Recognition of Pedestrian Traffic Light using Tensorflow And SSD MobileNet V2

To cite this article: Evan *et al* 2020 *IOP Conf. Ser.: Mater. Sci. Eng.* **1007** 012022

Recognition of Pedestrian Traffic Light using Tensorflow And SSD MobileNet V2

Evan[1], Meirista Wulandari[2], Eko Syamsudin[3]

[1]Electrical Engineering Department, Faculty of Engineering
Universitas Tarumanagara, Jakarta.

[2]Electrical Engineering Department, Faculty of Engineering
Universitas Tarumanagara, Jakarta.

[3]Electrical Engineering Department, Faculty of Engineering
Universitas Tarumanagara, Jakarta.

E-mail:evan.525160004@stu.untar.ac.id

Abstract. As a government needs to facilitate the whole public including traffic public facilitate. A pedestrian traffic light should be perceived by pedestrian to safe them to cross the street including someone with the vision deficiency problem. Nowadays, a technology supports human being to build some innovations or application more practical. An innovation which adopt the computer vision and deep learning technology can be built to recognize the pedestrian traffic light which is portable and practical in the context of light computing. The paper aim to evaluate threshold value and data parameters to recognition pedestrian traffic light by Tensorflow and SSD MobileNet V2. The sample are green light and red light on pedestrian traffic light. Based on result, more number of data train than data test and higher of threshold value, the accuracy obtained is more better. Object detection accuracy level reaches 97,98%. On the other result, earphone can produce sounds based on detected objects with output accuracy level reaches 100%.

1. Introduction

Based on the Data and Information Center of the Indonesian Ministry of Health, there are approximately 5.8 million people in Indonesia who had problem with vision. About 5.313 million Indonesian people had vision deficiency at low-level and about 507 thousand Indonesian people had vision deficiency at high-level. Some people have vision deficiency since they are born and some people have vision deficiency because of accidents [1].

The problem of vision deficiency might cause a decrease productivity. Because person needs other people, technology and facilities around to assist and to support him. One of many facilities that are created by government to assist someone who had vision deficiency is pedestrian traffic light. It makes pedestrian feel safe to cross the street. The pedestrian traffic lights show an information of street's condition by an indicator pedestrian traffic lights [2]. Red light shows that pedestrians are disable to cross the street and green light shows that pedestrians are able to cross the street. Because of vision deficiency the vision of pedestrian light can be replaced by sound of red and green indicators.

In this era of revolution Industry 4.0 with the rapid of technology development, there is an innovation instrumental area based on computer vision and deep learning. The innovation can support the people who had vision deficiency to obtain the information of pedestrian traffic light. The information is captured by camera and the image is processed by a microcomputer with deep learning algorithm inside of it to recognize the pattern. The innovation uses a Raspberry Pi 4 model B as the microcomputer and the Tensorflow and SSD MobileNet as the deep learning algorithm which is light of computing [3].
The pedestrian traffic light recognition used RGB images as the input. Some algorithms have been applied to recognize the traffic light based on color [4], edge [5], shape and texture [6]. The algorithms which is based on color, it computed the color density information to recognized the color of pedestrian traffic light [7]. However, the recognition of pedestrian traffic light uses many parameters to obtain the optimum result.
The optimum result of pedestrian traffic light recognition can be shown by accuracy value. Factor on level accuracy of recognition depends on the configuration of the threshold values and comparison in the number of train data and test data. In this paper, we discuss the method, result and discussion and conclusion.

2. Theory

2.1. Pedestrian Traffic Light

Pedestrian traffic lights are facilities for pedestrians to assists pedestrians cross the street safely and comfortably. Pedestrian traffic lights have 2 signal such as green light and red light. Generally, the width of vehicle lanes is 3.5 meters, so the lane has free space of 0.5 meters if there are vehicles with a maximum width of 2.5 meters. The range time when the green signal on the pedestrian traffic light is between 7 to 60 seconds depends on the volume of vehicles and pedestrians at the street [8].

Figure 1. Pedestrian Traffic Light

2.2. Image Processing

Image processing is a method to manipulating and analyzing images. The input of image processing is the image and the output is a processed image. This image processing uses certain information or the number of pixels on an image. The purpose of image processing is to improve image quality (Image Enhancement) by adding certain features and restoring images (Image Restoration) by removing noise that causes image quality degradation [9].

