11 (1) 2025 13-40
ISSN 2502-3357 (online) | ISSN 2503-0477 (print)
Register: *Jurnal Ilmiah Teknologi Sistem Informasi*
Contents lists available at www.journal.unipdu.ac.id
Register
Journal Page is available to www.journal.unipdu.ac.id/index.php/register

Research article

# One-Way Communication System Using CNN for Interaction between Deaf and Blind People

*Juli Sulaksono [a], Ida Ayu Dwi Giriantari [b*], Made Sudarma [c], Ida Bagus Alit Swarmardika [d]*

[a,b,c,d] *Engineering Science Doctoral Programme, Udayana University, Jl. Raya Kampus Unud, Kec. Kuta Selatan, Kabupaten Badung, 80361, Indonesia*

*email:* [a,*] jsulaksono@gmail.com, [b] dayu.giriantari@unud.ac.id , [c] imasudarma@gmail.com , [d] gusalit@unud.ac.id

* *Correspondence*

**ARTICLE INFO**

***Article history:***
Received April 7th, 2024
Revised June 19th, 2024
Accepted June 20th, 2025
Available online June 30th, 2025

***Keywords:***
Deaf
Blind People
SIBI
Braille
CNN

***Please cite this article in IEEE style as:***
J. Sulaksono, I. A. D. Giriantari, M. Sudarma, and I. B. A. Swarmardika, "One-Way Communication System Using CNN for Interaction between Deaf and Blind People," *Register: Jurnal Ilmiah Teknologi Sistem Informasi*, vol. 11, no. 2, pp. 29-40, 2025.

**ABSTRACT**

Communication is essential for everyone, including for individuals who are deaf and blind. People with disabilities must have equal rights to communicate, just like the general public. A one-way communication system between deaf and blind people is therefore necessary. The input to the system is in the form of spoken language, and the output is in the form of Braille. The input uses SIBI (Indonesian Sign Language System), which is recorded with a camera and then processed using a Convolutional Neural Network (CNN). The CNN is divided into three parts: the Training Process using a Teachable machine, the SIBI DataSet model, and the Detection Process. The output of this process is text. The conversion of text into Braille is conducted using an image index. The resulting Braille can be read by blind users. System performance is analysed using a Confusion Matrix. The analysis results show an accuracy of 85%, a precision of 90%, and a recall of 82%.



## 1. Introduction

Communication is very important for everyone. Good communication can open one's mind. It is also essential for people with disabilities [1][2]. With effective communication, there can be equality between people with disabilities and non-disabled individuals. Among the most commonly encountered groups of people with disabilities in the community are deaf and blind individuals [3][4].

Deaf people are individuals who are unable to hear [5][6][7]. A lack of communication skills can negatively impact their psychological well-being, academic development, and level of independence. Deaf people communicate with each other using SIBI, which is Indonesian that has been converted into sign language. SIBI is expressed using hand movements in the area from the head to the waist. In addition to being used among the deaf community, SIBI can also be used by deaf people to communicate with non-disabled individuals because it is a visual language [8]. What becomes a big problem is when deaf people communicate with blind people. Communication between deaf and blind individuals is nearly impossible because blind people cannot see [6][9].

The novelty of this problem lies in developing a one-way translator system to transfer information from deaf individuals to blind individuals. The aim of this research is to build one-way communication between deaf and blind people. This research aims to facilitate knowledge sharing between deaf and blind individuals.

Research related to deaf and blind people, especially communication tools, remains very limited. Existing studies mainly focus on translating spoken language into sign language [10], Braille to Voice

ISSN 2502-3357 (online) | ISSN 2503-0477 (print)
11 (2) 2025 29-40

Translator [11], and the Introduction of SIBI using CNN [12]. The reason for using CNN is because if its strong capability in recognizing complex visual patterns in images. This is because CNN is able to capture important features of images through convolution layers that filter input using kernels or filters that can detect edges, textures and other patterns [12]. So far, there is no research specifically addressing the translation of SIBI into Braille, although the current research is still in its early stages. The researchers hope this research raise awareness among scientists about the needs of people with disabilities and serves as an embryo for communication tools for deaf and blind people. The significant need among deaf people for the development of this communication tool is the reason for the researchers to develop this translator.

