E-ISSN: xxxx-xxxx
DOI: https://doi.org/10.64803/cessmuds.v1i1.15
Herdianto.

Detecting Monkey Pests Using Convolutional Neural Networks

Herdianto Herdianto[1*]

[1]Faculty of Computational Science and Digital Intelligence, Artificial Intelligence, Universitas Pembangunan Panca Budi, Medan, Indonesia
E-mail: [1]*herdianto0108047703@gmail.com
*E-mail Corresponding Author: herdianto0108047703@gmail.com

Abstract

Coffee is one of the agricultural products that has a fairly high value, making it one of the export commodities. Ironically, however, the amount of exports from 2021 to 2024 has continued to decline, partly due to monkey infestations. Efforts to prevent monkey attacks have been made by designing monkey detection based on PIR-HC-SR04 sensors, ultrasonic sensors, and combined with a microcontroller as a controller, but the results have not been able to distinguish between monkeys, cows, and humans passing through the plantation area. For this reason, the use of cameras will be tried as a means of sensing the plantation area and replacing the use of old sensors. Therefore, the purpose of this study is to optimize the detection of monkey pests based on digital images. To achieve this research objective, the convolution neural network (CNN) method is used. In order to obtain maximum results, the CNN will be trained beforehand so that it can correctly distinguish between monkeys, cows, humans, and so on. The results obtained are not yet optimal due to the occurrence of monkey pest attacks.
.

Keywords: *neural network convolution; monkey pests; coffee plants; classification.*
.

I. INTRODUCTION

Data from BPS in 2025 (Statistik, 2025) shows that coffee is one of the plantation commodities exported to countries in Asia and Europe, such as Japan, England, and Germany. However, Table 1 shows that coffee exports to Japan from 2021 to 2024 have continued to decline. This could be due to several factors, such as declining harvests or reduced demand. The decline in harvests could be due to attacks by pests such as monkeys, as has occurred in various regions in Indonesia.

Table 1. Indonesia's coffee export volume 2021-2024 (Statistik, 2025)

Country	2021	2022 ton	2023	2024
Japan	27,297.00	18,813.50	15,316.80	15,258.60
Singapore	5,377.00	6,573.20	5,522.30	4,394.80
Malaysia	29,180.70	26,117.30	22,690.90	28,687.10
England	12,259.50	20,778.00	4,339.00	4,029.30
German	13,334.90	36,976.40	9,460.40	16,442.20
Italy	24,590.00	24,006.20	18,122.00	6,549.90
Romania	509.4	340	372.7	403.3
Georgia	13,398.00	15,902.60	11,536.40	8,505.60
Netherlands	2,243.60	3,598.00	3,795.00	5,016.30
Belgium	14,434.00	22,179.80	3,430.90	21,298.20

To help combat monkey pests that attack various crops such as corn, coffee, oil palm, and others, research has been conducted to design a prototype for detecting and repelling monkey pests, as done by (Tamia & Zafia, 2022), (Herdianto & Nasution, 2023), (Hasibuan et al., 2023), (Hidayati et al., 2024). Research conducted (Tamia & Zafia, 2022) Designing a motion detector using an HC-SR04 PIR sensor combined with an NODE MCU as the controller. Then (Herdianto & Nasution, 2023) in their research continued what had been produced (Tamia & Zafia, 2022) by adding SIM 800 to the system designed as a short message sender (SMS) to the mobile phone number of the farm owner or designated person.

Next, there is (Hasibuan et al., 2023) who designed a monkey pest prototype using a HC-SR04 PIR sensor with an Arduino UNO controller and an additional power source derived from solar energy. Next is (Hidayati et al., 2024) who designed a monkey detection system using a PIR HC-SR04 sensor and equipped with an ultrasonic frequency generator to disrupt the monkey's hearing. The results of the four studies above indicate that the prototype designs produced are functional. In addition to experimental research, there are also other studies that are educational in nature for the community regarding monkey pest prevention, such as those conducted by (Dewi et al., 2023), (Sahri et al., 2024). Detecting an object that passes by can be done in various ways, such as with PIR sensors, ultrasonic sensors, infrared sensors, or cameras. The use of PIR sensors, ultrasonic sensors, and infrared sensors cannot distinguish the type of object passing by, whether it is a monkey, tiger, rabbit, cow, or other animal. However, this weakness can be overcome by using cameras for object detection (Arshad, 2021; P. Felzenszwalb et al., 2008, 2013; P. F. Felzenszwalb et al., 2010; R. Girshick et al., 2014; R. B. Girshick et al., 2011; Viola & Jones, 2001, 2004; Zou et al., 2019) because the features of objects such as monkeys, tigers, rabbits, cows, and

.

humans are different. The key is that in order for a computer to correctly distinguish between monkeys, cows, humans, and tigers, it must first be trained to recognize these objects.