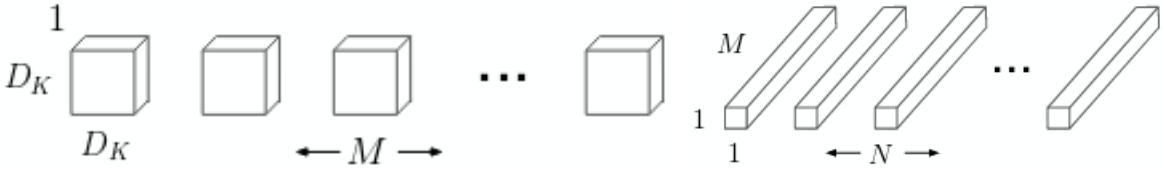
2.3. Tensorflow

Tensorflow is an open source library that used to build, develop and conduct training data on machine learning models. Machine learning is a computer science to help and solve problems automatically. Tensorflow requires data to be trained [10].
Tensorflow uses a variety of tools and libraries so Tensorflow can be widely implemented. Tensorflow can be used for number processing, image processing, sound processing, etc. Tensorflow can be converted to Tensorflow Lite for mobile devices user [11].

2.4. SSD Mobilenet V2

SSD MobileNet V2 is one of the machine learning architectural models. SSD MobileNet V2 is designed to detect more than one or multiple objects in real time. SSD MobileNet V2 is used for mobile devices with light specifications. The architecture of SSD MobileNet V2 has the features of linear bottlenecks and shortcut connections [12].

SSD MobileNet V2 is one of the models of the Convolutional Neural Network (CNN) architecture. Convolutional Neural Network is a computer science for the development of artificial neural networks by adopting human neural networks to recognizing and detecting objects. SSD MobileNet V2 uses depthwise convolution and pointwise convolution [13].



(a) (b)

Figure 2. Architecture of CNN Model (a) Depthwise Convolution and (b) Pointwise Convolution [12]

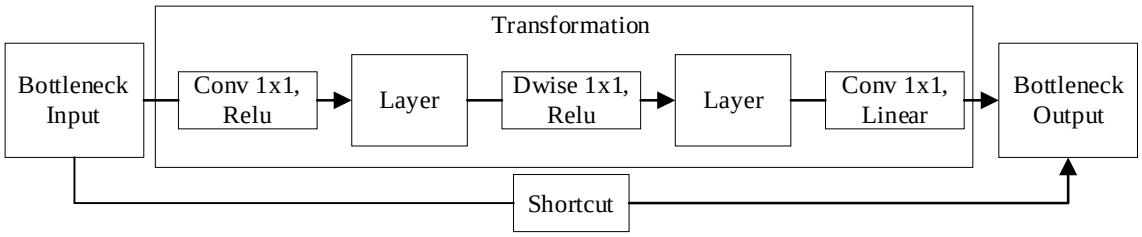


Figure 3. Architecture of SSD MobileNet V2

2.5. Raspberry Pi 4 Model B

Raspberry Pi 4 Model B has similar components to Personal Computer (PC). Raspberry Pi 4 Model B has components such as CPU, GPU, RAM, USB Port, HDMI, GPIO and Audio Port. Raspberry Pi 4 Model B requires power 5V/3A to operate. Raspberry Pi 4 Model B uses Micro SD Card as a data storage [14].

Raspberry Pi 4 Model B can be connected and control other devices or modules such as sensor modules, camera modules, earphone etc. Raspberry Pi 4 Model B needs to be programmed to control other modules or devices.

Figure 4. Raspberry Pi 4 Model B

3. Results and Discussion

In this paper, the data of pedestrian traffic light recognition are obtained at crossing street in front of the Bundaran Hotel Indonesia, Jakarta. The width of the street up to 12.5 meters. Time estimated of the signal green on pedestrian traffic light up to 20 seconds depends on the volume of vehicles and pedestrians on the road. The data consist of 530 images of red light and green light pedestrian traffic light. The training of deep learning is based on video that captured during the data obtaining. The data video divided into 99 frames. The images are captured and labeled as the green light or red light. The labeled activity in the purpose of training the SSDMobileNet. The result of recognition by SSDMobileNet depends on the confidence value which is called as threshold. The value of threshold prevents the misclassification of the result. In the range of 55% up to 85% threshold value is investigated in this paper. The value of comparison data training and data testing is investigated on this paper. The result with 70:30, 80:20 and 90:10 comparison training data and testing data respectively is shown in figure 5. Data results of pedestrian traffic lights detection can be seen in Table 1.