The input of this system is the video footage of SIBI gestures performed by deaf people. These SIBI videos are processed using machine learning models built with Teachable Machine, which utilizes a convolutional neural network (CNN) algorithm to create a machine learning model. The SIBI words that were input to the Teachable Machine totaled 14 words. Each word was captured in three stages. The stages of taking pictures include front view, left-side view, and right-side view. The images were used as training data. The testing data for the communication systems consists of SIBI gestures performed by deaf people during real communication scenarios [13][14]. The testing data is then recognized by matching with the training data. If the system can recognize the testing data and convert it to text, it converts the result into text, which is then translated into Braiile letters. The system is analyzed using the confusion matrix method. This method allows us to assess the accuracy, precision, and recall of a data. [15][16][17].

## 2. Materials and Methods

### 2.1. Materials

The research data refers to the 2002 SIBI dictionary issued by the Education Office. The data in the book is in the form of images. In this process, the researchers converted the images into videos. There were three video actors with different ages and postures. The purpose of choosing three people with different postures is to represent several physical forms of people [16][18].

Video recording is conducted indoors using a Logitect C270 camera [19]. The purpose of using this camera was to clarify the results when filming. The distance between the camera and the subject is 2 meters [20]. The filming scenarios are made in two types. The first scenario was taken outdoors, and the clothes used by the actors were not determined. The reason for using two scenarios is to obtain the best shooting scenario, so that the system can translate correctly. If, when collecting data, the scenario used is not correct, the accuracy, precision, and recall of the system will be very poor. The following figures illustrate the process of taking pictures in the first scenario.

Fig. 1. Video filming of Scenario 1

Fig. 2. Video filming of Scenario 2

Referring to Figure 1, the first scenario takes place outdoors, the cloth worn by the model is white. In Scenario 1, lighting is done using lamps. Referring to Figure 2, the second scenario takes place indoors, the clothes worn by the model are white. The following section describes the shooting process for the second scenario. The SIBI words used in both scenarios can be seen in the table 1.

### 2.2. Methods

One-Way Communication System for the Deaf and Blind invividuals requires detailed and structured steps in processing SIBI words into Braille letters. The flowchart can be seen in Figure 3.


http://doi.org/10.26594/register.v11i2.4556

ISSN 2502-3357 (online) | ISSN 2503-0477 (print)
11 (2) 2025 29-40

Table 1. SIBI words taken

| No | Word |
|---|---|
| 1 | Me |
| 2 | Eat |
| 3 | Mangosteen |
| 4 | Chicken |
| 5. | Congratulations |
| 6. | Morning |
| 7. | You |
| 8. | Love |
| 9. | I |
| 10. | Like |

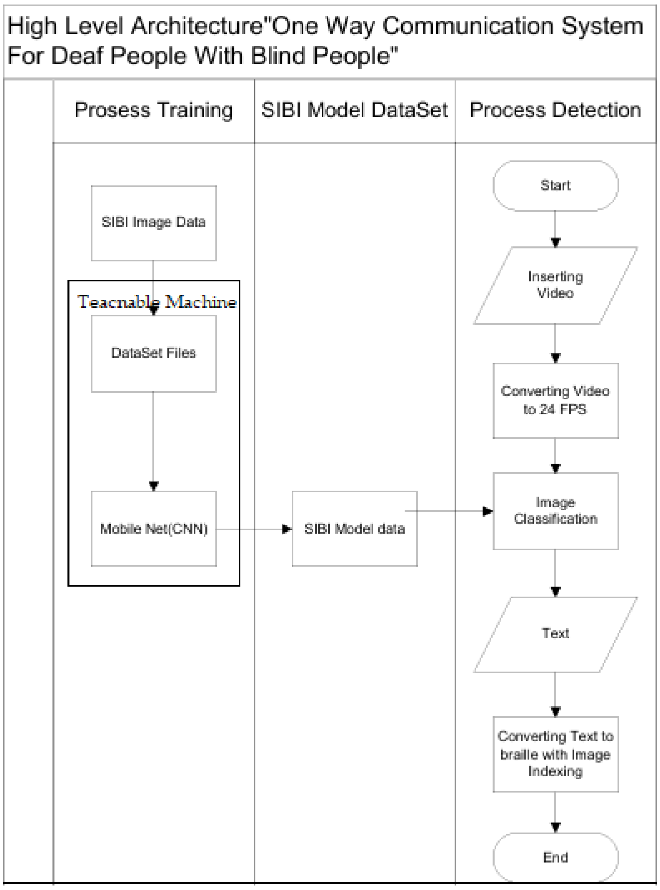


Fig 3. Flowchart of One-way Communication System for Deaf and Blind People

### 2.2.1 Process Training

Training Data, also known as training datasets, learning sets and training sets, is a part of a collection of datasets that are provided to be the learning material for the model. This allows the model to generalize patterns in the data so that it can later be used to predict new inputs [21]. The training data consists of 10 SIBI words, including both base words and affixed words. The words used are eat, mangosteen, chicken, congratulations, you, morning, love, me, ber, and ball.

