Recognition of Balinese letters with convolutional neural network

Cite as: AIP Conference Proceedings **2542**, 030009 (2022); https://doi.org/10.1063/5.0103942
Published Online: 10 November 2022

Kadek Adi Sukma Wijaya and Wiwien Widyastuti

Recognition of Balinese Letters with Convolutional Neural Network

Kadek Adi Sukma Wijaya[a)] and Wiwien Widyastuti[b)]

Universitas Sanata Dharma, Paingan, Maguwoharjo, Depok, Sleman, Yogyakarta, Indonesia

[a)]Corresponding author: kandasorata69@gmail.com
[b)]wiwien@usd.ac.id

Abstract. Indonesia is one of the countries which is famous for having various kinds of distinctive cultures. Almost every region has a unique culture, one of which is the province of Bali which is known as "the island of the gods". Bali Province has very strong customs and culture. This study aims to train computers to recognize Balinese letters. Prior to this study, there had been several studies with similar aims to other recognition methods. This study uses the Convolutional Neural Network (CNN) method. The process of recognizing Balinese script has its own difficulty level in terms of collecting data that will be used for the training process for the Balinese script recognition system. Broadly speaking, the process of introducing Balinese script can be divided into 3 stages, namely; the Pre-Processing stage, the Training stage (Training), and the Testing stage. The introduction of Balinese script was carried out using the Tensorflow and Keras frameworks with a dataset of 1440 images. The dataset is divided into 80% training data and 20% testing data. This Balinese script recognition system is made using the Python programming language. The research was conducted using variations of the epoch in training, namely 50, 100, 150, 200, 250, 400, and 500. The highest testing accuracy was obtained at the number of epoch 500 with an accuracy rate of 89.93%.

INTRODUCTION

Balinese script is an ancestral heritage and a regional identity that must be preserved. The young generation of Bali today has difficulty learning Balinese writing or script. Learning about Balinese script needs to be done in order to stay sustainable. One effort can be made by utilizing technological advances to help learning Balinese script by creating an automatic Balinese script recognition system.

The automatic recognition of Balinese script can be done with machine learning technology. Many algorithms/models used in machine learning technology include the Hidden Markov Model, Neural Network, and Deep Learning [1]. Conventional machine learning techniques have limited capabilities in processing neural data in its raw form. Representation learning is a method that allows a machine to process raw data and automatically find the representation needed for detection or classification.

One of the machine learning methods that is often used is the convolutional neural network. The use of the convolutional neural network method for image recognition or classification has been carried out by many researchers. Lorentius *et al*. used a convolutional neural network to recognize Javanese script with the best test results reaching an accuracy of 97.52% [2]. Prihatiningsih *et al*. used a convolutional neural network to analyze the performance of numeric handwriting recognition. They obtained 100% accuracy in the 1000th training iteration [3]. All these studies show that the convolutional neural network method gives good results in pattern recognition/classification.

Based on the description above, a system to detect or recognize the handwriting of the Balinese Wreastra script will be created automatically using the Convolutional Neural Network.

The 4th International Conference on Science and Science Education (IConSSE 2021)
AIP Conf. Proc. 2542, 030009-1–030009-8; https://doi.org/10.1063/5.0103942
Published by AIP Publishing. 978-0-7354-4259-7/$30.00

LITERATURE REVIEW

Balinese Script

Balinese script is one of the traditional Indonesian scripts that developed on the island of Bali. This script was brought by Indians who embraced Hinduism to Indonesia through the politics of colony expansion, trade, religion, and culture [4]. In Indian culture itself, there is the oldest script called Karosthi which developed again into the Devonegari script and the Pallowo script.

The two scripts entered the archipelago through the Sriwijaya Kingdom. In line with the development of Hindu/Buddhist religion, the script then gave a social/cultural influence in the archipelago. In its development, the two scripts are transformed into a new form called the Kawi script. This Kawi script developed into Balinese script and Javanese script. There are two Balinese scripts, namely the Wreastra script and the Swalelita script. Wreastra is used to write general Balinese, while Swalelita is used to write Sanskrit. Balinese Wreastra script can be seen in Fig. 1.

