5. Faster RCNN ResNet-50 Detection Results for recyclable waste classification

3.2.1 Faster RCNN ResNet-101 Detection Results The results of detecting recyclable waste on the Faster R-CNN ResNet-101 network architecture can be seen in the image below:



Figure 6. Faster RCNN ResNet-101 detection results for metal can waste



Figure 7. Results of Faster RCNN ResNet-101 detection of paper waste



Figure 8. Faster RCNN ResNet-101 detection results for plastic bottle waste



Figure 9. Faster RCNN ResNet-101 detection results for recyclable waste classification

Network	Amount of	ount of Total Loss Step tO-			_
Architecture	Training Data	3000	4000	5000	x
	80	0.2101	0.2058	0.1808	0.1989
Faster R- CNN ResNet 50	120	0.2081	0.1988	0.1798	0.1956
	200	0.2003	0.1805	0.1536	0.1781
		0.1909			
	80	0.2485	0.2181	0.1883	0.2183
Faster R- CNN ResNet 101	120	0.22	0.1634	0.1873	0.1902
	200	0.1958	0.1037	0.1834	0.1610
		0.1898			

3.3	Total	Loss in	the	Training	Process

Table 3. Total Loss Results in Faster R-CNN ResNet-50 and ResNet-101 Network Architecture Models

Based on the results obtained in Table 3, the smallest average total loss was obtained in the Faster R-CNN ResNet-101 network architecture of 0.1898, while the largest average total loss was in the Faster R-CNN ResNet-50 network architecture model of 0.1909 [8].



Figure 10. Effect of Training Steps and Total Loss on Faster R-CNN ResNet-50 and ResNet-101 Network Architecture Models with 80 Training Data

Based on the results of Figure 10, it shows that the Faster R-CNN ResNet-50 and ResNet-101 network architecture types at steps 3000 to 5000 show decreasing loss values. However, the smallest average total loss was obtained in the Faster R-CNN ResNet-101 step 5000 network architecture [9].



Figure 11. Effect of Training Steps and Total Loss on Faster R-CNN ResNet-50 and ResNet-101 Network Architecture Models with 120 Training Data.



Figure 12. Effect of Training Steps and Total Loss on Faster R-CNN ResNet-50 and ResNet-101 Network Architecture Models with 200 Training Data

Based on the results of Figure 11 and Figure 12, it shows that the Faster R-CNN ResNet-50 network architecture type at steps 3000 to 5000 shows a decreasing loss value. However, the Faster R-CNN ResNet-101 network architecture type from steps 4000 to 5000 experienced an increase in the total loss value except from steps 3000 to 4000 which experienced a decrease in total loss. The smallest average total loss is obtained in the Faster R-CNN ResNet-101 network architecture at step 4000 [10].

The results obtained in Table 3 to Figure 12 show that the number of steps in the training process influences the total training loss value. As can be seen in the Faster R-CNN ResNet-50 network architecture type at steps 3000 to 5000, the loss value is decreasing. However, the Faster R-CNN ResNet-101 network architecture type from steps 4000 to 5000 experienced an increase in the total loss value except from steps 3000 to 4000 which experienced a decrease in total loss. The smallest average total loss was obtained in the Faster R-CNN ResNet-101 network architecture, while the largest average total loss was in the Faster R-CNN ResNet-50 network architecture model [11].

3.4 Total Loss in the Testing Process

Table 4. Results of Total Loss Testing on Faster R-CNNResNet-50 and ResNet-101 Network Architecture Models

Network Architecture	Ct.	Trial-					
	Step	1	2	3	4	5	6
	3000	0.493	0.492	0.433	0.430	0.413	0.408
Faster R- CNN BasNat 50	4000	0.448	0.445	0.406	0.401	0.386	0.386
inter et et	5000	0.439	0.436	0.371	0.370	0.367	0.364
Faster R- CNN ResNet 101	3000	0.254	0.251	0.221	0.219	0.211	0.211
	4000	0.226	0.223	0.203	0.202	0.187	0.189
	5000	0.226	0.224	0.219	0.211	0.194	0.193

Based on the results obtained in Table 4, it shows that the number of steps. The training process influences the total training loss value. As can be seen in the Faster R-CNN ResNet-50 network architecture type at steps 3000 to 5000, the loss value is decreasing. However, the Faster R-CNN ResNet-101 network architecture type from steps 4000 to 5000 experienced an increase in the total loss value except from steps 3000 to 4000 which experienced a decrease in total loss. The smallest average total loss was obtained in the Faster R-CNN ResNet-101 network architecture, while the largest average total loss was in the Faster R-CNN ResNet-50 network architecture model [12].