These words were taken in video format and then converted into images. Each video was was converted into 540 images. the reason for using 540 images per word is to reduce the risk of overfitting. If overfitting occurs, the system may not be able to recognize new images properly. Overfitting can occur for several reasons, including (a) The size of the training data is too small and does not contain enough data samples to accurately represent all possible input data values, (b) The training data contains a large amount of irrelevant information or meaningless information, (c) The model is trained for too long on a set of sample data. The epochs used in the training process amounted to 50, while the BatSize was 16. Figure 4 is an overview of the Training Process.


http://doi.org/10.26594/register.v11i2.4556

ISSN 2502-3357 (online) | ISSN 2503-0477 (print)
11 (2) 2025 29-40

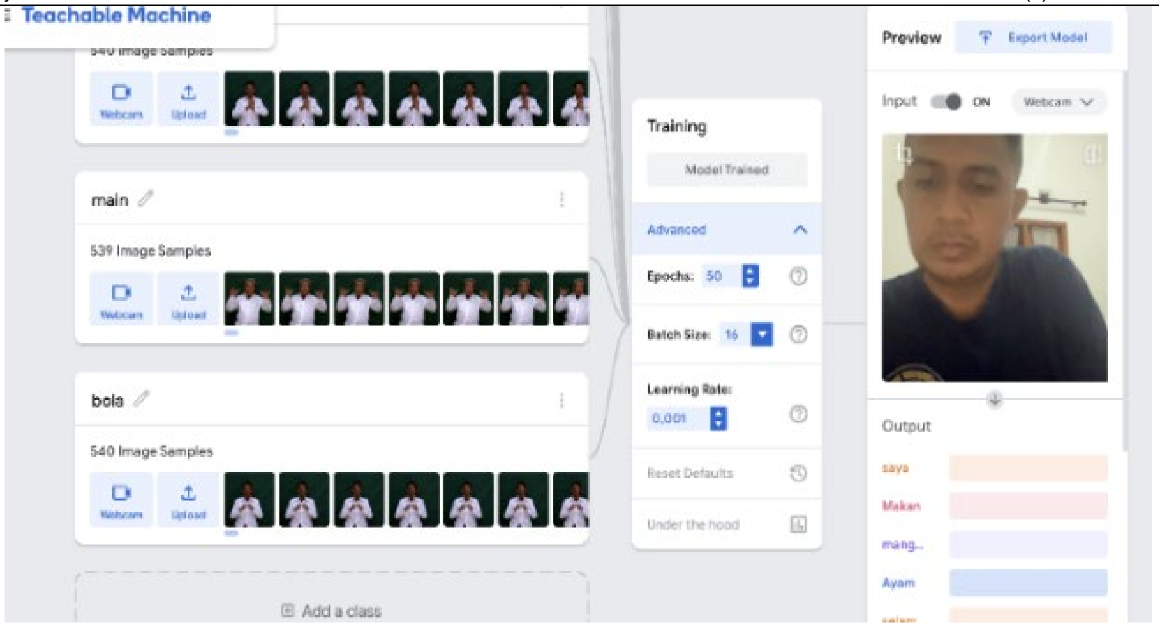


Fig 4. Training Process.

Figure 4 illustrates the algorithm used is Convolutional Neural Network (CNN). CNN is a class of deep learning that is capable of image recognition and image classification. The CNN method is a class of neural networks that specializes in processing data that has a grid-like topology, such as images. The CNN method can be used in face recognition, document analysis, image classification, video classification. The following steps outline the use of training data in this study.

**SIBI Image data Set**

This is the first step in the training process. An image dataset that contains various hand signals or Indonesian sign language (SIBI) gestures is used. This dataset contains images of various hand gestures representing words or phrases in sign language. This dataset is used as training data to teach computer models to recognize and classify sign language signs.