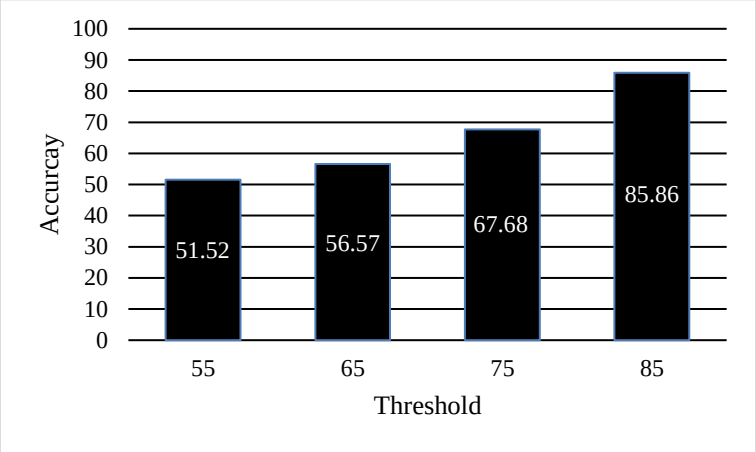


Figure 5. Accuracy with 70% : 30% of Training Data and Testing Data

From the data result in Figure 5, The highest accuracy of object detection is 85.86% with threshold value at 85%. The lowest accuracy of object detection at 51.52% with threshold value at 55%.

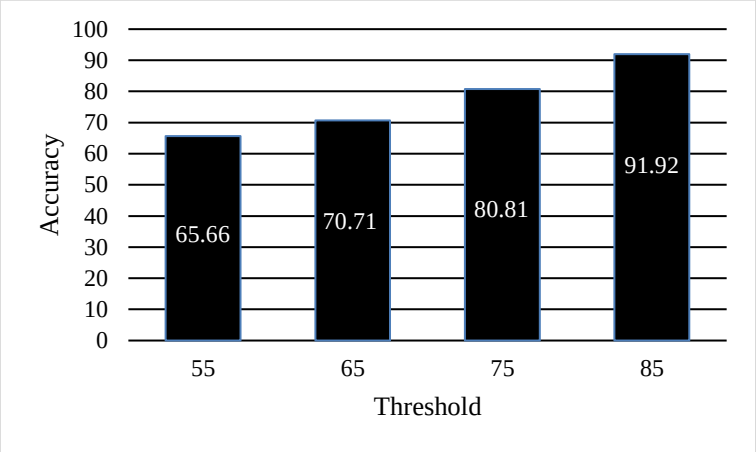


Figure 6. Accuracy with 80% : 20% of Training Data and Testing Data

From the data result in Figure 6, The highest accuracy of object detection is 91.92% with threshold value at 85%. The lowest accuracy of object detection at 65.66% with threshold value at 55%.

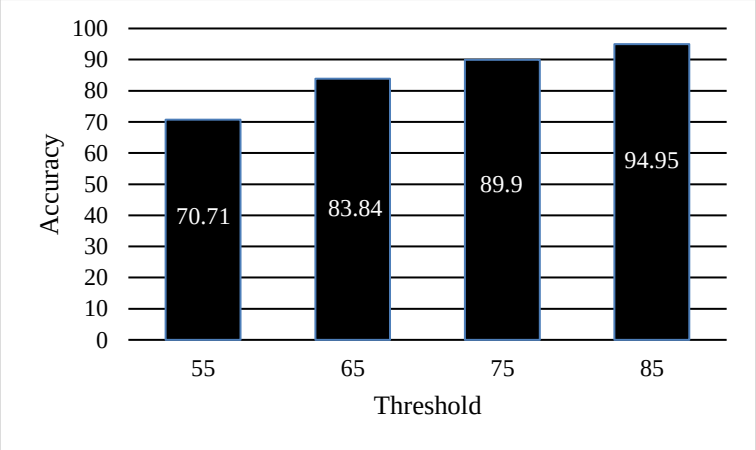


Figure 7. Accuracy with 90% : 10% of Training Data and Testing Data

From the data result in Figure 7, The highest accuracy of object detection is 94.95% with threshold value at 85%. The lowest accuracy of object detection at 70.71% with threshold value at 55%.

From the Figure 5, Figure 6 and Figure 7 can be seen that the number of comparison training data and testing data gave a variation of accuracy. However, those figures showed the trend of increasing accuracy proportionally comparable with the value of threshold.

Table 1. Result of Accuracy of Output Sound of Red Light and Green Light

Data Train : Data Test (%)	Threshold (%)	Number of False Detection	Object Detection Accuracy Level (%)	Output Sound Accuracy Level (%)
70 : 30	85	14	85,86	100
80 : 20		8	91,92	100
90 : 10		5	94,95	100

Table 1 shows the result of accuracy of output sound of red light and green light. The output sounds based on the pedestrian traffic light detected. All output sound accuracy levels are 100% due to the compatibility between the object detection results and the sound produced.
Some results of detection of pedestrian traffic lights can be seen in Figure 8 and Figure 9. The object detection result matched with the pedestrian traffic light in Figure 8. The figure 8 shows that the green light is successful to be recognized as green light at above and the red light is successful to be recognized as red light at the bottom. On the other hand, the object recognition result which is mismatched can be seen in Figure 9. Figure 9 shows that some mismatched would be happen when the system false detected object and the system could be wrong recognized the color of the pedestrian traffic light.