3.5 Pressicion, Recall, dan F1 Score

At the testing stage for recyclable waste detection, the parameters that will be created to group the detection results are True Positive (TP), False Positive (FP), and False Negative (FN). In this study, test results were obtained from 30 and 50 images which were used as test images for detecting recyclable waste objects with a maximum amount of training data of 200 images [13]. System capabilities based on the Faster R-CNN ResNet-50 and ResNet-101 architectural network models are shown in Table 5 to Table 6 respectively.

Table 5.	Experim	ental Re	sults on	the H	aster	R-CNN
Re	sNet 50	Network	Archite	ecture	Mod	el

Faster R-CNN ResNet-50								
т	Trial							
1	riai-	3000	4000	5000				
	ТР	13	14	15				
1	FP	14	13	13				
	FN	3	3	2				
	ТР	22	25	26				
2	FP	24	23	21				
	FN	4	2	3				
	ТР	14	19	20				
3	FP	11	6	5				
	FN	5	5	5				
	ТР	24	28	35				
4	FP	18	19	14				
	FN	8	3	1				
	ТР	20	22	23				
5	FP	8	4	4				
	FN	2	4	3				
	ТР	34	38	40				
6	FP	12	9	7				
	FN	4	3	3				

Table 6. Experimental Results on the Faster R-CNN ResNet-101 Network Architecture Model [14].

Faster R-CNN ResNet-101						
T		Step				
11	riai-	3000	4000	5000		
	ТР	14	15	17		
1	FP	13	13	11		
	FN	3	2	2		
	ТР	24	26	29		
2	FP	23	22	20		
	FN	2	2	1		
	ТР	16	20	18		
3	FP	12	8	8		
	FN	2	2	4		
	ТР	28	35	31		
4	FP	18	8	11		
	FN	4	7	8		
	ТР	21	24	23		
5	FP	5	4	3		
	FN	4	2	4		
	ТР	36	42	40		
6	FP	9	7	8		
	FN	5	1	2		

Based on the values obtained in Table 5 and Table 6, this can be done by calculating the precision, recall and F1 Score values. To calculate these values, you can look at equations (3.1), (3.2), and (3.3) [15]. The results of precision, recall and F1 Score calculations based on the ResNet-50 and ResNet-101 architectural network models with the Faster R-CNN algorithm are respectively shown in Table 7 and Table 8.

Table 7. Precision, Recall and F1 Score results	based	on
the ResNet-50 architecture network model	[16]	

Trial-	Step	Precision	Recall	F1 Score
	3000	48%	81%	60%
1	4000	52%	82%	64%
1	5000	54%	88%	67%
		Average		64%
	3000	48%	85%	61%
2	4000	52%	93%	67%
2	5000	55%	90%	68%
		Average		65%
	3000	56%	74%	64%
2	4000	62%	80%	70%
5	5000	80%	80%	80%
		71%		
	3000	57%	75%	65%
4	4000	60%	90%	72%
4	5000	71%	97%	82%
		73%		
	3000	71%	91%	80%
5	4000	85%	85%	85%
5	5000	85%	88%	87%
		Average		84%
	3000	74%	89%	81%
6	4000	81%	93%	86%
0	5000	85%	93%	89%
		85%		
	63%			

Table 8. Precision, Recall and F1 Score results based on the ResNet-101 architecture network model

Trial-	Step	Precision	Recall	F1 Score
	3000	52%	82%	64%
1	4000	54%	88%	67%
1	5000	61%	89%	72%
		Average		68%
	3000	51%	89%	65%
2	4000	54%	93%	68%
2	5000	59%	97%	73%
		Average		69%
	3000	57%	89%	70%
3	4000	71%	91%	80%
5	5000	69%	82%	75%
		75%		
	3000	61%	88%	72%
4	4000	81%	83%	82%
-	5000	74%	79%	77%
		77%		
	3000	81%	84%	82%
5	4000	86%	92%	89%
5	5000	88%	85%	87%
		Average		86%
6	3000	80%	88%	84%

Trial-	Step	Precision	Recall	F1 Score
	4000	86%	98%	91%
	5000	83%	95%	89%
		Average		88%
Average				77%

Based on Table 7 and Table 8, the ResNet-50 and ResNet-101 network architecture models using the Faster R-CNN algorithm obtained average F1 Score values of 63% and 77%, respectively. This shows that ResNet-101 performs better when used in the Faster R-CNN algorithm [17]. The F1 Score value for experiments with 50 test data is better than experiments with 30 test data for each method. This is because, 20 images from the test data of 50 images were obtained from images that had been trained. The best F1 Score value was obtained on the ResNet-101 network architecture model in experiment 6 and step 4000.

