Reptile Recognition based on Convolutional Neural Network

Condro Kartiko, Agi Prasetiadi, Elisa Usada

Abstract*: Indonesian people are less interested in reptile animals. These are because most Indonesian people have the mindset that reptiles are difficult to tame and are focused on things about the ferocity of these animals in their natural habitat. Therefore it is necessary to have the means to identify reptile objects as one of the educational tools for introducing reptiles to the public. This research aims to produce a specialized Convolutional Neural Network model for recognizing reptile species. We also expand the model for recognizing another reptile species such as Snake, Crocodile, Turtle, and Gecko. Thousands of reptile images are being trained inside our model in order to obtain a kernel that can be used to automate reptile species recognition based on ordinary camera images. Our model currently reaches 64.3% accuracy for detecting 14 different species. Finally, as a suggestion for the next research, further enrichment especially from the background extraction process is needed to increase the accuracy of reptile detection.

Keywords: Reptile; Species Recognition; Automatic Detection; Convolutional Neural Network.

I. INTRODUCTION

Indonesia has various species of reptiles including snakes, crocodiles, lizards, turtles, etc. In Indonesia, several types of reptiles are often discovered around residential areas. Each reptile requires a different treatment, some just need to be left alone, while the others need to be handled specifically when it appears in the presence of humans. Therefore, classifying reptiles becomes one important step when people come into contact with reptiles. Furthermore, groups of reptile animals need to be studied and protected because of their benefits to the environment and humans. Reptiles are a significant component in ecosystems and as a reference indicator of the status of environmental damage. Reptiles have an essential role in human life and the environment, as objects of agriculture and animal husbandry, and used as supplements in the field of medicine [1] [2] [3]. Consequently, introducing reptiles to society is needed. Nowadays, there are reptile communities that are now scattered in almost every city in Indonesia. The existence of these reptile communities is very helpful for society to understand various species of reptiles and to help people interact with reptiles safely. However, information about the reptile community itself is not well known by the public at large. The community activities are still limited to temporary gatherings and short education.

Digital Data Reports 2019 states that there are currently 5.112 billion mobile device users and 67% penetration of the world's population [4]. Recently, Android users reach 87% of all smartphone users globally. Android is an operating system for Linux-based smartphones [5]. One of the advantages of Android compared to other smartphone operating systems is that Android is open-source code. Thus, it enables people to customize features that are not available in the Android operating system as desired. Through technology that is increasingly developing now, the utilization of mobile applications could be done more effectively, efficiently, and optimally. Therefore, this research attempts to provide a solution by developing a reptile recognition application using an Android smartphone. The use of Android smartphone technology enables the form widely available for free so that it can be used by anyone and at any time. All the conveniences provided by the Android operating system encourage many people to use it.

The reptile recognition application requires the implementation of a deep learning method to form a reptile object recognition model. Presently, deep learning has become one of the crucial topics in the world of machine learning because of its significant capabilities in modeling complex data such as imagery and sound. Machine learning is one of the disciplines of computer science that studies how to make a computer or machine have a deep intelligence. A computer or device must be able to learn to have intelligence. In other words, machine learning is a scientific field that contains computer or machine learning to be smart. The high focus of machine learning research is how to automatically recognize intricate patterns and make intelligent decisions based on data. Based on these capabilities, this research has the potential to improve the quality of data processing which is currently still not optimal. This research processes thousands of image data on dozens of reptile species. Based on these data, machine learning algorithms make it possible to produce reptile objects, recognition models.

Revised Manuscript Received on January 5, 2020.
***** Correspondence Author

Condro Kartiko*, Department of Software Engineering, Institut Teknologi Telkom Purwokerto, Indonesia.
E-mail: condro.kartiko@ittelkom-pwt.ac.id

Agi Prasetiadi, Department of Informatics, Institut Teknologi Telkom Purwokerto, Indonesia. E-mail: agi@ittelkom-pwt.ac.id

Elisa Usada, Department of Informatics, Institut Teknologi Telkom Purwokerto, Indonesia. E-mail: elisa@ittelkom-pwt.ac.id

The deep learning method that currently has the most significant results in image recognition is the Convolutional Neural Network (CNN). That is because CNN tries to imitate the image recognition system in the human visual cortex so that it can process image information. CNN is one of the technologies used extensively in image classification. Facial recognition using CNN produces an accuracy of 97.2% [6]. It can also recognize diseases in plants up to 99.53% accuracy [7].