FIGURE 1. Balinese Wreastra script [4].

Convolusional Neural Network

Convolutional Neural Network (CNN) is one type of artificial neural network that is commonly used in image data [5]. CNN can be used to detect and recognize objects in an image. Broadly speaking, CNN utilizes the convolution process by moving a convolution kernel (filter) of a certain size onto an image. The computer gets new representative information from multiplying that part of the image with the filter used.

Broadly speaking, CNN is not much different from an ordinary artificial neural network. CNN consists of neurons that have weight, bias and activation functions. The convolutional layer consists of neurons arranged in such a way

that they form a filter with length and height (pixels). The processes carried out by CNN are Convolutional Layer, Non-Linearity Layer (ReLU Layer), Pooling Layer, and Fully-Connected Layer.

The CNN algorithm consists of 3 processes, namely pre-processing, training, and testing. The training process is repeated until the stopping conditions are met. The termination criteria used in general are iteration and error. The iteration will stop when the iteration exceeds the specified iteration or when the error value is less than the specified.
Input: training dataset and number of iterations
Output: training model/network parameters
The steps in the training process are as follows:

1. Set the maximum number of iterations
2. The training data that has been labeled enters the convolution layer and the ReLU layer
3. The data (feature map) resulting from the convolution and ReLU process then enters the pooling layer with a down-sampling operation.
4. The results of the pooling layer, then enter the fully-connected layer or the last layer. Where the data in the form of an x-dimensional matrix is converted into a linear (one-dimensional) matrix, so that identification can be done more easily.
5. Update the weight until it reaches the desired value
6. Repeat steps 2 to 5 until the maximum number of iterations is reached.

METHOD

Block Diagram System

The recognition system of Balinese script can be divided into three stages. The first stage is pre-processing to cut Balinese script from a document into an image of one Balinese script. The second stage is the training process which aims to extract features from each segmented Balinese script. The third stage is the introduction stage (testing process) using the CNN method which is used to recognize Balinese script. The block diagram of the script recognition system is illustrated in Fig. 2.

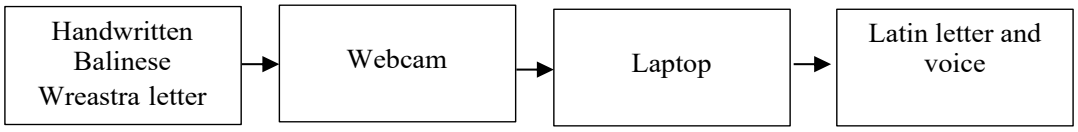

FIGURE 2. Block diagram system.

CNN Architecture

The CNN architecture is illustrated in Fig. 3. The network input is an image of the Balinese Wreastra script with a size of 143x143 pixels. The convolution layer uses a 1x1 kernel/filter with stride 1. The “dot” operation is performed every time the stride shifts between the inputs of the kernel/filter value. The output of the "dot" operation is usually called an activation map or feature map (downsampling). Feature Map functions to speed up computing because fewer parameters are updated and overcome overfitting. Feature maps that have been obtained from the convolution layer will be activated using the ReLU function.

The next layer is the pooling layer. The pooling layer is useful for maintaining the available features so that CNNs have spatial independence. Feature maps obtained after going through a combination of convolutions will be converted into a single feature vector through a flatten layer. The feature vector functions as an input to the fully connected layer which is an Artificial Neural Network architecture.

The last layer is the fully connected layer. Networks generally use fully connected layers, with each pixel considered a separate neuron. The final layer will completely contain as many neurons as the number of classes has to be predicted. During the fully-connected layer process, all activation neurons from the previous layer are connected to all neurons in the next layer just like a normal artificial neural network. Each activity from the previous layer needs to be converted into one-dimensional data before it can be connected to all neurons in the fully-connected layer. The stages in the CNN Architecture are illustrated in Fig. 3.

