Design and Simulation Transfer Learning on Image Processing for Determining Condition of Robot Based on Neural Network

Asep Khoerudin
Department of Electrical Engineering
UIN Sunan Gunung Djati Bandung
Bandung, Indonesia
asep@student.uinsgd.ac.id

Edi Mulyana
Department of Electrical Engineering
UIN Sunan Gunung Djati Bandung
Bandung, Indonesia
edim@uinsgd.ac.id

Rina Mardiati
Department of Electrical Engineering
UIN Sunan Gunung Djati Bandung
Bandung, Indonesia
r_mardiati@uinsgd.ac.id

Aan Eko Setiawan
Department of Electrical Engineering
UIN Sunan Gunung Djati Bandung
Bandung, Indonesia
aaneko37@gmail.com

***Abstract*— Neural Networks or Artificial Neural Networks (ANN) are an interesting topic in the last decade. This is due to the ability of ANN to mimic the nature of the input system. ANN is an information processing system that has similar characteristics to biological neural networks. This research will discuss the application of the Neural Network method for image processing which aims to classify the condition of the robot. The input for dataset are RSTRAIGHT, RTURNRIGHT, and RTURNLEFT, with a total dataset of 300 images. The image was taken using a camera with various conditions of the robot in various positions. The output of simulation is a learning rate precision which will use as a reference for whether the system can recognize the robot's condition or not. In this study, the learning rate precision output was 98% which was obtained in the 12th iteration with an epoch of 50 and a batch size of 10 using the Adam optimizer.**

Keywords*— *Neural Network, Image Processing, Transfer Learning.

I. INTRODUCTION

Neural Networks or Artificial Neural Networks (ANN) is an interesting topic in the last decade. This is due to the ability of ANN to mimic the nature of the input system. ANN is an information processing system that has characteristics similar to biological neural networks. Several applications has implemented ANN method such as image processing, pattern recognition, medicine, speech recognition, and business [1]. Backpropagation is one of algorithm that very suitable for image processing because of its advantages in repeating the classification process in producing output if an error occurs. Nevertheless, the low quality of image in image processing was one of the main problems that have an impact on the level of accuracy of the classification of an image [2].

Neural networks are built from many nodes/units connected by direct links. The link from one unit to another is used to propagate the activation from the first unit to the next. Each link has a numeric weight. This weight determines the strength and marker of connectivity [3]. Transfer learning is a technique or method that utilizes a model that has been trained on a dataset to solve other similar problems by using it as a starting point, modifying and updating its parameters so that it fits the new dataset [4].

Image processing is a form of signal processing where the input is in the form of images and transformed into other images as output with certain techniques. Image processing was performed to correct image signal data errors that occur due to transmission and during signal acquisition, as well as to improve the quality of image appearance so that it is easier to interpret by the human visual system either by manipulating it [5].

The research about image classification which taken from web or the internet has been done by Zhu et al. [6]. Zhu et al. use the image from web and try to determine the maximum possible learning rate through the transfer learning method. Meanwhile, Thai et al. try to classification the images using a combined method, namely using a Support Vector Machine and Artificial Neural Network which produces a precision level of 86% [7].

Another research was conducted by Markovic et al. that using a camera that will follow a moving object, changing the direction into an image that will be used as a determinant of the movement of the servo direction [8]. Another research used fisheye lenses to classify moving objects and stationary objects with a deep neural network method [9].

In this study, one of the implementations of Artificial Neural Networks will be discussed further in the field of digital image processing. This model will be made to determine the position of the robot direction with a camera device that is used as a sensor. The purpose of using this Artificial Neural Network is to determine the output in the form of a precision learning rate with the highest result as a reference to whether the system can identify the position of the robot or not. This study uses the python programming language which will be written and executed using Google Colab software.

The paper is organized as follows. Section 1 gives a general introduction, while Section 2 described the design and implementation. In Section 3 describes the results. Finally, Section 4 presents the conclusion and future works.

978-1-6654-4402-6/21/$31.00 ©2021 IEEE

II. Design and Simulation

A. Block Diagram System

The main process in this research is divided into three step, which are the design of input system, the processing system, and the output as shown in Figure 1. The process of data input, we use a robot as an object, which the image will be caught by camera. Then, using neural network method we obtained the learning rate of this system.

