Advances in Social Science, Education and Humanities Research, volume 520
Proceedings of the 6th UPI International Conference on TVET 2020 (TVET 2020)

Construction of SIBI Datasets for Sign Language Recognition Using a Webcam

Dandhi Kuswardhana, Iva Rachmawati, Ghina Nadifah Amani, Hendra Fauzi, Agus Heri Setya Budi*
Electrical Engineering Education Department
Universitas Pendidikan Indonesia
Bandung, Indonesia
*agusheri@upi.edu

***Abstract*—SIBI (Sistem Isyarat Bahasa Indonesia) is a sign language established by the Government of Indonesia as a standard of communication that used in Special School. Hence, not everyone can understand the sign language, even though people have ability to hear. They also must be able to understand the sign language in order to conduct a communication. Currently, technology of recognition has been applied to hand gestures that can translate sign language. Various studies regarding sign language recognition have been developed and there are many applications that can recognize sign language through hand gesture as datasets using webcam, but there still restricted on the datasets in SIBI. Therefore, the study aims to create SIBI datasets via hand gesture using experimental method. So that, the application can be used by SIBI users. SIBI datasets were successfully created with signs that consisting of 8 static word. Based on the confusion matrix, all the signs that has been created have a precision of 100%, meanwhile based on the test results with sufficient lighting conditions and plain background, SIBI datasets has a prediction success rate of 97,5%.**

Keywords—SIBI datasets, sign language, recognition

I. INTRODUCTION

Naturally, human uses sign or cues when communicating with others. However, for deaf and speech impaired person, making a sign is the most natural language used to communicate [1]. There are several sign language systems in the world such as ASL (American Sign Language), BSL (British Sign Language), FSL (French Sign Language), CSL or ZGS (Chinese Sign Language), Brazilian Sign Language (Libras), and others. In Indonesia alone, there are BISINDO (Bahasa Isyarat Indonesia) and SIBI (Sistem Isyarat Bahasa Indonesia).

SIBI is a sign language that has been standardized and normalized corresponding with grammar, syntax, and word morphology, so that almost all basic words have their own signs, and to add vocabulary, SIBI have also been equipped with signs that represent affixes, and also has been set by the Government of Indonesia as a standard communication used in Special School [2]. On the other hand, not everyone can understand sign language, even though people has ability to hear, they also should be able to understand sign language in order to conduct a communication. Learning SIBI is one of good steps to develop communication skills. Thus, people can do various activities related to sign language.

Currently, technology of recognition has been applied to hand gestures that can translate sign language and make a better communication between the sign language user and non-sign language user. Various studies regarding sign language recognition have been developed and there are many applications that can recognize sign language through hand gesture as datasets using device that looks like a glove or camera, [3] successfully created glove like device that can translate ASL, but the article entitled 'Why Sign-Language Gloves Don't Help Deaf People' states that instead of simplify sign language user, this kind of recognition technology only facilitates non-sign language user.

Nowadays, camera based sign language recognition is a state of the art in recognition technology. Various kind of camera used in many studies, such as [4] successfully created SIBI alphabet recognition using Kinect Depth camera which allows the detection process is more accurate, with an accuracy of 96.53%, [5] successfully created SIBI alphabet recognition using DSLR camera with an accuracy of 88.8%, [6] successfully created SIBI recognition using smartphone camera, it has an increased accuracy until the highest predictive accuracy is obtained of 98.89%, but this study is limited to numbers, and [7] succeeded in making a ASL recognition with an accuracy of 99.92% by using a webcam.

Based on those studies the highest accuracy is obtained by the recognition that used a webcam, but there is yet to be system or application that can recognize SIBI by using a webcam because there still restricted on the datasets in SIBI. Whereas, by utilizing the camera that available in various type of PCs will make it easier to use and communicate. Therefore, this study aims to create SIBI datasets via hand gesture so that the sign language system or application can be used by SIBI users.