**Data SetFile**

The SIBI image that has been entered into the data set will later be converted into a Data SetFile. This Data SetFile is then processed and trained with mobile net (CNN)

**Mobile Net(CNN)**

MobileNet is a CNN architecture developed by Google. The advantage of using MobileNet is it is efficient in terms of computing resource usage. At this stage, the SIBI data set is formed into training data.

**2.2.2 SIBI Model Data Set**

In Teachable Machine, the method used is Transfer Learning, and the model architecture applied is MobileNet as the "backbone" or foundation of the model. A set of classes or labels were then defined to represent the SIBI signals to be recognized. the model was trained using the uploaded SIBI image dataset. The training process allows the model to learn to recognize patterns and features in images that match the labels assigned.

The Convolutional Neural Network (CNN) algorithm consists of neurons that are designed to work similarly to the frontal lobe, specifically the visual cortex area in human and animal brains [22], [23]. Visual cortex is the area responsible for processing information in the form of visual stimuli. This makes CNN particularly effective in image processing compared to similar neural network algorithms [24]. Convolution in CNN is a mathematical operation on two functions which then produces a third function. This function combines two sets of information and shows how the shape of one function is modified by the other.

In the model formation process, the input data consists of the word "I (me)" in SIBI form. This word is used to match the input data with the training data. In this research, the number of Epochs used


http://doi.org/10.26594/register.v11i2.4556

ISSN 2502-3357 (online) | ISSN 2503-0477 (print)
11 (2) 2025 29-40

is 50, with a the Batch Size of 16 and a learning rate of 0.0001. A more detailed illustration is provided in the figure 5.

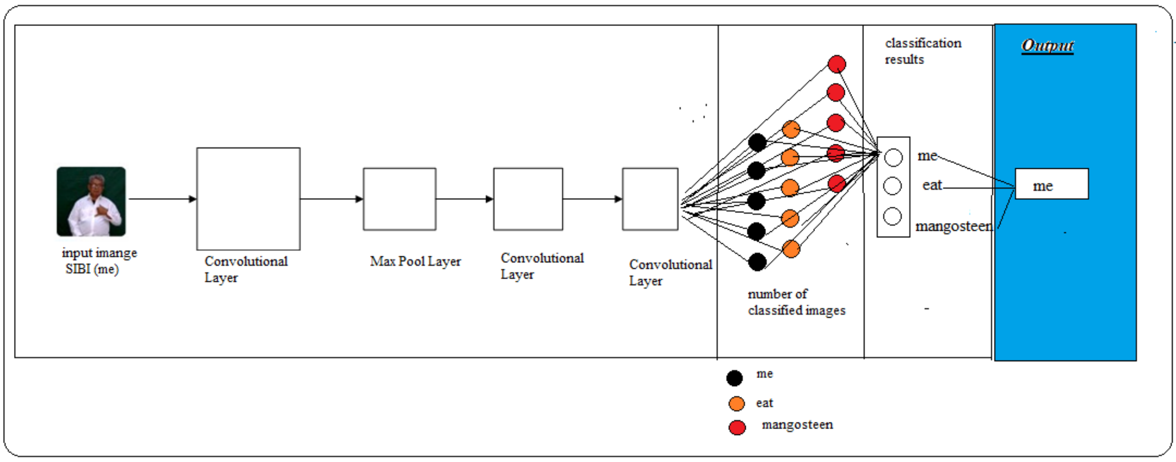


Fig 5.Image Recognition Process with CNN

The input from Figure 5 is the SIBI display. The video is first processed through the Colovutional Layer process, followed by the Max Pool Layer, and then passed through another Convolutional Layer again. Next, the testing data is compared with the training data. Once a match is found, the CNN determines the corresponding output.