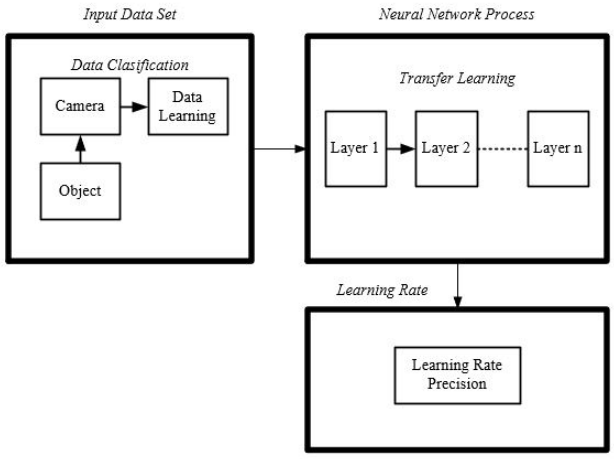


Fig. 1. The block diagram of design system.

In the input section, the object is a stationary robot in different conditions that will be photographed using a camera, then the condition of the robot is divided into three conditions, namely straight, right oblique, and left oblique conditions which will later become data learning. These data learning will become a part of the input data set which then the data will be stored on Google Drive.

After data input process, the next process is using a Neural Network with the Transfer Learning method which consists of several layers. This programmed is using the Python programming language in the Google Collabs application. Programming is done through the Google Collabs application by retrieving learning data that has been stored on Google Drive. Furthermore, the data will be trained by the program with the aim that the system can read the condition of the robot. The result of the programming is in the form of a percentage of "Learning Rate Precision" which will later become a reference whether the system can read and classify the condition of the robot or not.

B. Artificial Neural Network

The basic concept of the development of artificial intelligence Neural Network is to adopt the thinking mechanism of a system or application that resembles the human brain, both for processing various received element signals, tolerance for errors, and also parallel processing.

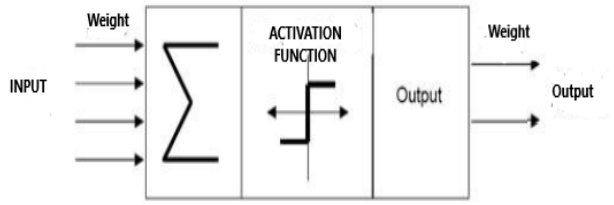


Fig. 2. ANN's structure.

In Fig. 2 the neural network process at each layer begins with an input which will become a weight, which then enters the layers contained in the ANN and passes the activation function which is used to get the output of the neuron from the input value, which then becomes the output. to be forwarded to the next laye

C. Data Set, Data Train, and Data Test Design

The retrieval of Dataset and Data Learning is done by taking pictures using a camera with a mobile robot object with 3 different classifications of conditions, namely RSTRAIGHT, RTURNRIGHT, and RTURNLEFT as shown as illustration in Fig. 3, Fig. 4, and Fig. 5.

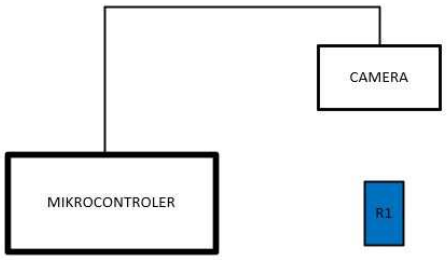


Fig. 3. Dataset RSTRAIGHT

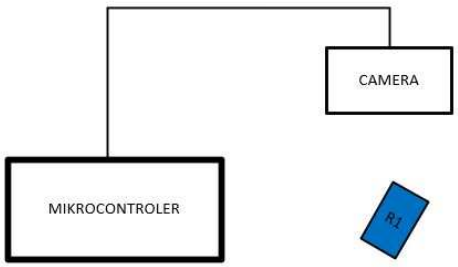


Fig. 4. Dataset RTURNRIGHT

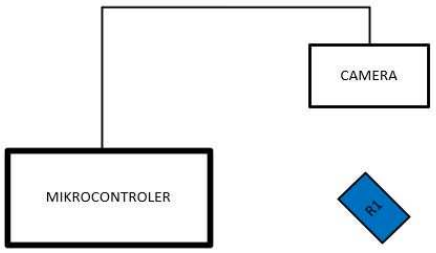


Fig. 5. Dataset RTURNLEFT

The retrieval of the Dataset is done by taking the image object 100 times in each classification. The total inputted dataset is as many as 300 images which then the data will become a Dataset. Furthermore, it will be determined for Train Data and Test Data with a distribution of 75% for Train Data and 25% for Test Data.

