Evolution in Electrical and Electronic Engineering Vol. 2 No. 2 (2021) 629-636

EEEE

Homepage: http://publisher.uthm.edu.my/periodicals/index.php/eeee
e-ISSN : 2756-8458

Development of a Malaysian Sign Language Interpreter by using Image Recognition Techniques for the Community to Understand the Deaf

Nurul Ashikin Mohammad[1], Tee Kian Sek[1]

[1]Faculty of Electrical and Electronic Engineering,
Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Johor, MALAYSIA

*Corresponding Author Designation

DOI: https://doi.org/10.30880/eeee.2021.02.02.075
Received 04 July 2021; Accepted 27 September 2021; Available online 30 October 2021

Abstract: In Malaysia, people with hearing problems still suffer going through their daily life because they struggle to have a conversation with the community who does not know or how to do sign language efficiently. This project aims to develop a Malaysian Sign Language interpreter to convert hand sign language to texts, in order to facilitate the conversation between the deaf and the community. Basically, when a hand sign is detected by the camera device, a text will be shown to its corresponding meaning. The requirement for this project is a laptop or desktop that has camera access for image recognition and internet access to connect to Jupyter Notebook to run the specific coding. The system would implement a camera and vision system to capture images of hand signs. These images would be further been annotated correlated to their meaning. It will undergo a training process digitally and eventually being digitally recognized using a series of image processing algorithms and recognition algorithms. Lastly, the system will detect hand signs and interpret the meaning of the hand sign by using the camera. Hence, the community can understand them well and able to communicate. At the end of this project, the system able to convert hand sign into visible text and manage to help the deaf to connect to community in their daily life.

Keywords: Sign Language Interpreter, Image Recognition Techniques, Hand Sign

1. Introduction

Communication is a two-way process on how peoples exchange messages between each other in order to have the same understanding or subject. There are two types of communication; verbal communication and non-verbal communication. Verbal communication is about language; it can be either spoken or written. Meanwhile, nonverbal communication refers to communication that occurs through means over words [1]-[2]. For example, sign language, body language and facial expression, in simple words it is a wordless signal. For deaf community, they use a non-verbal language to deliver

Corresponding author: tee@uthm.edu.my

messages to other parties. In 2008, Malaysia Sign Language (MSL) was acknowledged to be the main sign language for deaf community in Malaysia [3]. Even though MSL is used widely by the deaf in Malaysia, individuals with hearing problems have difficulty with communication in their daily life, especially with non-sign language users. This is because most verbal language users are not exposed to sign language as they can speak well to interact with others. Thus, this creates a communication barrier toward community and deaf.

1.1 Deaf Community in Malaysia

The total population in Malaysia during 2019 was 31,949,777 people and the total of registered disability people were 548,195 people, which only 1.72% out of Malaysian. Based on the Malaysia statistic report 2019 of Department of Social Welfare, Ministry of Women, Family and Community Development, Malaysia [4] 38,676 out of 548,195 disable people have hearing problem [4].

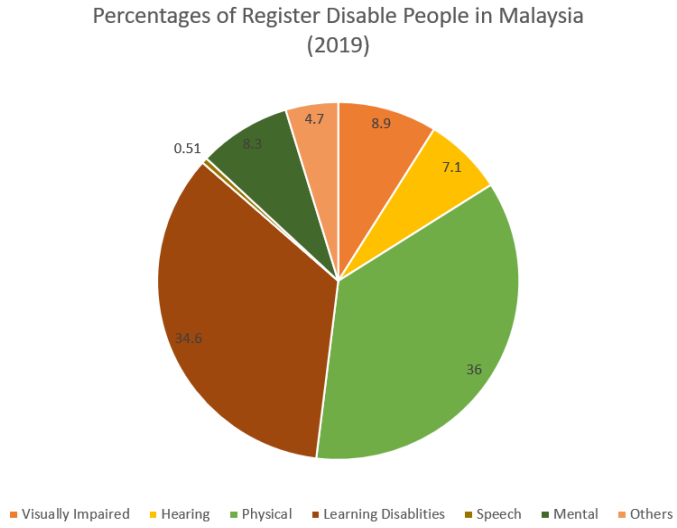


Figure 1: The Percentages of Registered Disable People in Malaysia (2019) [4]

From Figure 1, the total percentages of disable people who have hearing problems is 7.1%, which is the top 5 in disabilities in Malaysia. However, the World Health Organization (WHO) has declared that each country must have at least 15% disable people out of the total population, which means there are 4,792,466 estimated disable people that still have not registered yet in Malaysia. Therefore, the government needs to take serious action to discover unfortunate people in Malaysia [5].