Fig. 1. Final result of hand histogram.

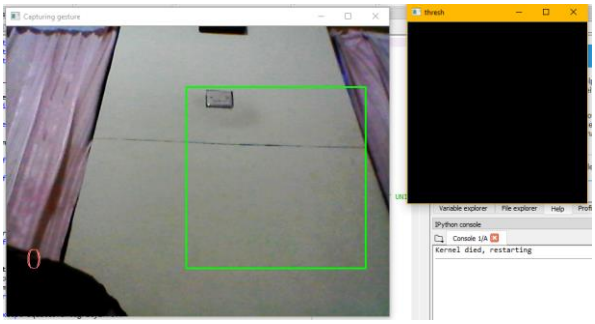

Fig. 2. SIBI datasets making process.

II. METHODS

This study used experimental method to construct SIBI Datasets for sign language recognition using a webcam, with steps consisting of hand histogram adjustment, datasets making, and datasets training and testing.

A. Hand Histogram Adjustment

Hand histogram is the pre-processing stage of the system that used to detect skin colour on the hand. Captured images in certain areas need to be resized, before converting to HSV (Hue Saturation Value). HSV is one of the colour spaces used to detect colour, it is dominance and brightness level [8], after that, the histogram image will be calculated and the image will be normalized before doing the back projection and structuring. Afterward, the picture will be convoluted, blurred using Gaussian Blur and Median Blur, later thresholding using binary and otsu.

The final result of this stage shows black and white colours. The hand or something recognized similar as skin colour that has been recorded by the webcam in the hand histogram window will display as white, while other colours will look black that can be seen in Fig. 1. Hand histogram adjustment is a process that will determine how well sign language recognition works, therefore this steps must be done properly in order to obtain a hand histogram that can detects skin colour properly.

B. SIBI Datasets Making

After hand histogram adjustment was done and a good hand histogram is obtained, SIBI Datasets were created by capturing images by webcam. It can be seen in Fig. 2, the hand that makes sign is placed on the green box, the sign must be maintained, but the position of the hand that make the sign can be shifted so that the image obtained can vary. The number of images taken by the webcam will be displayed on the window.

C. Training dan Testing SIBI Datasets

To make SIBI Datasets available for recognition, training and testing the datasets to the system or application is needed so there will be confusion matrix and recognition testing to know how well SIBI Datasets work in sign language recognition. The system or application used in this steps is the same model used in [7], the dataset was divided into train data, validation data, and test data before the training process. Then, the training and testing was conducted by using Spyder, Tensor Flow, OpenCV, and Keras in Anaconda Environment.

III. RESULTS AND DISCUSSION

SIBI is a sign language that has a very broad scope, there are alphabet and number signs that are mostly static, also static and dynamic word signs that require several movements. However, in this study, the datasets that has been created are static word signs because the datasets stored is in the form of images captured by the webcam. There are 8 static word signs of SIBI that has been created such as ‘bagus/baik’, ‘kamu’, ‘saya’, ‘nama’, ‘ingat’, ‘cinta’, ‘tolong’ and ‘aku cinta kamu’ that can be seen on Fig. 3.

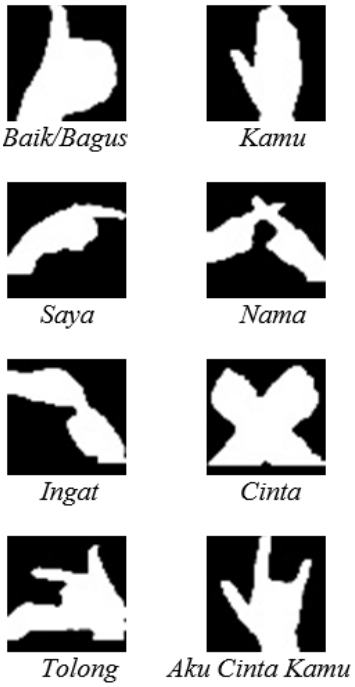

Fig. 3. SIBI datasets.