D. Simulation

After doing the design, a simulation will be carried out with the aim of seeing how much "learning rate" is generated by the program to be able to read the condition of the robot. For the input itself, 300 datasets are used which are divided into 3 position classifications, namely RSTRAIGHT, RTURNRIGHT, and RTURNLEFT as we seen in Fig. 6. Where each there are 100 images that will later be used as learning data.

(a) (b) (c)

Fig. 6. Dataset Input

Then all the datasets are stored in Google Drive and made into one folder in which there are 3 position classification folders. The Google Drive will be called by the program through the Google Collab application by Mount File.
After that, all the images that become the dataset will be resized by the program to a size of 224 x 224 pixels. The amount of Train data and Test data is 75% for Train data and 25% for Test data. So that we get 225 data for the Train Data and 75 data for the Test Data.

```
(225, 224, 224, 3)
(75, 224, 224, 3)
```

Fig. 7. Percentage of Data Train and Data Test

In this program, we use patience with a value of 10, and this program uses an optimizer, namely "Adam Optimizer". Adam is an optimization algorithm that can be used instead of the classic stochastic gradient descent procedure to update the weighted network iteratively based on training data [10].

Other parameters that exist in the program are time which is given a value of 50. And batch size is given a value of 10. Epoch itself is the iteration limit for learning data which means 50 times the maximum. The batch size is the amount of data drained in one iteration. For optimal results. On the program that applies the es callback. This aims to measure when he does learning but the data cannot go up again so he will stop without having to wait until the total of the epoch is finished. In this study, based on experiments that have been carried out after conducting research and executing the program, the results of the program are obtained in the form of a learning rate precision. The input given for this research is as many as After all the parameters have been inputted, program execution is carried out on Google Collabs to see the level of accuracy and the resulting learning rate

III. RESULT AND ANALYSIS

In this study, based on experiments that have been carried out after conducting research and executing the program, the results of the program are obtained in the form of a learning rate precision. The input given for this research is as many as 300 Learning Data which then appears by the program. Some of the parameters inputted in this program include a lot of patience with a value of 10, an epoch value of 50, and a batch size of 10.

With several parameters that have been entered earlier, an execution program is carried out which will produce an output in the form of a learning rate. The output produced in this study amounted to 98% with a total of 12 iterations. And produce two graphs that depict the comparison of time and the relationship of loss with epoch.

```
Epoch 00006: val_accuracy did not improve from 0.98667
Epoch 7/50
23/23 [==============================] - 2s 70ms/step - loss: 0.1286 -

Epoch 00007: val_accuracy did not improve from 0.98667
Epoch 8/50
23/23 [==============================] - 2s 71ms/step - loss: 0.1640 -

Epoch 00008: val_accuracy did not improve from 0.98667
Epoch 9/50
23/23 [==============================] - 2s 71ms/step - loss: 0.2827 -

Epoch 00009: val_accuracy did not improve from 0.98667
Epoch 10/50
23/23 [==============================] - 2s 71ms/step - loss: 0.0330 -

Epoch 00010: val_accuracy did not improve from 0.98667
Epoch 11/50
23/23 [==============================] - 2s 71ms/step - loss: 0.0698 -

Epoch 00011: val_accuracy did not improve from 0.98667
Epoch 12/50
23/23 [==============================] - 2s 70ms/step - loss: 0.0276 -

Epoch 00012: val_accuracy did not improve from 0.98667
```

Fig. 8. Learning rate precision.

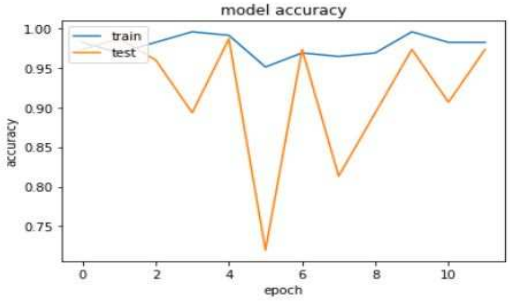


Fig. 9. Graphics of accuracy model

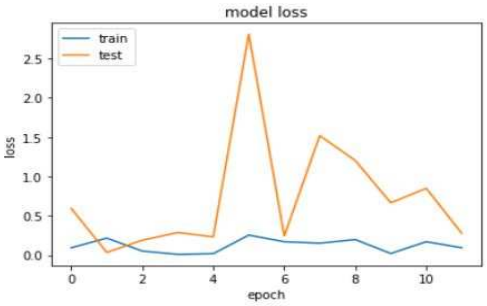
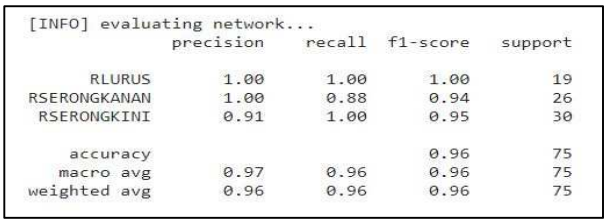