1.2 Common Problem Faced by Deaf

The deaf face obstacles in daily life. This is because the community is lacking exposure to sign language and some of the community still have prejudice toward the deaf because they are considered disable people [6]. This effect of communication breaks down between the community and the deaf and causes the deaf to have less chances to survive in the society for example, in the workplace or studies [7].

In Malaysia, people with hearing problems still face hardship in their life because they have trouble with communication especially with the people who do not understand sign language. This is because the community themselves lack the awareness towards the disable people and have prejudice towards the deaf. Despite all ages and society, people with hearing problems still face obstacles in daily life. This is because the community is lacking exposure to sign language and some of the community still have prejudice toward the deaf because they are considered disable people [8]. This effect of communication breaks down between the community and the deaf and causes the deaf to have less chances to survive in the society for example, in the workplace or studies [9].

2. Materials and Methods

Figure 2 indicate the flowchart of the architecture framework for Malaysian Sign Language interpreter. The first phase is creation of database. The images can be taken by using any type of camera deceives, for example smartphones , digital camera or webcam. The Image properties for all data collection need to have the same spatial resolution such as 640 x 480 pixels, 320 x 240 pixels or 1280 x 720 pixels based on preferences. This is crucial especially during the training section of the image collection. In this project will be using 640x 640 pixels as the resolution.

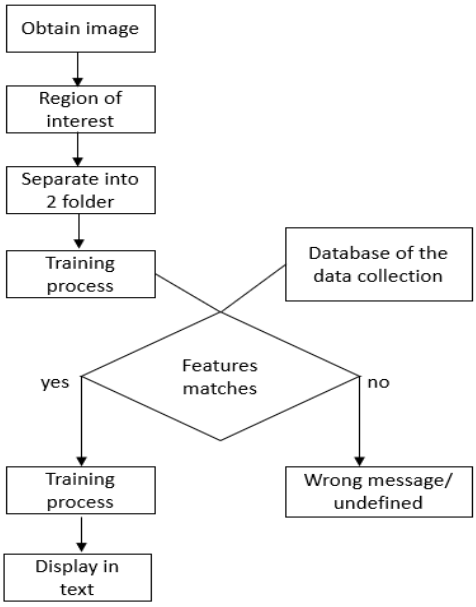


Figure 2: Flowchart of the proposed framework

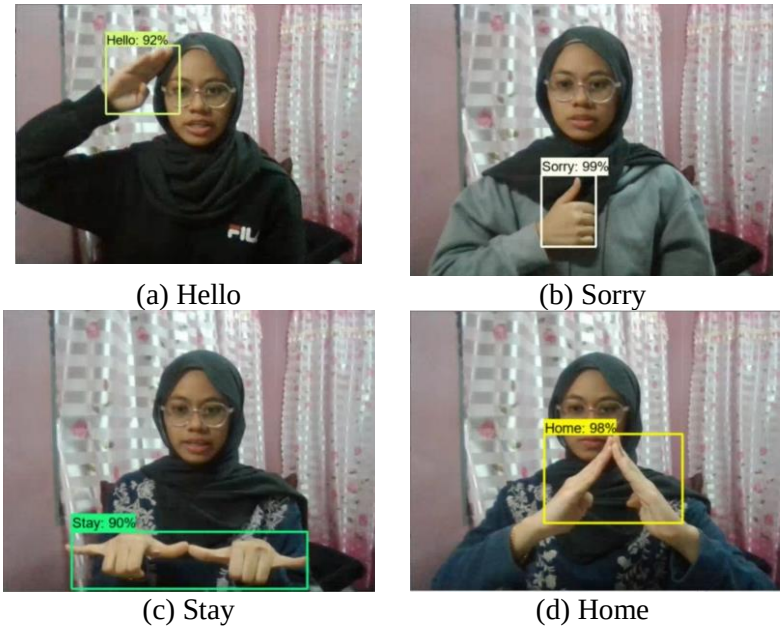
Next step is feature extraction, needed to highlight or focus on the region of interest, which can be done by using a graphical annotation tools (GAT). The tool was created with usability in mind, with the goal of reducing user input by predicting future region and semantic class selections. In this project will be using a special Graphical Annotation Tools, LabelImg as the medium. After the all images have been annotated, the data collection needs to be split into few different compartments. There are two type of dataset which are Training Dataset and Test Dataset .