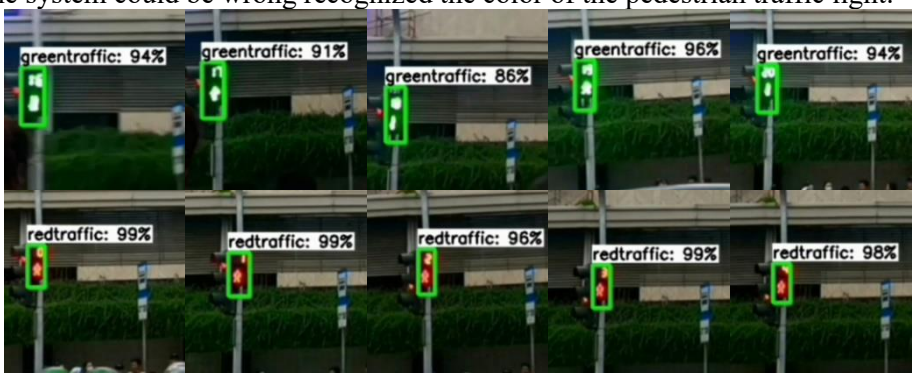


Figure 8. The Object Detection Result Does Matches With The Pedestrian Traffic Light

Figure 9. The Object Detection Result Which is Mismatch With The Pedestrian Traffic Light

4. Conclusion

From the data result of this paper, The highest accuracy of object detection is 94.95% has a comparison on training data and testing data at 90% : 10% of all data and the threshold value at 85%. The lowest accuracy of object detection at 51.52% has a comparison on training data and testing data 70% : 30% of all data and the threshold value at 55%. It can be concluded that more number of training data than testing data and higher of threshold value, give a better result in accuracy. On the other hand, output sounds can be produced successfully 100% based on the object detected. Some computer vision and deep learning application can be applied to obtain the real time system that is portable and very helpful for someone with a vision deficiency. This innovation can be more developed.

5. References

[1] Pusdatin Kemkes. Situasi Penyandang Disabilitas 2014, Jakarta, 2014
[2] E. Soleimani-sefat, M. Rostami, S. Amani, and G. Movallali, "The Needs and Problems of Students with Visual Impairment," vol. 2, no. 2, pp. 8–16, 2016.
[3] S. Phon-Amnuaisuk, K. T. Murata, P. Pavarangkoon, K. Yamamoto, and T. Mizuhara, "Exploring the Applications of Faster R-CNN and Single-Shot Multi-box Detection in a Smart Nursery Domain," 2018.
[4] G. Siogkas, E. Skodras, and E. Dermatas, "Traffic Lights Detection in Adverse Conditions using Color, Symmetry and Spatiotemporal Information.," in *VISAPP (1)*, 2012, pp. 620–627.
[5] M. Omachi and S. Omachi, "Traffic light detection with color and edge information," in *2009 2nd IEEE International Conference on Computer Science and Information Technology*, 2009, pp. 284–287.
[6] C. C. Chiang, M. C. Ho, H. S. Liao, A. Pratama, and W. C. Syu, "Detecting and recognizing traffic lights by genetic approximate ellipse detection and spatial texture layouts," *Int. J. Innov. Comput. Inf. Control*, vol. 7, no. 12, pp. 6919–6934, 2011.
[7] T. H. P. Tran, C. C. Pham, T. P. Nguyen, T. T. Duong, and J. W. Jeon, "Real-time traffic light detection using color density," in *2016 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia)*, 2016, pp. 1–4.
[8] F. Lestari, "Analis Yuridis Lampu Merah Penyeberangan (Pelican Crossing) dalam Perspektif Hukum Pengangkutan," vol. 2, no. 4, pp. 1235–1256, 2019.
[9] R. C. Gonzalez, R. E. Woods, A. H. Scott, and P. P. Hall, *Digital Image Processing*. .
[10] S. Das and N. Roy, "Applications of Artificial Intelligence in Machine Learning : Review and Applications of Artificial Intelligence in Machine Learning : Review and Prospect," no. April, 2015.
[11] W. G. Hatcher and W. E. I. Yu, "A Survey of Deep Learning : Platforms , Applications and Emerging Research Trends," *IEEE Access*, vol. 6, pp. 24411–24432, 2018.
[12] W. Liu *et al.*, "SSD : Single Shot MultiBox Detector."
[13] M. Xin and Y. Wang, "Research on image classification model based on deep convolution neural network," vol. 8, 2019.
[14] A. Nayyar, "Raspberry Pi- A Small, Powerful, Cost Effective and Efficient Form Factor Computer: A Review," no. July, 2016.