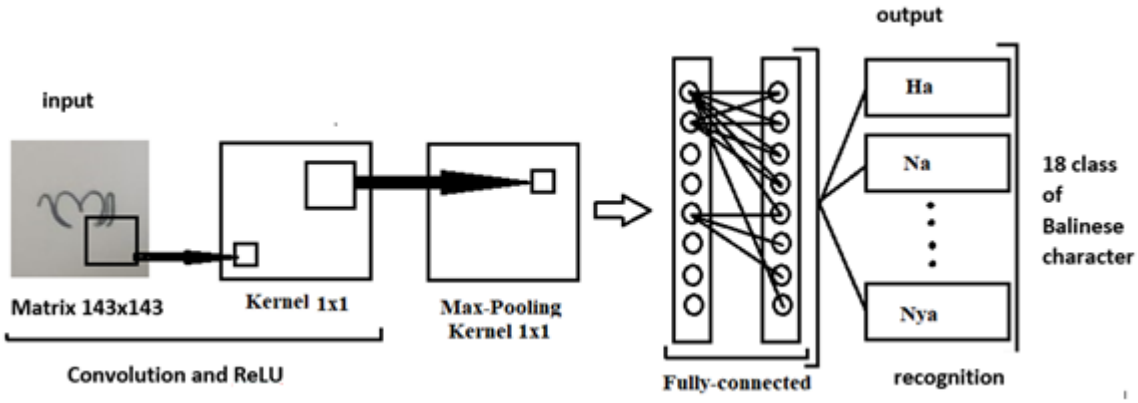

FIGURE 3. The CNN architecture.

Research Procedure

The research procedure consists of 3 stages as stage of creating data sets, modeling/training stage, and testing stage. The stage of making data sets begins with data collection. The data collected was in the form of 18 Balinese handwritten scripts from 20 different participants on a piece of paper for each participant. An example of a handwritten Balinese Wreastra script from a participant can be seen in Fig. 4. The writing is scanned and pre-processed. In this process, the scan results are cut into one by one Balinese script into 18 different scripts and labeled according to their spelling in Latin script. Eighteen Balinese scripts from 20 different participants, each rotated to 4 different positions, 0°, 90°, 180°, and 270° so that the total number of data is 1440 data. Furthermore, this 1440 data will be used 80% as training data and 20% as test data.

The next stage is the training stage. The training phase is used to obtain network models/parameters according to the architecture that has been designed. The flow diagram of the training process can be seen in Figure 5. The training process uses Python 3.7 software and hardware ASUSVivoBook Laptop with Intel Core i5-8250U Processor, 4GB DDR RAM with speeds up to 3.4 GHz, 1TB HDD, and Logitech*Webcam c170 (External webcam). The model studied was 7 variations of epochs, namely 50, 100, 150, 200, 250, 400, and 500 epochs.

The model that has been obtained at the training stage is then used at the testing stage. Testing is done in a way that is non real time and real time. In non-real time testing, the model is applied with input using 20% of the stored data set. Real time testing is carried out with input in the form of handwritten Balinese Wreastra script written on a piece of paper and photographed with a webcam connected to a laptop recognition system.