This is also used to implement optical character recognition. It achieved 97% accuracy for the number of characters in the introduction of Bank Checks [8], 93.3% accuracy for Handwritten Character on Mobile Computing Devices, 96.5% accuracy for Chinese uppercase characters in the application of internet of things [9], 98.4% accuracy for Invoice Classification [10], as well as 99.97% accuracy for Japanese handwriting recognition. This paper investigates the possibility of recognizing reptile types based on CNN.

II. MODELLING

A. Reptile

In this research, three orders of reptile class, namely crocodilia (crocodiles), squamata (lizards and snakes), and testudines (turtles), were used as training objects for building a kernel recognizing reptile class via input image. Since squamata is the largest reptile order, the focus of data collection will be focused more on this order. Species that are used for training object consists of two species for the crocodilia order, one species for the family gekkonidae, 10 species for the suborder serpentes, and one species for the order testudines. All these animals' images are taken at zoo.

B. Design Application

This study focuses on software development which also includes experimentation on the system. Fig. 1 shows the flow of the study.

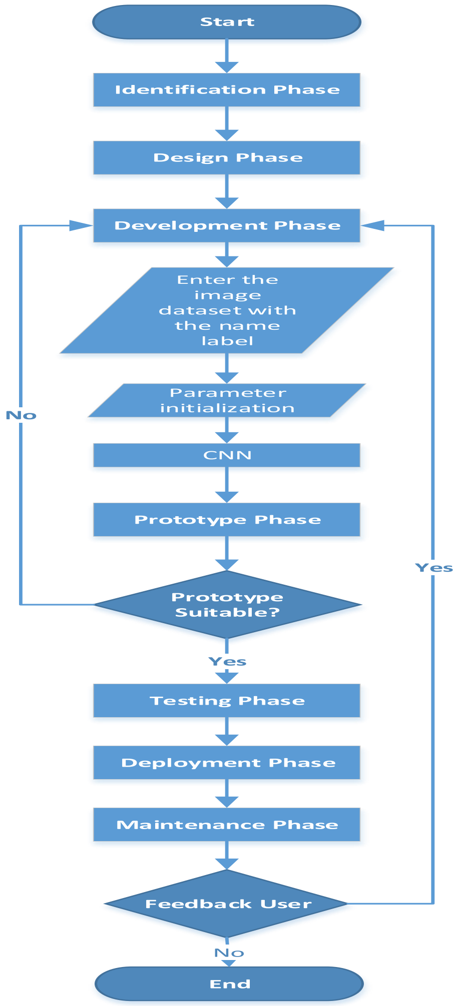

Fig. 1 The flow of the study

C. CNN

CNN is a type of Neural Network that has unique characteristics compared to other types of Neural Networks. CNN exploits feature extraction from an image by performing systematic extraction based on multi-convolution. Compared to ordinary NN, this technique is very efficient when applied in image recognition. In ordinary NN, millions of rows of data will be processed as they are in the hidden layer. But in CNN, only image features are processed in hidden layers. CNN is famous for its very high image classification level of up to 99% in recognizing plant diseases through leaf imagery [8].

Convolution is one of the basic feature extraction techniques in image processing using a simple small matrix. The bluring process, line detection, to image sharpening [9], are the basic effects that we usually get from the convolution process. In CNN, an image will be carried out a chain convolution process until a type of feature is extracted at the end of the calculation. Because not only one type of kernel is used in this image convolution, we will get a feature chain at the end of the calculation.

This feature chain will be obtained by ordinary NN to be classified.

Simply put, CNN is a combination of feature extraction in image pre-processing combined with the usual NN method for certain group classifications. Illustration 3 of CNN's architectural ideas, in general, can be drawn as Figure 1.

Overall, the CNN architecture consists of one input layer (I), three hidden layer types, and one output layer (O). In the hidden layer, there are 3 types of layers, namely the convolution layer, the sub-sampling layer, and the full connection layer. These three hidden layer types are *C* for the convolution layer, *S* for the sub-sampling layer, and *F* for the full connection layer where x is the index layer. No standard requires these layers to be arranged to form a particular architecture. So that it could be that in one architecture found the arrangement of *I-C-S-F-O* or $I\text{-}C_1\text{-}S_1\text{-}C_2\text{-}S_2\text{-}C_3\text{-}S_3\text{-}F\text{-}O$. The composition and parameters of each layer differs for each case. Following is the process that occurs at each layer.