Therefore, in this study, the author will attempt another method to detect monkey pests using a CNN approach.

II. RESEARCH METHODOLOGY

The steps used by the author to complete the research are shown in Figure 1.

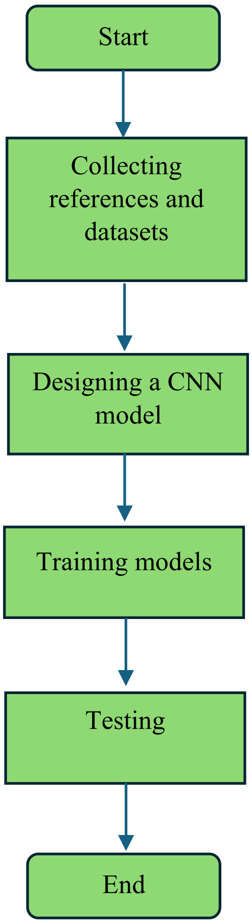


Figure 1. Research process flowchart

Start: The author begins the research.

Collecting references and datasets: The author collects references to articles related to the research topic from journals, proceedings, and reliable websites. In addition to collecting datasets on monkeys, tigers, humans, and cows that will be used during training and testing.

Designing a CNN model: creating a CNN model that will be used to recognize monkeys, birds, humans, and cows.

Training models: training the network model that has been created so that the network becomes intelligent and can be used.

Testing: The trained CNN network that has been declared good is tested with 80% training data and 20% test data to determine the accuracy of the prediction.

End: Drawing conclusions and completing the research.

III. RESULTS AND DISCUSSION

CNN is a development of a fully connected neural network. The architecture of this CNN network is as shown in Figure 2.

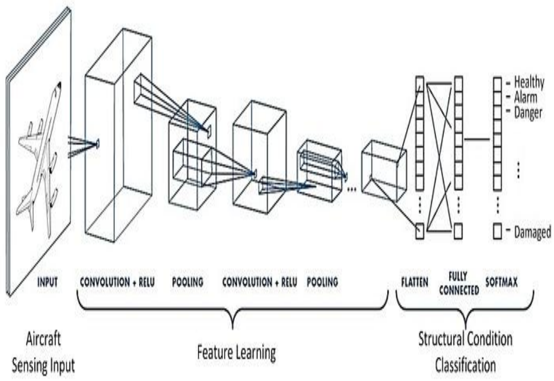


Figure 2. Architecture CNN (Convolutional Neural Network, n.d.)

CNN architecture consists of two parts: feature extraction and classification. Feature extraction contains convolution, ReLU, and pooling, which can number more than ten to hundreds depending on the size of the training data used. The feature extraction part functions to obtain the characteristics of each object. Furthermore, the classification contains flatten, fully connected, and softmax, which function to determine the class of the input object.

Convolusi

The objects to be classified are digital images represented in the form of matrices with specific sizes. For example, a 4*4 image with pixel values as shown in Table 2.

Table 2. 4x4 image

1	1	1	3
4	6	4	8
30	0	1	5
0	2	2	4

Next, the 4*4 image is multiplied by a filter of a certain size, for example 2*2.

Table 3. Filter size 2*2

1	0
0	1

so that a new image is formed with a size of 3*3 with pixel values as shown in Table 3.

1	1	1	3
4	6	4	8
30	0	1	5
0	2	2	4

7	5	9
4	7	9
32	2	5

x

1	0
0	1

= (1*1) + (1*0) + (4*0) + (6*1) = 7

Then, when all pixels have been processed, a new 3*3 image is formed. This process is called convolution.