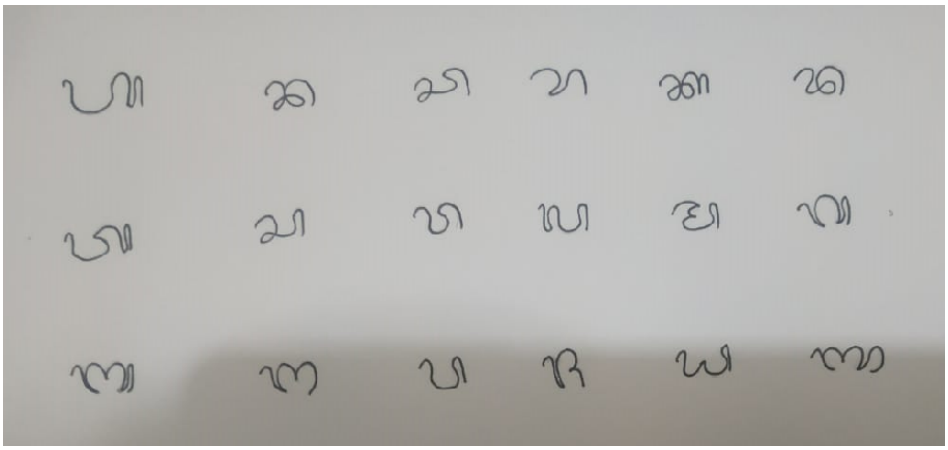

FIGURE 4. Example of handwritten Balinese Wreastra script

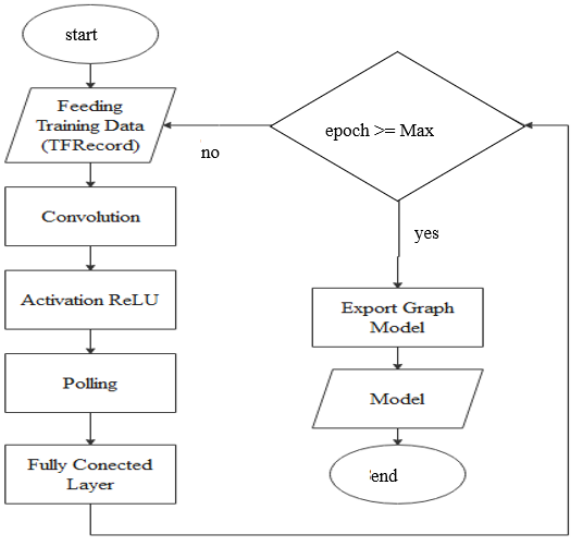

FIGURE 5. The flow chart of training proccess.

RESULT AND DISCUSSION

Testing is done in a way that is non real time and real time. In non-real time testing, the model generated from training using various epoch variations was tested using 20% of the data set. This test produces a level of accuracy and the precision. The level of accuracy shown in the graph in Fig. 6. The vertical axis shows the level of accuracy in percent. The horizontal axis shows the number of epochs used by the model. The level of accuracy is calculated by the formula of equation (1).

$$\text{Accuracy} = \frac{\text{Correct number of recogniton}}{\text{total data}} \times 100\% \quad (1)$$

The precison is calculated by the formula of equation (2).

$$\text{Precision} = \frac{\sum_{i=1}^{l} TP_i}{\sum_{i=1}^{l} (FP_i + TP_i)} \times 100\% \quad (2)$$

where TP is True Positive and FP is False Positive.

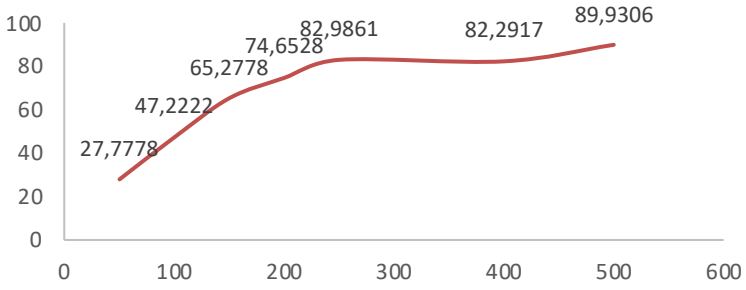

FIGURE 6. Testing Accuracy

Table 1 is a confusion matrix of non-real time testing using the 500 epoch model. From Table 1, it appears that the new model can recognize with an accuracy rate of 89.93% and the precision of 89.86%. This is probably because the shapes of these characters are similar to one another. For example, in recognizing the character ha, 12 characters are recognized correctly, 3 characters are recognized as ra, and 5 characters are recognized as la. If you look at it with

your eyes, it is true that the letters ha, ra and la have parts that are indeed similar. Likewise for ca character recognition, 3 characters are recognized correctly, but there are 5 characters recognized as sa. This is also because of the shape similarity of the two characters.