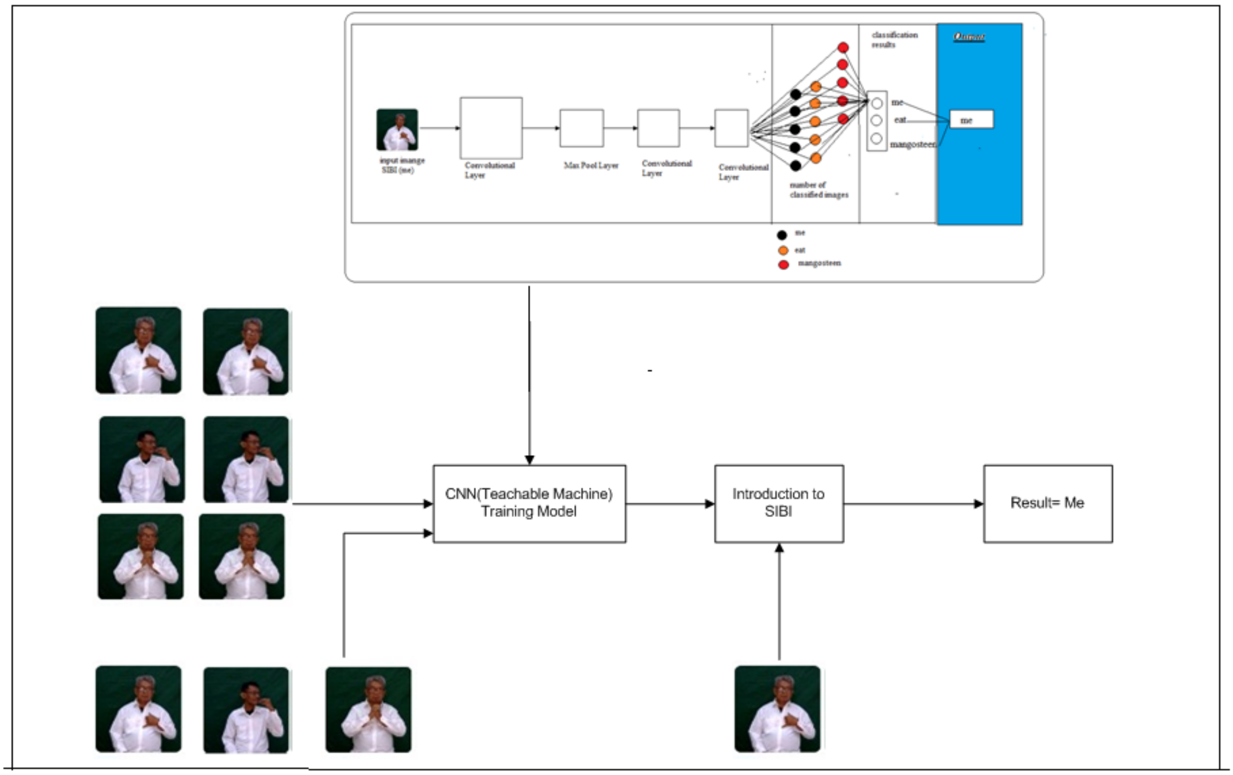


Fig 6. Convolutional Neural Network algorithm for SIBI recognition

Here is an explanation of Figure 6. This study uses Teachable Machine, which uses CNN algorithm to classify images. The data used for training data consists of 10 SIBI images, and each SIBI image consists of 540 images. The results of this labeling will later be used as a SIBI training model with CNN. Once the model is created, the next step is to enter the SIBI recognition process. The SIBI image data entered is processed by matching it with the training data. If the data is appropriate and correct, the system can determine that the image corresponds to a specific SIBI word.

**2.2.3 Process Detection**

**Video (Input With Camera)**

ISSN 2502-3357 (online) | ISSN 2503-0477 (print)
11 (2) 2025 29-40

If deaf people want to communicate with blind people, they can use this system. The method involves the deaf person expressing information in SIBI form, which the system will automatically recognize [25][26].

**Change video to 24 fps**

This section converts the video to images, as CNN is unable to recognize videos. Each SIBI video capturing each movement lasts 6 seconds. Captured videos are converted into 24 frames. One of these frames is then tested and matched to the training data.

**Image Classification**

This section is used to recognize SIBI Images. In this process, the input images are classified by comparing them with the images stored in the SIBI Model DataSet. Once classified, the SIBI image can be recognized. The output of this process is text

**Text**

After processing by the CNN, sentences in the form of SIBI are converted into text; this text is then converted into Braille letters. Each text entry has a unique id, which is later used in the image indexing process.

**Converting Text to Braille with Image Indexing**

Image indexing is the process of converting text into braille images. In this process, indexing is done for each letter, because Braille represents characters one by one. After the characters are converted into letters, the baille characters are combined into words. After that, the braille image is printed and read by deaf people. A more detailed explanation is provided below.