Each dataset contains 2.400 images, 1.200 of them were images captured by the webcam and the rest were vertically mirrored images. Thus, there are 19.200 images in total.

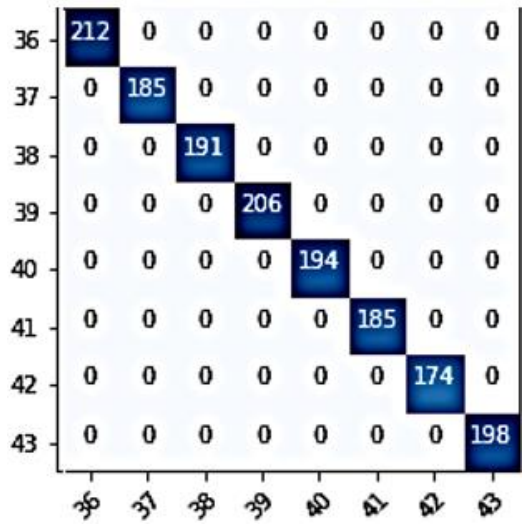

Fig. 4. Confusion matrix.

Fig. 5. Recognition process.

After that, training and testing the datasets to the system or application was done with an accuracy of 99.97%. Compared to the previous model [7], this model has better accuracy, so it can be concluded that creating datasets can improve the accuracy of the model.

After training and testing, a confusion matrix was obtained. Confusion matrix is one of the evaluation method that is often used in recognition system that utilize classification model. The performance of SIBI Datasets on test data is visualized with a confusion matrix that can be seen on Fig. 4. The column in the figure shows the target and the row shows the prediction. The confusion matrix can be explained as follows, dataset number 36 is 'baik/bagus' sign obtained 212 true prediction, number 37 is 'kamu' sign obtained 285 true prediction, number 38 is 'saya' sign obtained 191 true prediction, number 39 is 'nama' sign obtained 206 true prediction, number 40 is 'ingat' sign obtained 194 true prediction, number 41 is 'cinta' sign obtained 185 true prediction, number 42 is 'tolong' sign obtained 174 true prediction, and lastly number 43 'aku cinta kamu' sign obtained 198 true prediction.

From the confusion matrix, a precision value can be calculated by dividing the number of true prediction by the number of prediction on the target. Based on the confusion matrix, it appears that all signs have a precision of 100%, because there is no sign that has wrong prediction on the target, all of the table on the matrix except the diagonal row has zero value.

In addition, testing on the SIBI Datasets in the recognition system or application was also done by the researcher as a user five times on each signs. The steps of sign language recognition can be seen in Fig. 5. First is hand histogram adjustment, this must be done every time there is a change in lighting condition or change of place, but if the condition and place are still the same, adjustment is no longer necessary. It was previously said that this step is very important, because it will determine how well sign language recognition works, especially in the testing process. It must be ensured that the hand making the sign is fully detected, so that all parts of the hand are white and the rest is black.

Second, is the recognition process. This process used the model which obtained from training and testing results. In Fig. 6., the window that recognizes the signs can be seen along with the recorded image from the webcam, in which there is a green box to put a hand when making a sign. The hand histogram window also be displayed. To obtain a prediction result, the hand position was held inside the green box for a few seconds.

Lastly, is the recognition result. After a few seconds, the recognition result appeared on the right window. The obtained result is almost real time, when the hand makes another sign, then results will immediately change according to the recognized sign. The recognition result will not be displayed if the hand is not inside the green box.

Based on the test results on Table I, with a total of 40 experiments on all signs, it was obtained 39 true recognition results and 1 false recognition results. False recognition results came from 'nama' sign. In this test, 'nama' sign is mistakenly recognized as 'ingat'.