There are many kinds of deep learning method for the training process that can be choose. For example, Classic Neural Networks, Convolution Neural Networks and Deep Reinforcement Learning, . Since each method has different built, the type of method needs to choose wisely by using the benefit of the method and requirement for the Classifier Training and Validation. Therefore, this project will be using SSD MobileNet V2 FPNLite 320x320 with properties 22ms, 22.2 COCO mAP and will have and boxes output.

Real-time image recognition can be done by using a camera device right after the training process finished. This process will be using the result of the training and will try to recognize if there is a Malaysian Sign Language that have been performed in front of the camera. There are two possible result which are if image recognition able to recognize the Malaysian Sign Language, the result will show a correct sign language translation. However, if the image recognition failed to detect, it will show either wrong sign language translation or will not show at all.

2.1 Experiment Setup

A simple webcam from laptop is used as a medium to capture images during the collections of images. Pre-defined images of resolution 640 x 480 were saved. Four MSL messages (also known as the classes for SSD) as illustrated in Figure 3, are experimented, i.e. “Hello”, “Sorry”, “Stay”, “Home”. We have deployed open-source mobile inference box, SSD MobileNet v2 320x320, shared by TensorFlow 2 Detection Model Zoo, for the purpose of hand signal recognition. The setup is illustrated in Figure 4.



(a) Hello (b) Sorry

(c) Stay (d) Home

Figure 3: The classes of hand signs

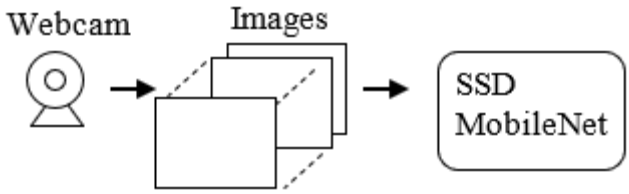


Figure 4: Experiment setup

2.2 Preparation for model training

For the purposes of train and test, five images were taken for each class where four for train, one for test. A total of 4 classes x 5 images = 20 images, were taken. These images were labelled and annotated. For each image, its class was labelled and a bounding box was drawn manually at the location of the hand signal. Image annotation was conducted using LabelImg [9]. A metadata of each image annotation was generated and was saved as a XML file. The images and their image annotations were sorted and saved in known folders labelled as “train” and “test”. These data were fed into SSD for train and test.

3. Results and Discussion

In this part, it will be discussing about the result from the training model based on the four selected Malaysian Sign Language, which are ‘hello’, ‘sorry’, ‘stay’ and ‘home’.

3.1 Training results

SSD training was conducted on HP Laptop 15s, running on AMD Ryzen 5 3500U with Radeon Vega Mobile Gfx 2.10GHz, RAM 4GB. By default, TensorFlow V2 turns on GPU mode. Total training time took 1 hour and 45 minutes.

3.2 Classification Loss

The loss method is used to measure the difference between the model's output data and the actual sample data, and its purpose is to guide the model towards convergence during the training process, where minimizing the loss value is essentially to achieve model fitting of training data and the model's minimum test error, and eventually to accurately classify new samples [10]. Based on the previous figure, the precision for this model is 0.2618 which is 26.18 % in percentages at 2000 steps.