3.6 Comparison of Test Results of Two Architectures

Based on several test results that have been carried out, a table summarizing the overall data from the two network architecture models used can be created which is shown in Table 9 [18].

Table 9 Comparison of Test Results from the Faster R-CNN ResNet-50 and Faster R-CNN ResNet-101 Network Architecture Models

Architectural Models	Parameter	Mark
Faster R-CNN ResNet-50	Average Computational	20 minute 37
	Training Time	second
	Average Test Time	01 minute 29
	Computation	second
	Average Total Training Loss	0.1909
	Average F1 Score	63%
Faster R-CNN ResNet-101	Average Computational	26 minute 01
	Training Time	second
	Average Test Time	01 minute 34
	Computation	second
	Average Total Training Loss	0.1898
	F1 Score Average	77%

Based on Table 9, it can be seen that computing the Faster R-CNN Resnet-101 for both the training and testing processes takes longer than the Faster R-CNN Resnet-50 architectural model. However, the average total loss of Faster R-CNN Resnet-101 training is smaller than Faster R-CNN Resnet-50, namely 0.1898 and the average F1 score value is greater, reaching 77% [19]. On the other hand, the Faster R-CNN Resnet-50 computation for both the training and testing processes requires faster time compared to the Faster R-CNN Resnet-101 architecture model. However, the average total training loss produced was the largest, namely 0.1909 and the average F1 score produced was low, namely 63%.

3.7 Effect of Average Loss on Average F1 Score

The influence of the average total loss on the average F1 score results is shown in Figure 13. The graph of the average total loss and the average F1 score is the overall average of the two architectures used.



Figure 13. Graph of the influence of the average total loss on the average F1 Score

Based on Figure 13, the smaller the average total loss in the training process, the better the model is at learning object features so that the average F1 score value is greater. Recyclable waste detection errors occur because the object is too small or the image resolution is not good. Apart from that, detection errors are also caused by less complex data annotation processes [20].

4. CONCLUSION

From the series of tests that have been carried out, various things can be concluded that the two network architecture models studied resulted in an average F1 score of 63% and 77% respectively, where the best F1 score is found in Faster R-CNN ResNet-101, but with a longer computation time when compared to the Faster R-CNN ResNet-50 architecture model, where the average computation time for the training process of Faster R-CNN ResNet-50 and ResNet-101 is 20 minutes 37 seconds and 26 minutes 01 seconds respectively. While the average computation time for testing Faster R-CNN ResNet-50 and ResNet-101 is 01 minute 29 seconds and 01 minute 34 seconds respectively. The best F1 Score value was obtained on the ResNet-101 network architecture model in experiment 6 and step 4000, namely 91%.

5. REFERENCES

- F. F. Putra and Y. D. Prabowo, "Low resource deep learning to detect waste intensity in the river flow," *Bull. Electr. Eng. Informatics*, vol. 10, no. 5, pp. 2724– 2732, 2021, doi: 10.11591/eei.v10i5.3062.
- [2] Z. Zou, Z. Shi, Y. Guo, and J. Ye, "Object Detection in 20 Years: A Survey," pp. 1–39, 2019.
- [3] W. Zhiqiang and L. Jun, "A review of object detection based on convolutional neural network," *Chinese Control Conf. CCC*, pp. 11104–11109, 2017, doi: 10.23919/ChiCC.2017.8029130.
- [4] D. S. W. Ting *et al.*, "Artificial intelligence and deep learning in ophthalmology," *Br. J. Ophthalmol.*, vol. 103, no. 2, pp. 167–175, 2019, doi: 10.1136/bjophthalmol-2018-313173.
- [5] M. Aghilan, A. M. Kumar, M. T. Aafrid, N. A. Kumar, and S. Muthulakshmi, "Garbage Waste Classification Using Supervised Deep Learning Techniques," *Int. J. Emerg. Technol. Innov. Eng.*, vol. 6, no. 3, pp. 77–82, 2020.
- [6] R. Sultana, R. D. Adams, Y. Yan, P. M. Yanik, and M. L. Tanaka, "Trash and Recycled Material Identification using Convolutional Neural Networks (CNN)," *Conf. Proc. IEEE SOUTHEASTCON*, vol. 2020-March, 2020, doi: 10.1109/SoutheastCon44009.2020.9249739.
- [7] I. Salimi, B. S. B. Dewantara, and I. K. Wibowo, "Visual-based trash detection and classification system

for smart trash bin robot," Int. Electron. Symp. Knowl. Creat. Intell. Comput. IES-KCIC 2018 - Proc., pp. 378– 383, 2019, doi: 10.1109/KCIC.2018.8628499.