Table 4. Pixel values of the new image resulting from convolution

5	17
32	24

ReLU

ReLU is one of the activation functions that can be used to determine the final value of the convolutional image. The ReLU function will cause the output to be 0 if the input value is negative or zero, and if the input value is positive, the output will be the same as the input value. So, the final value of the new image resulting from convolution with the ReLU activation function will be the same as the convolution.

Figure 3. Graph of the ReLU activation function

Table 5. New pixel value output by ReLU

7	5	9
4	7	9
32	2	5

Pooling

Pooling is used to reduce image input spatially using down-sampling operations. Many pooling methods can be used, including max pooling (which takes the largest value from all values in the output matrix of the activation function). In addition, there is also average pooling, but for this study, max pooling is used.

Table 6. Pixel values of the new image resulting from ReLU

1	4	3	2
2	5	10	17
10	32	24	20
15	20	11	6

Table 7. Pixel values of the new image resulting from max pooling

7	5	9
4	7	9
32	2	5

Image pixel values Table 6 were obtained from the process of using a 2x2 filter and stride 2.

Flatten

Flattening is the process of converting the output of a two-dimensional matrix image in the pooling layer into a single column (a one-dimensional vector).

Table 8. The process of forming a flatten

5	17
32	24

2 dimensions

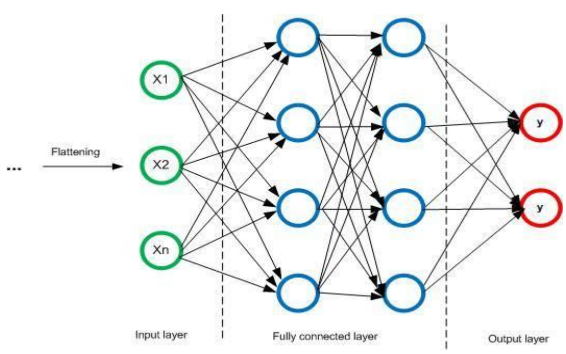
→

5
17
32
24

1 dimensions

Fully Connected

Fully connected is a form of neural network where images that have been converted into one dimension are used as input layer values for prediction.



E-ISSN: xxxx-xxxx
DOI: https://doi.org/10.64803/cessmuds.v1i1.15

Figure 4. Flattening relationship with fully connected

Softmax

Softmax is a function for calculating the probability of each class in the target for all possible target classes and for determining the target class from the given input. The advantage of softmax is that it works with an output probability range of 0 to 1.

Testing

Before CNN can be used for object classification, the CNN network must first be trained. The training data consists of 90 images, comprising 70 images of monkeys and 20 images of other animals. The test data consists of 15 images, comprising 10 images of monkeys and 5 images of other animals.

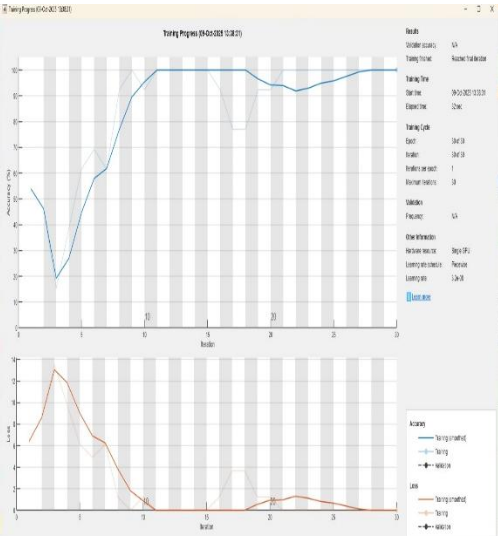


Figure 5. Graph data training accuracy results

Figure 5. shows the CNN training process, which achieved 98% training accuracy with 30 epochs.