TABLE 1. Confusion matrix of non-real time testing using the 500 epoch model.

	ha	na	ca	ra	ka	da	ta	sa	wa	la	pa	ja	ya	nya	ma	ga	ba	nga
ha	12			2						5								
na		18																
ca			3					5										
ra				11				1	1									
ka		1			11		1											
da						18												
ta							17		1		3						1	
sa								17										
wa									12	1								
la										19								
pa											12							
ja												16						
ya												1	19	3		1		
nya														19				
ma															15			
ga																17		
ba						1		1						1			14	
nga					1													7

Real time testing is done by making a simple hardware system as shown in Fig. 7. The system consists of a laptop and a webcam. Webcam is used as input, laptop as a place of processing and displaying output. The system is made with Python programming. Figure 8 shows an example of the Graphical User Interface display when testing. The model used in real time testing is a model that is trained with 500 epochs because it has the highest level of accuracy.

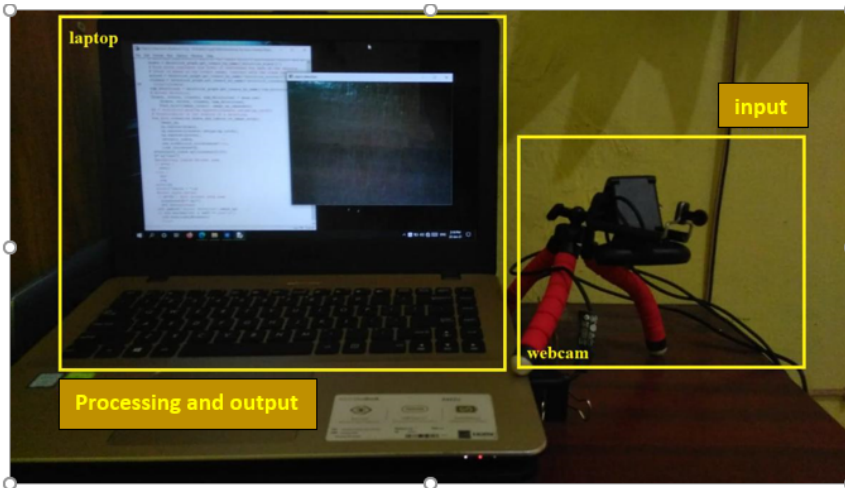

FIGURE 7. The hardware of recogniton system.

FIGURE 8. The GUI of recogniton system.

Testing is done by providing handwritten Balinese script photographed with a webcam as input from the system. Furthermore, the image is processed by the program on the laptop and the recognition results are displayed on the GUI. In addition to output in the form of Latin writing, the program also produces sounds according to recognized characters. The text entered into the system is rotated in 4 variations of positions, 0º, 90º, 180º and 270º. The overall real time test results are presented in the confusion matrix in Table 2.

TABLE 2. Confusion matrix of real time testing using the 500 epoch model.

	ha	na	ca	ra	ka	da	ta	sa	wa	la	pa	ja	ya	nya	ma	ga	ba	nga
ha	3			1														
na		3				1												
ca			1					3										
ra				4														
ka			1		1		2											
da						4												
ta		1					3											
sa								4										
wa									3						1			
la										4								
pa								1			3							
ja												4						
ya													3	1				
nya														4				
ma				1											3			
ga								1								3		
ba														3			1	
nga													1	1				2

From Table 2, the accuracy of the real time recognition system with 72 new test data can be calculated using equation 1. The calculation results get an accuracy value of 75% and the precison of 73.61%. The recognition system gives different accuracy results when the recognized object changes position. This is due to the lack of a number of

datasets for training, the number of epochs during the training process, taking the position of the detected object varies and many other possibilities (clarity of writing, similarity between one character and another, and light intensity during testing).