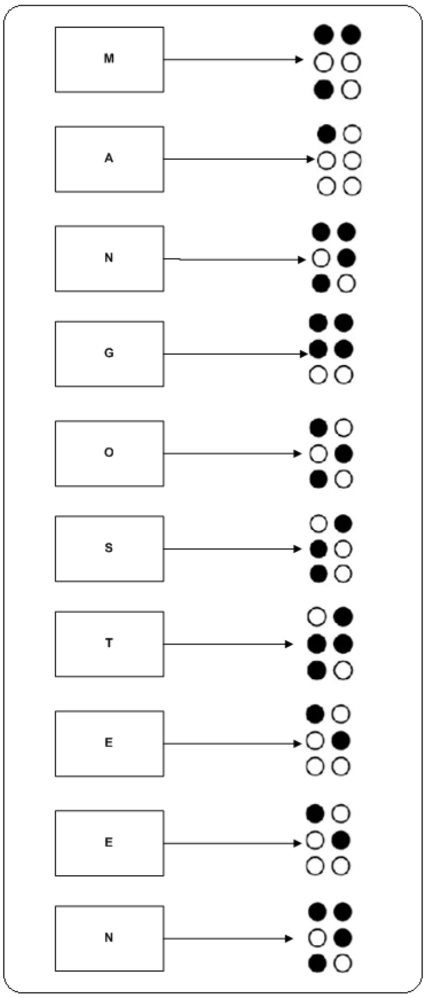


Fig 7. Indexing words with SIBI

Figure 7 illustrates that blind people, while unable to see, are able to hear well. They are able to read special writing designed for the blind, namely Braille. Therefore, this study utilizes braille to


http://doi.org/10.26594/register.v11i2.4556

ISSN 2502-3357 (online) | ISSN 2503-0477 (print)
11 (2) 2025 29-40

provide information to the blind. The following is the process of converting text into Braille. Figure 8. displays the result of converting text to Braille. The word shown as output is "mangosteen."

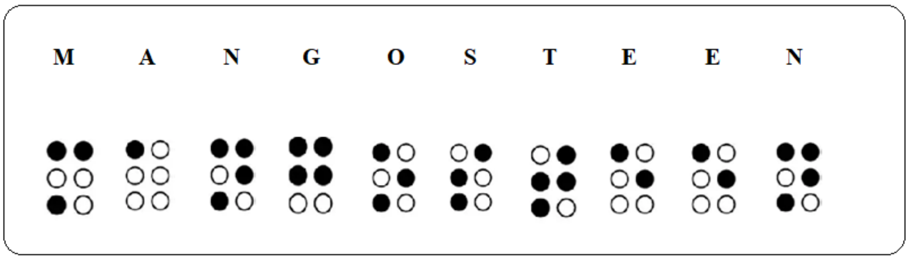


Fig 8. Output System

## 3. Results and Discussion

### 3.1 Shooting Method

The video recording was conducted using two scenarios. In the first scenario, there were no specific constraints. The first scenario was carried out outdoors without a background), the clothes worn by the actors are casual, and the environment was open. The second scenario used strict restrictions in taking pictures using a green background. The reason for using a green background was because proficient SIBI performers use green. The actors wore white clothing, which was chosen because white contrasts well with green, making gestures clearer. This decision was made after direct consultation with a resource person, Mrs. Masunah, a teacher at SMP-LB in Surabaya. The use of a green background and indoor setting also ensured more consistent lighting.

Table 2. Shooting Scenario

| Id | Scenario | Background | Color of clothes | place |
|---|---|---|---|---|
| 1. | First, no free outside | There aren't any | First, no free outside | outdoor |
| 2. | Third green color: white Inside the room | green color | Third green color: white Inside the room | indoor |

Referring to Table 2, there are two methods of capturing images. The number of words captured is 14 words. The following are the SIBI words captured

Table 3. Words in the data set

| Id | Word | Id | Word |
|---|---|---|---|
| 1. | Me | 8. | Love |
| 2. | Eat | 9. | I |
| 3. | Chicken | 10. | Like |
| 4. | Mangosteen | 11. | Ber |
| 5. | Congratulations | 12. | Play |
| 6. | Morning | 13. | Ball |
| 7. | You | 14. | Vegetable |

Table 3 shows that the method used to see the accuracy, precision and recall values is the Confusion Matrix [27]. The reason for using a confusion matrix for analysis is because the confusion matrix can test the effectiveness of deep learning. The equation is as follows.

$$\text{Accuracy} = \frac{(TP+TN)}{TP+TN+FP+FN} \tag{1}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{2}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{3}$$