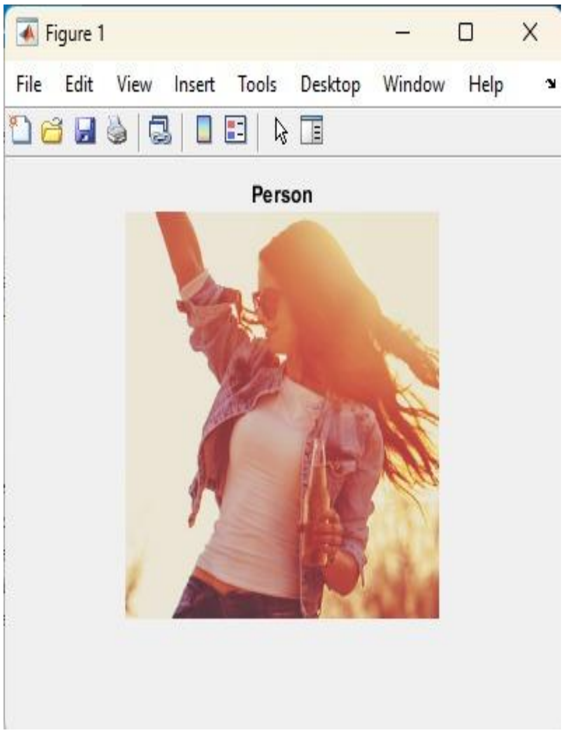


Figure 6. Testing with human images

Figure 6. shows the CNN testing process with human image objects, where the test results show that CNN can correctly determine that the image is of a human.

Figure 7. Testing with rabbit images

Figure 7. shows the CNN testing process with rabbit image objects, where the test results show that CNN can correctly determine that the images are rabbits.

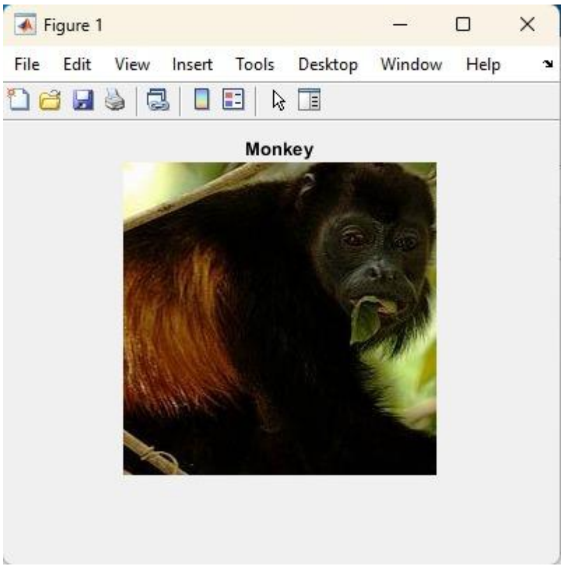


Figure 8. Testing with monkey images

Figure 8. shows the CNN testing process with monkey image objects, where the test results show that CNN can correctly determine that the image is a monkey object.

Figure 9. Testing with bird images

Figure 8. shows the CNN testing process with bird image objects, where the test results show that CNN can correctly determine that the image is a bird.

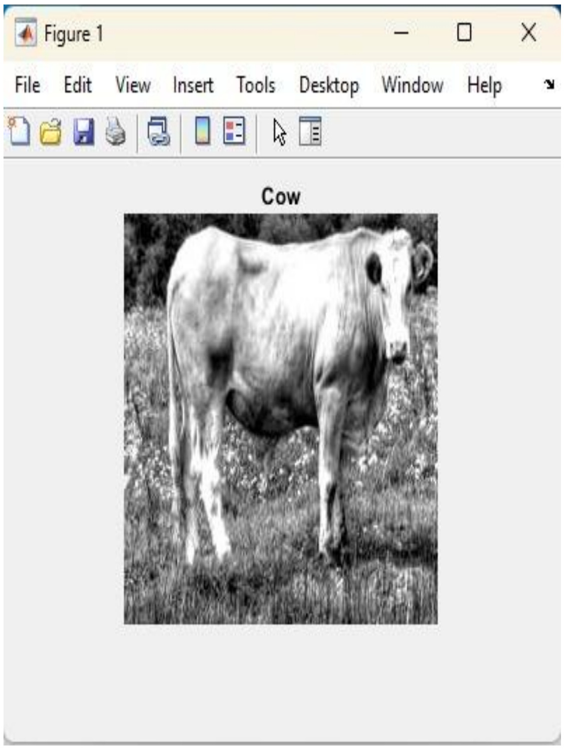


Figure 10. Testing with images of cows

Figure 10. shows the CNN testing process with images of cows, where the test results show that CNN can correctly determine that the image is of a cow.