CONCLUSION

Based on the results of the analysis that has been done, some conclusions can be drawn as follows. The Balinese script recognition system on a digital image using the Convolutional Neural Network can work well. The more epochs used in training, the recognition accuracy tends to be higher. The level of accuracy of testing obtained in non-real time is 89.93% and in real time is 75% using the 500 epoch model. The precision of testing obtained in non-real time is 89.86% and in real time is 73.61% using the 500 epoch model.

ACKNOWLEDGMENTS

Thank you to the co-authors who have helped in completing this paper and to the Faculty of Science and Technology of Sanata Dharma University who have funded so that this paper can be published.

REFERENCES

1. W. Widyastuti, Media Teknika Jurnal Teknologi **12**, 115–123 (2017).
2. C. A. Lorentius, R. Adipranata, and A. Tjondrowiguno, "Pengenalan Aksara Jawa dengan Menggunakan Metode Convolutional Neural Network", Bachelor thesis, Universitas Kristen Petra, 2019.
3. S. Prihatiningsih, N. S. M., F. Andriani, and N. Nugraha, J. Ilmiah Teknologi Rekayasa **24**, 58–66 (2019).
4. NN, "*Mengenal Aksara Bali, dari Sejarah hingga Jenisnya Lengkap*", available at https://news.detik.com/berita/d-5078465/mengenal-aksara-bali-dari-sejarah-hingga-jenisnya-lengkap (accessed on 28 March 2020).
5. M. E. Paoletti, J. M. Haut, J. Plaza, and A. Plaza, ISPRS J. Photogramm. Remote Sens **158**, 279–317 (2017).
6. A. Deshpande, "*A Beginner's Guide To Understanding Convolutional Neural-Network*," available at https://adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks/.
7. NN, "*Machines that can see: Convolutional Neural Networks*" available at https://shafeentejani.github.io/2016-12-20/convolutional-neural-nets/ (accessed on 10 August 2020).
8. NN, "*Pengenalan Deep Learning Part 7 : Convolutional Neural Network (CNN)*", available at https://medium.com/@samuelsena/pengenalan-deep-learning-part-7-convolutional-neural-network-cnn-b003b477dc94 (accessed on 2 April 2020)
9. NN, "*Pengertian Webcam Beserta Fungsi dan Cara Kerjanya*" available at https://www.galerinfo.com/pengertian-webcam/ (accessed on 10 Agustus 2020).
10. NN, "*Apa Itu Python dan Fungsinya di Dunia Nyata*?" available at https://medium.com/purwadhikaconnect/apa-itu-python-dan-fungsinya-di-dunia-nyata-d5b533117c63. (accessed on 28 March 2020).
11. Noprianto, *Python & Pemrograman Linux* (ANDI, Yogyakarta, 2002).
12. Wikipedia, "*Tensor Flow*," available at https://en.wikipedia.org/wiki/TensorFlow/ (accessed on 28 March 2020).
13. Keras *Documentation*., "*Keras: The Python Deep Learning Library*," available at https://keras.io/ (accessed on 28 March 2020).

Faculty of Science and Mathematics
Universitas Kristen Satya Wacana

Certificate

Award to:

Ir. Wiwien Widyastuti, S.T.,M.T.

as

Participant (Co-Author)

at the 4th International Conference on Science and Science Education (IConSSE) 2021
organized by Faculty of Science and Mathematics-Universitas Kristen Satya Wacana

Salatiga, September 7-8, 2021

Dr. Adi Setiawan, M.Sc.
Dean-Faculty of Science and Mathematics

Dr. Adita Sutresno, M.Sc.
Chairman-The 4th IConSSE 2021