In conclusion, the score of the correct category should be bigger by some margin than the total of the scores of the erroneous categories (usually one). As a result, for maximum-margin classification, hinge loss is utilised, most notably for support vector machines. Despite the fact that it is not differentiable, it is a convex function, making it simple to work with common convex optimizers in the machine learning area. In conclusion, classification loss is loss calculated based on accurately classifying the object. Hence, based on Figure 5, at 1997 steps the loss was 0.0552 or 5.52% in percentages

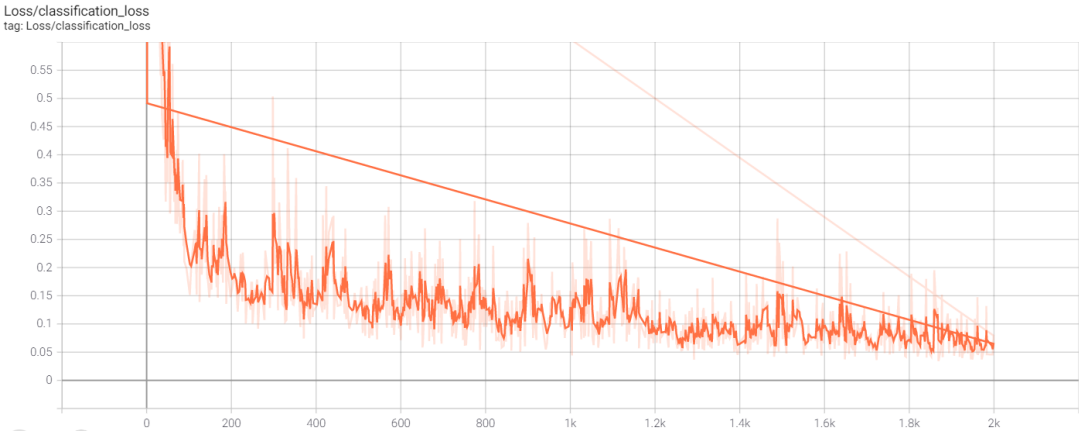


Figure 5: Classification Loss

3.3 Localization loss

The localization loss is a smooth L1 loss between the projected bounding box adjustment and the reality values. The coordinate correction transformation is the same as R-bounding CNN's box regression transformation. Basically, localization loss is loss attached to unnormalized position so a measure of how far or close the detect object was from the real object placement. Localization Loss graph is shown in Figure 6.

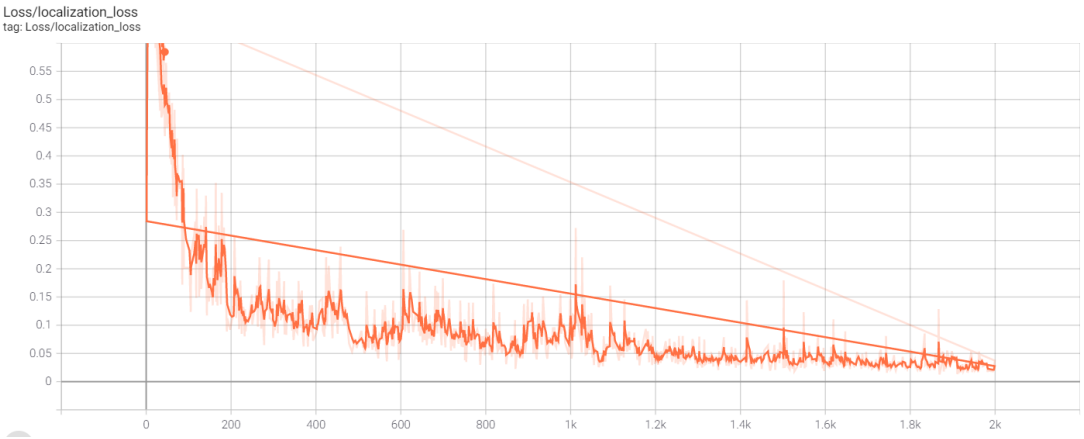


Figure 6: Localization Loss

3.4 Expected Results

Based on Figure 7, the model able to detect most of the data collection from the test category by using the webcam as a medium to take picture. The image taken from the data collection contain four Malaysian Sign Language and different position using different person.

Figure 7: The classes of hand signs using image collection

Based from both Table 1 and Table 2, the model gains various result from different sign language and different T-shirt. From Table 2, the accuracy of the result from black T-shirt is better than the other two T-shirt because the plain black T-shirt reflected less light from the lamp. The grey T-shirt and

pattern T-shirt have a decent result as the grey T-shirt reflect more light and pattern T-shirt have decoration that may interrupt the model.