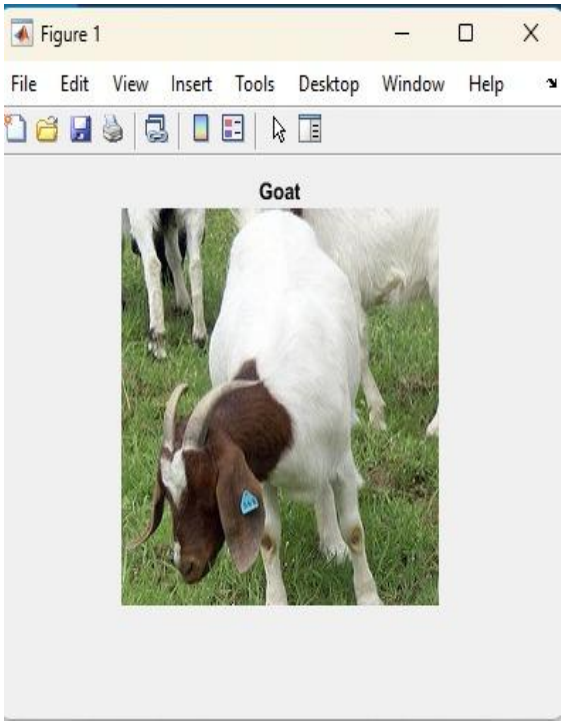


Figure 11. Testing with goat images

Figure 11. shows the CNN testing process with goat image objects, where the test results show that CNN can correctly determine that the image is a goat.

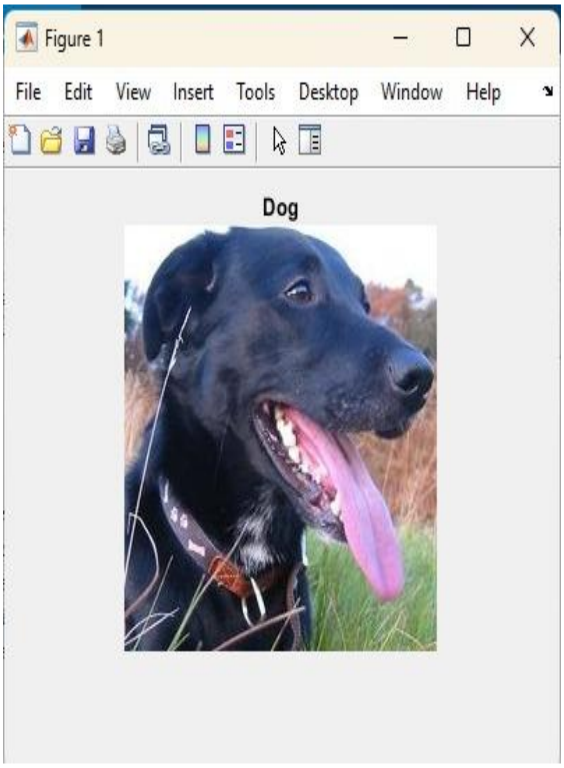


Figure 12. Testing with dog images

Figure 12. shows the CNN testing process with dog image objects, where the test results show that CNN can correctly determine that the image is a dog.

Figure 13. Testing with horse images

Figure 13. shows the CNN testing process with horse image objects, where the test results show that CNN can correctly determine that the image is a horse.

IV. CONCLUSION

After testing the CNN method in classifying image data of monkeys, rabbits, cows, goats, humans, birds, and dogs, it was found that the accuracy rate reached 97.2%. This means that this accuracy value is quite high and can be implemented.