Within covolution layer, the input image will be extracted its features by various kernel filter. The size of its filters are differents, ranging from 3 x 3 matrix, 5 x 5 matrix, and so on. The sub-sampling layer is used for helping the network to identify object even if the object position are slightly different in regards of pixel position. This part tolerates object various position during its training process. The fully connection layer performs basic neural network job for finishing the neuron weight calculation before deciding the input's type.

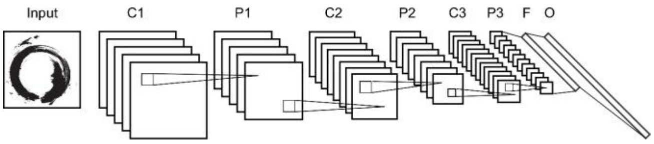

Fig. 2. CNN Architecture

D. Architecture

The following architecture is used to build the kernel of reptile recognition. This kernel aims to guess reptile type without background removal processing.

Table- I: Matrix Layer

Layer	w x h x n	Filter size
Input	64 x 64 x 1	-
C_1	60 x 60 x 64	5 x 5
P_1	30 x 30 x 64	2 x 2
C_2	26 x 26 x 96	5 x 5
P_2	13 x 13 x 96	2 x 2
C_3	9 x 9 x 128	5 x 5
P_3	1 x 1 x 128	5 x 5

The first convolution layer will receive 64 x 64 pixel input image. This resizing is considered safe since the resized input image still retains the detail of each animal pattern. Then, 64 number of filters sized 5 x 5 pixel will be convoluted with its input image resulting on 64 variation results on 60 x 60 pixel image. In this stage, 2 x 2 pixel pooling kernel will be applied. The data will be feed again into the second convolution layer with 96 filters with sized 5 x 5 pixel. Its output will be pooled again with 5 x 5 pixel kernel. Then the result will be feed again into the third convolution layer with 128 filters sized 5 x 5 pixel. The output will be pooled by 5 x 5 pixel pool kernel. The data will be connected into 256 nodes in its Fully Connected Layer. Then it is narrowed down again into 64 nodes. After that the output will be connected into its final stage, where 14 output nodes will be used to decides reptile's category.

III. RESULTS AND DISCUSSION

For each 14 species used in this research, 100 sampling of the images are collected, resulting in 1400 sampling of images. Then this dataset is divided into two groups, training images and testing images. The training images are then feed into kernel training process. After the kernel has been trained, then the remaining testing images are used to test the kernel accuracy in guessing each reptile type.

In this research, the model is trained for 24 epochs. Based on the simulation result, the following performance is obtain in relation of epoch vs training accuracy and testing accuracy.

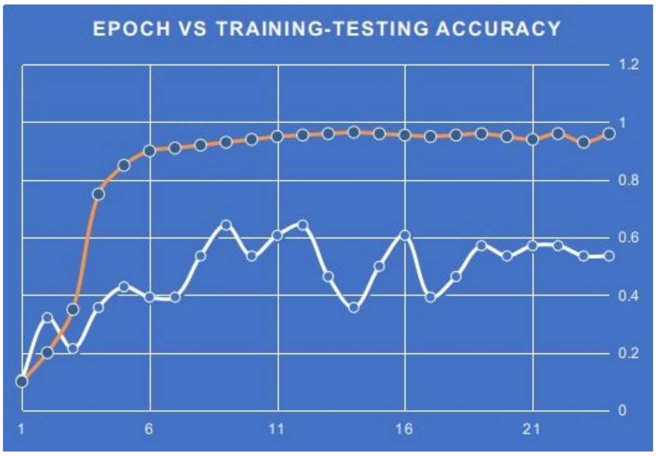

Fig. 3: Performance of training and testing accuracy on each epoch

The orange line represents the accuracy of training phase in each epoch. It can be seen that the network reach 90% training accuracy after 6th epoch. The training phase then shows stable learning outcome in each phase.

However, when the network being tested with the test data, the highest accuracy is obtained while the epoch are still at its 9th epoch, which shows 64.3%. Then the accuracy sways in between after that. It reachs the same accuracy as 9th epoch during its 12nd epoch. Then, the accuracy somehow consistenly decrease into nearby 53% accuracy. Thus it can be understood that the model experienced overfitting after its 9th epoch training.

Since overfitting is happened, the model actually learn the species characters quite well and fast within short epoch. However, the nature of the input images which are taken from the zoo may has homogenous pattern within the dataset. The result is that when the kernel is fed with same animal with completely different background, the result may not satisfy the real condition. Thus, image of animals with more various background and

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

condition could be needed to address this overfitting problem.