Table 1: The classes of hand signs using Laptop Webcam

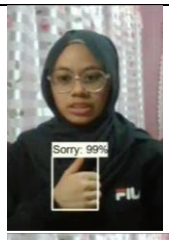
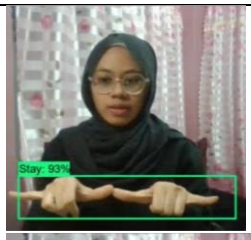
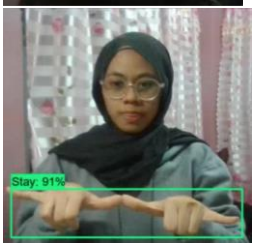
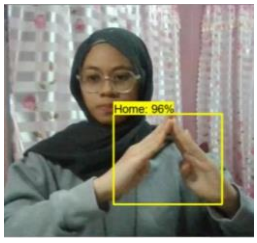
	a) Hello	b) Sorry	c) Stay	d) Home
1				
2				
3				

Table 2: Maximum and minimum accuracy between difference outfit

	hello		sorry		stay		home	
	min	max	min	max	min	max	min	max
Black T-shirt	80%	98%	93%	100%	82%	95%	85%	99%
Grey T-shirt	86%	96%	83%	99%	81%	96%	82%	97%
Pattern T-shirt	80%	97%	81%	97%	84%	95%	85%	95%

4. Conclusion

For future works, this project can be improve by adding more Malaysian Sign Language and implement it into a system. The more the Malaysian Sign Language that has being added, the better the community will able to understand the deaf. Thus, it can be help deaf people in daily routine especially when they need to ask or talk to person who have zero knowledge in Malaysian Sign Language. Next, the system can be added more features for the translation by adding a voice to indicate the sign language instead of a visible text. Lastly, this project has the capability to be implemented into embedded system by uploaded the training data into the integrated circuit. Hence, the system can be worked easily everywhere with the camera device. In conclusion, the main purpose of PSM 2 to obtain the image as an input and converting it into image segmentation was a success. The project able to detect 4 type Malaysian Sign Language with average 86% accuracy.

Acknowledgement

The authors would like to thank the Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia for its support.

References

[1] "Difference Between Verbal and Non-Verbal Communication."

[2] S. M. Mohd Rashid, M. H. Mohd Yasin, and N. S. Ashaari, "Penggunaan Bahasa Isyarat Malaysia (BIM) di dalam terjemahan maksud Surah Al-Fatihah," GEMA Online J. Lang. Stud., vol. 17, no. 4, pp. 209–224, 2017, doi: 10.17576/gema-2017-1704-14

[3] S. Ghanem, C. Conly, and V. Athitsos, "A survey on sign language recognition using smartphones," ACM Int. Conf. Proceeding Ser., vol. Part F1285, no. June 2017, pp. 171–176, 2017, doi: 10.1145/3056540.3056549

[4] National Welfare Department, "Statistic Report 2018," Stat. Rep. 2018, vol. 1, no. 1, p. 1, 2018

[5] "Jumlah OKU berdaftar JKM masih rendah." https://www.sinarharian.com.my/article/23353/BERITA/Nasional/Jumlah-OKU-berdaftar JKM-

[6] A. F. B. M. Sahid, W. S. W. Ismail, and D. A. Ghani, "Malay Sign Language (MSL) for Beginner using android application," in ICICTM 2016 - Proceedings of the 1st International Conference on Information and Communication Technology, Mar. 2017, pp. 189–193, doi: 10.1109/ICICTM.2016.7890798

[7] S. Meeze, M. Rashid, M. Hanafi, and M. Yasin, "Di Dalam Terjemahan Maksud Surah Al-," no. May, pp. 935–942, 2013

[8] A. S. Amin, "Isu Pekerjaan Orang Kurang Upaya Pendengaran : Kajian Kes di Selangor," vol. 2, no. 2, pp. 1–10, 2018

[9] G. Atinc et al., "Pelaksanaan Dasar Inklusif Orang Kurang Upaya di Institusi Pendidikan Tinggi," J. Finance, vol. 11, no. 1, pp. 128–137, 2019

[10] W. Hao, R. Bie, J. Guo, X. Meng, and S. Wang, "Optimized CNN Based Image Recognition Through Target Region Selection," Optik (Stuttg)., vol. 156, pp. 772–777, 2018, doi: 10.1016/j.ijleo.2017.11.153