Reference

Arshad, U. (2021). Object Detection in Last Decade - A Survey. *Scientific Journal of Informatics*, *8*(1), 60–70. https://doi.org/10.15294/sji.v8i1.28956

Convolutional Neural Network. (n.d.). The MathWorks, Inc. https://www.mathworks.com/discovery/convolutional-neural-network-matlab.html

Dewi, R. N. H., Ariyani, A. M., Widodo, R. C., Miharjo, E. S. R., Mutohhar, A., & Nursyahidah, F. (2023). Pencegahan Hama Kera sebagai Upaya Mengatasi Permasalahan Petani Alpukat Desa Sumberahayu. *Jurnal Pemberdayaan Masyarakat*, *2*(2), 80–88. https://doi.org/10.46843/jmp.v2i2.287

Felzenszwalb, P. F., Girshick, R. B., McAllester, D., & Ramanan, D. (2010). Object Detection with Discriminatively Trained Part Based Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(9), 1627–1645. https://doi.org/10.1109/TPAMI.2009.167

Felzenszwalb, P., Girshick, R., McAllester, D., & Ramanan, D. (2013). Visual Object Detection With Deformable Part Models. *Communications of the ACM*, *56*(9), 97–105. https://doi.org/10.1145/2494532

Felzenszwalb, P., McAllester, D., & Ramanan, D. (2008). A discriminatively trained, multiscale, deformable part model. *2008 IEEE Conference on Computer Vision and Pattern Recognition*, 1–8. https://doi.org/10.1109/CVPR.2008.4587597

Girshick, R. B., Felzenszwalb, P. F., & McAllester, D. (2011). Object Detection with Grammar Models. In Curran Associates Inc (Ed.), *Annual Conference on Neural Information Processing Systems 2011, NIPS 2011* (pp.

442–450). Curran Associates Inc. https://dl.acm.org/doi/10.5555/2986459.2986509

Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 580–587. https://doi.org/10.1109/CVPR.2014.81

Hasibuan, S. H., Andriani, T., Darmawan, I., & Hidayatullah, M. (2023). Rancang Bangun Alat Pengusir Hama Monyet Pada Ladang Jagung Design of a Monkey Pest Repellent Tool in Corn Fields. *Jurnal Elektronika, Sains Dan Sistem Energi*, *02*(02), 130–136.

Herdianto, H., & Nasution, D. (2023). Rancang Bangun Sistem Pengusir Hama Monyet Pada Perkebunan Kopi Menggunakan Arduino Uno. *Prosiding Seminar Nasional Unars*, 624–633. https://unars.ac.id/ojs/index.php/prosidingSDGs/article/view/3475

Hidayati, Q., Jamal, N., Veronika, M., & Dzulfiqar Raihan, A. Z. (2024). Sistem Keamanan Trafo Gardu Induk 150 KV Dari Hewan Kera di PT PLN UPT Kaltimtara. *Journal of Electrical Engineering and Computer (JEECOM)*, *6*(2), 417–426. https://doi.org/10.33650/jeecom.v6i2.9564

Sahri, M. S., Salsabila, A., Hasanah, L. M., Daulay, A. M., Al Hazza, M. R. R., & Rahmadani, K. (2024). Pencegahan Hama Kera Ekor Panjang Sebagai Upaya Mengatasi Permasalahan Petani di Dusun Tekik, Desa Ngloro, Kapanewon Saptosari, Kabupaten Gunungkidul. *Welfare : Jurnal Pengabdian Masyarakat*, *2*(3), 606–610. https://doi.org/10.30762/welfare.v2i3.1585

Statistik, B. P. (2025). *Ekspor Kopi Menurut Negara Tujuan Utama, 2000-2024*. https://www.bps.go.id/id/statistics-table/1/MTAxNCMx/ekspor-kopi-menurut-negara-tujuan-utama--2000-2024.html

Tamia, E. B., & Zafia, A. (2022). Rancang Bangun Prototype Pengusir Hama Kera Pada Perkebunan Berbasis Internet Of Things. *LEDGER : Journal Informatic and Information Technology*, *1*(1), 25–38. https://doi.org/10.20895/ledger.v1i1.775

Viola, P., & Jones, M. (2001). Rapid Object Detection using a Boosted Cascade of Simple Features. *Computer Society Conference on Computer Vision and Pattern Recognition*, 1–9. https://doi.org/10.1109/CVPR.2001.990517

Viola, P., & Jones, M. J. (2004). Robust Real-Time Face Detection. *International Journal of Computer Vision*, *57*(2), 137–154. https://doi.org/10.1023/B:VISI.0000013087.49260.fb

Zou, Z., Shi, Z., Guo, Y., & Ye, J. (2019). *Object Detection in 20 Years: A Survey*. *May*, 1–39. http://arxiv.org/abs/1905.05055