TABLE I. SIBI DATASETS TESTING

Sign	Result		Note
	True	*False*	
Bagus/Baik	5	0	
Kamu	5	0	
Saya	5	0	
Nama	4	1	Ingat
Ingat	5	0	
Cinta	5	0	
Tolong	5	0	
Aku cinta kamu	5	0	
Total	**39**	**1**	

This error might be caused by many things, such as, poor hand histogram, improper hand position on the green box, hand position that keeps moving, and similarity of signs. In this case, some signs have similarities to each other or almost the same. For example, based on this test result, the sign for 'nama' is the similar to 'ingat', and the test results on both signs might be exchanged with each other. The other sign, such as 'baik/bagus', 'kamu', 'saya', 'cinta', 'tolong', and 'aku cinta kamu' do not have a similarity with each other, so there will be minimum possibility of getting false recognition. Therefore, in this test, the percentage of recognition success is obtained by comparing the true recognition results and the total result of overall recognition, which is 97.5%.

Fig. 6. SIBI datasets recognition testing.

Fig. 6 shows the recognition process of 'nama' sign. The recognition should be done on a plain wall background and sufficient lighting, because these two factors greatly affect the hand histogram adjustment which is the most important step and used for the recognition process. For several times, researcher had difficulty in getting a good hand histogram since the background was not plain and the colour of the surrounding objects were similar to skin colour. This problem can be handled by using a green screen or using gloves that are plain coloured but have a contrasting colour against the background. With this solution, the webcam will easily detect the hand so that the hand histogram adjustment is easier, and the results that is obtained on the hand histogram window will also good for the recognition process.

IV. Conclusion

SIBI datasets were successfully constructed with signs that consisting of 8 static word. Based on the confusion matrix, all the signs that has been created have a precision of 100%, meanwhile based on the test results with sufficient lighting conditions and plain background, SIBI datasets has a prediction success rate of 97,5%.

Acknowledgment

The authors would like to thank the Electrical Engineering Education Department, Faculty of Technology and Vocational Skills Education, Universitas Pendidikan Indonesia.

References

[1] L. Yusnita, R. Rosalina, R. Roestam and R. Wahyu, "Implementation of Real-Time Static Hand Gesture Recognition Using Artificial Neural Network," CommIT (Communication Inf. Technol., vol. 11, no. 2, pp. 85–91, 2017.

[2] D.R. Kurnia and T. Slamet, "Menormalkan yang Dianggap 'Tidak Normal' (Studi Kasus Penertiban Bahasa Isyarat Tunarungu di Sekolah Luar Biasa [SLB] dan Perlawananya di Kota Malang)," IJDS, vol. 3, no. 1, pp. 34–43, 2016.

[3] J. Bukhari, M. Rehman, S.I. Malik and A.M. Kamboh, "American Sign Language Translation through Sensory Glove ; SignSpeak," vol. 8, no. 1, pp. 131–142, 2015.

[4] C.V. Angkoso, M. Fuad and D.R. Hadiwineka, "Pengenalan Abjad Sistem Isyarat Bahasa Indonesia (SIBI) Berbasis Kamera Depth," Link, vol. 24, no. 1, p. 6, 2018.

[5] A.A. Gafar and J.Y. Sari, "Sistem Pengenalan Bahasa Isyarat Indonesia dengan Menggunakan Metode Fuzzy K-Nearest Neighbor," Ultimatics, vol. IX, no. 2, pp. 122–128, 2017.

[6] M. Bagus, S. Bakti and Y.M. Pranoto, "Pengenalan Angka Sistem Isyarat Bahasa Indonesia Dengan Menggunakan Metode Convolutional Neural Network," pp. 11–16, 2019.

[7] S. D, Sign-Languange (Version 1): A Very Simple CNN Project, 2018,

[8] D.J. Bora, A.K. Gupta and F.A. Khan, "Comparing the Performance of L * A * B * and HSV Color Spaces with Respect to Color Image Segmentation," Int. J. Emerg. Technol. Adv. Eng., vol. 5, no. 2, pp. 192–203, 2015.