IV. CONCLUSION

A reptile recognition kernel based on CNN architecture has been built in this research. The model currently reaches 64.3% accuracy for detecting 14 different species. Higher input data is needed to increase the model's network accuracy. Another possibility to increase the accuracy is by doing further image refinement processing such as background removal or specialized reptile detection. In order to build specialized reptile detection, R-CNN families can be utilized to exploit this need, resulting more specialized image training with minimum background noise interference and higher accuracy in guessing reptile species.

ACKNOWLEDGMENT

We are grateful for the Institute for Research and Development, Institut Teknologi Telkom Purwokerto, who supported and gave an opportunity in this research. In addition, special thanks to the colleagues which aid in completing this research

REFERENCES

1. F. H. Pough, R. M. Andrews, J. E. Cadle, M. L. Crump, A. H. Savitzky and K. D. Wells, "Herpetology," New Jersey, Prentice-Hall, 1998, pp. 138-169.
2. G. R. Zug, Herpetology: An Introductory Biology of Amphibians and Reptiles, San Diego: Academic Press, 1993.
3. L. J. Vitt and J. P. Caldwell, Herpetology: An Introductory Biology of Amphibians and Reptiles. 3rd ed., San Diego: Academic Press, 2009.
4. S. Kemp, "Global Digital Insights," 31 January 2019. [Online]. Available: https://datareportal.com/reports/digital-2019-global-digital-overview. [Accessed 27 November 2019].
5. IDC, "Smartphone Market Share," IDC, 25 October 2019. [Online]. Available: https://www.idc.com/promo/smartphone-market-share/os. [Accessed 27 November 2019].
6. E. Zangeneh, M. Rahmati and Y. Mohsenzadeh, "Low resolution face recognition using a two-branch deep convolutional neural network architecture," *Expert Systems with Applications*, vol. 139, pp. 1-11, 2020.
7. K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Computers and Electronics in Agriculture*, vol. 145, pp. 311-318, 2018.
8. S. Srivastava, J. Priyadarshini, S. Gopal, S. Gupta and H. S. Dayal, "Optical Character Recognition on Bank Cheques Using 2D Convolution Neural Network," *Applications of Artificial Intelligence Techniques in Engineering*, vol. 697, pp. 589-596, 2019.
9. Y. Weng and C. Xia, "A New Deep Learning-Based Handwritten Character Recognition System on Mobile Computing Devices," *Mobile Networks and Applications*, vol. 24, no. 127, pp. 1-10, 2019.
10. Y. Yin, W. Zhang, S. Hong, J. Yang, J. Xiong and G. Gui, "Deep Learning-Aided OCR Techniques for Chinese Uppercase Characters in the Application of Internet of Things," *IEEE Access*, vol. 7, pp. 47043-47049, 2019.

AUTHORS PROFILE

Condro Kartiko, is a lecturer at Institut Teknologi Telkom Purwokerto, Department of Software Engineering, Faculty of Informatics. He graduates from Universitas Jenderal Soedirman majoring in Informatics Engineering as Bachelor Degree and from Universitas Indonesia majoring in Information Technology as Graduate Degree. He was involved in research and development of social media monitoring and analytics in industrial field. His publications cover topics such as software requirements, software design, software construction, software testing, and software quality. Currently his research interests range from software requirements, software testing, and software quality. He is an associate member of IEEE and APTIKOM (Indonesian Association of Higher Education in Informatics and Computing).

Agi Prasetiadi, is a lecturer at Institut Teknologi Telkom Purwokerto, Department of Informatics, Faculty of Informatics. He graduates from Bandung Institute of Technology majoring in Electronics as Bachelor Degree and from Kumoh National Institute of Technology majoring in IT Convergence as Graduate Degree. He was involved in research and development of camera calibration and 3D reconstruction in industrial field. His publications cover topics such as VLSI, error correction, bio-inspired technology, wireless sensor network, and real time communication system. Currently his research interests range from image processing, machine learning, and deep learning. He is an associate member of IEEE and APTIKOM (Indonesian Association of Higher Education in Informatics and Computing).

Elisa Usada, is a lecturer at Institut Teknologi Telkom Purwokerto, Department of Informatics, Faculty of Informatics. He graduates from Universitas Sebelas Maret majoring in Mathematics as Bachelor Degree and from Institut Teknologi Bandung majoring in Digital Media Technology and Games as Graduate Degree. Currently his research interests range from Artificial Intelligence and Augmented Reality. He is an associate member of IEEE and APTIKOM (Indonesian Association of Higher Education in Informatics and Computing).