- [8] D. Atmajaya, N. Kurniati, W. Astuti, Y. Salim, and A. Haris, "Digital Scales System on Non-Organic Waste Types Based on Load Cell and ESP32," *Proc. - 2nd East Indones. Conf. Comput. Inf. Technol. Internet Things Ind. EIConCIT 2018*, pp. 308–311, 2018, doi: 10.1109/EIConCIT.2018.8878667.
- [9] Y. Dewilda and Julianto, "Kajian Timbulan, Komposisi, dan Potensi Daur Ulang Sampah Sebagai Dasar Perencanaan Pengelolaan Sampah Kawasan Kampus Universitas Putra Indonesia (UPI)," Semin. Nas. Pembang. Wil. dan Kota Berkelanjutan, vol. 1, no. 1, pp. 142–151, 2019, doi: 10.25105/pwkb.v1i1.5270.
- [10] M. F. Rahman and Bambang, "Deteksi Sampah pada Real-time Video Menggunakan Metode Faster R-CNN," *Appl. Technol. Comput. Sci. J.*, vol. 3, no. 2, pp. 117– 125, 2021, doi: 10.33086/atcsj.v3i2.1846.
- [11] Y. Wang and X. Zhang, "Autonomous garbage detection for intelligent urban management," *MATEC Web Conf.*, vol. 232, pp. 1–5, 2018, doi: 10.1051/matecconf/201823201056.
- [12] D. Patel, F. Patel, S. Patel, N. Patel, D. Shah, and V. Patel, "Garbage Detection using Advanced Object Detection Techniques," *Proc. - Int. Conf. Artif. Intell. Smart Syst. ICAIS 2021*, pp. 526–531, 2021, doi: 10.1109/ICAIS50930.2021.9395916.
- [14] X. Feng, Y. Jiang, X. Yang, M. Du, and X. Li, "Computer vision algorithms and hardware implementations: A survey," *Integration*, vol. 69, no.

April, pp. 309–320, 2019, doi: 10.1016/j.vlsi.2019.07.005.

- [15] Z. Zou, Z. Shi, Y. Guo, and J. Ye, "Object Detection in 20 Years: A Survey," pp. 1–39, 2019, [Online]. Available: http://arxiv.org/abs/1905.05055.
- [16] M. Sarosa, N. Muna, and E. Rohadi, "Performance of faster r-cnn to detect plastic waste," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 5, pp. 7756–7762, 2020, doi: 10.30534/ijatcse/2020/120952020.
- [17] A. Wahyudinata and H. B. Dirgantara, "Pengembangan Gim Edukasi 2D Pemilahan Sampah Daur Ulang Berbasis Android," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 20, no. 1, pp. 129– 138, 2020, doi: 10.30812/matrik.v20i1.860.
- [18] I. Febriadi, "Abdimas: Papua Journal of Community Service Pemanfaatan Sampah Organik Dan Anorganik Untuk Mendukung Go Green Concept Di Sekolah," vol. 1, no. 1, pp. 32–39, 2019.
- [19] A. Novianti and I. N. Muliarta, "Eco-Enzym Based on Household Organic Waste as Multi-Purpose Liquid," *AGRIWAR J.*, vol. 1, no. 1, pp. 12–17, 2021. M. Toğaçar, B. Ergen, and Z. Cömert, "Waste classification using AutoEncoder network with integrated feature selection method in convolutional neural network models," *Elsevier Meas.*, vol. 153, 2020, doi: 10.1016/j.measurement.2019.107459.
- B. Chitradevi and P. Srimanthi, "An Overview on Image Processing Techniques," Int. J. Innov. Res. Comput. Commun. Eng., vol. 2, no. 11, pp. 6466–6472, 2014, [Online]. Available: http://www.hindawi.com/isrn/sp/2013/496701/