Fig. 10. Graphics of loss model

For the results of the graph itself in Fig. 9, a graph is obtained from the test data with the smallest value on the fifth epoch with a value of 0.72 and the highest value is obtained on the fourth epoch with a value of 0.98. Then for the training data itself, the lowest value is obtained in the fifth epoch, but seen from the results it is still good because the value obtained is still above 0.90 and the highest value is obtained in the third and ninth epoch which is almost close to 1. From the graph, it can be seen that there are still increase and decrease in the data chart. This means that the data is not stable.

For the next graph depicting the loss model, it can be seen that the resulting loss model is the largest in the fifth epoch, which is above 2.5. For the two-graph model that has been described earlier, it is actually related where when the loss model is higher, the resulting accuracy will be lower. In this graph, it can be concluded that when doing training there is some accuracy and loss that fluctuates in both test and train data. This happens because of the dataset factor and the activation function used in the program.



```
[INFO] evaluating network...
              precision    recall  f1-score   support

       RLURUS       1.00      1.00      1.00        19
RSERONGKANAN        1.00      0.88      0.94        26
 RSERONGKINI        0.91      1.00      0.95        30

    accuracy                           0.96        75
   macro avg       0.97      0.96      0.96        75
weighted avg       0.96      0.96      0.96        75
```

Fig. 11. Network evaluation

For the detailed data itself in Fig. 11. It can be seen that for RSTRAIGHT data the best result is 1.00, and for the RTURNRIGHT results the final result is 0.94, for RTURNLEFT the final result is 0.95. Based on the above data, the system is easier to recognize RSTRAIGHT compared to RTURNRIGHT and RTURNLEFT.

IV. Conclusion

In this study, the inputs given were 3 data sets with 100 images each. For the condition classification of RSTRAIGHT, RTURNRIGHT, and RTURNLEFT robots with the number of rail data entered is 75% and test data is 25%. The results obtained in the form of "learning rate precision" of 98%. With epoch values of 50 and 12 iterations. The entered batch size is 10 and patience is 10.

for future implementation, this system can be used for Autonomous Mobile Robots that can move automatically at predetermined coordinate points with the help of a camera as a sensor.

References

[1] M. A, Neural Network Learning : Theoritical Foundation. Cambride University Press, New York, 1999.

[2] T. Guo, J. Dong, H. Li, and Y. Gao, “Simple convolutional neural network on image classification,” in 2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA), 2017, pp. 721–724.

[3] M. T. Islam, B. M. N. Karim Siddique, S. Rahman, and T. Jabid, “Food image classification with convolutional neural network,” in 2018 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), vol. 3, 2018, pp. 257–262.

[4] L. Shao, F. Zhu, and X. Li, “Transfer learning for visual categorization: A survey,” IEEE Transactions on Neural Networks and Learning Systems, vol. 26, no. 5, pp. 1019–1034, 2015.

[5] S. Annaduray, Fundamentals Of Digital Image Processing. Dorling Kindersley, India, 2007.

[6] Zhu, Y et al., “Heterogeneous transfer learning for image classification,” Proceedings of the AAAI Conference on Artificial Intelligence, 2011.

[7] Le Hoang Tai et al, “Image classification using support vectormachine and artificial neural network,” Published Online May 2012 in MECS, 2012.

[8] I. P. Ivan Markovic, Franois Chaumette, “Moving object detection, tracking and followingusing an omnidirectional camera on a mobile robot,” IEEE Int. Conf. on Robotics and Automation,ICRA14, 2014.

[9] I. Baek, A. Davies, G. Yan, and R. R. Rajkumar, “Real-time detection, tracking, and classification of moving and stationary objects using multiple fisheye images,” in 2018 IEEE Intelligent Vehicles Symposium (IV), 2018, pp. 447–452.

[10] P.K Diederik, L.B Jimmy, "ADAM : A Method For Stochastic Optimazion" Published as a Conference Paper at ICLR 2015