Information:
TP = Correct SIBI Word recognized correctly (Successful)
FP = Correct SIBI word recognized as the wrong word (Failing to recognize)
FN = Incorrectly recognized SIBI word , although the input was correct (Misrecognized)
TN = Incorrect SIBI word correctly recognized as wrong (True)


http://doi.org/10.26594/register.v11i2.4556

ISSN 2502-3357 (online) | ISSN 2503-0477 (print)
11 (2) 2025 29-40

### 3.1.1 First method of taking pictures

In the first scenario, no specific restrictions were applied. The shooting was done in an open space, and the clothing worn by the model was not specified. The testing process involved inputting 10 correct SIBI words and 10 incorrect SIBI words. The results of the analysis, using the Confusion Matrix, can be seen in the Table 4.

Table 4. Confusion Matric of each word of scenario one

| Id | Word | Number of Data Sets | Input Data | TP | FP | FN | TN | Accuracy | Precision | Recall |
|---|---|---|---|---|---|---|---|---|---|---|
| **1** | Eat | 416 | 20 | 3 | 7 | 6 | 4 | 0,35 | 0,30 | 0,33 |
| **2** | Chicken | 416 | 20 | 5 | 5 | 3 | 7 | 0,60 | 0,50 | 0,63 |
| **3** | Mangosteen | 416 | 20 | 6 | 4 | 6 | 2 | 0,44 | 0,60 | 0,5 |
| **4** | Congratulations | 416 | 20 | 7 | 3 | 6 | 4 | 0,55 | 0,70 | 0,54 |
| **5** | Morning | 416 | 20 | 4 | 6 | 8 | 2 | 0,30 | 0,40 | 0,33 |

The results from this table are presented in the graph as follows.

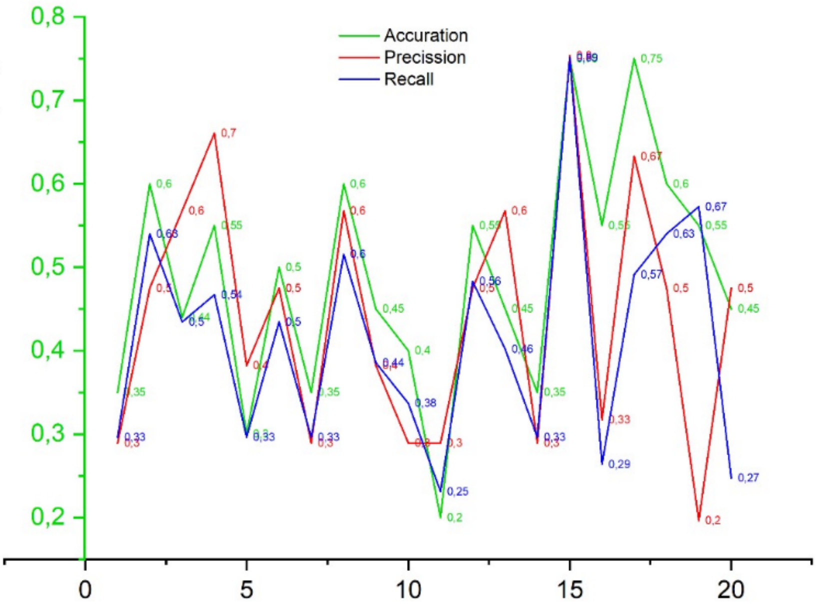


Fig 9. Confusion Matrix for data collection for the first scenario

Figure 9 shows the results of Scenario 1, which have not met the target outcome. In this scenario, the input SIBI images were frequently not recognized as SIBI images. Moreover, an incorrect image is sometimes recognized as a correct SIBI image. The average results of the confusion matrix are: Accuracy: 0.44, Precision: 0.45, and Recall: 0.44. Considering these results, significant improvement needs to be conducted, especially in the video shooting properties.

### 3.1.2 Video capture with Scenario 2

In the second scenario, the shooting was conducted indoors to ensure the lighting condition.

Table 5. Confusion matrix of each word in the second scenario

| Id | Word | Number of Data Sets | Input Data | TP | FP | FN | TN | Accuracy | Precision | Recall |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Eat | 416 | 20 | 8 | 2 | 1 | 9 | 0,85 | 0,80 | 0,89 |
| 2 | Chicken | 416 | 20 | 7 | 3 | 0 | 10 | 0,85 | 0,70 | 1,00 |
| 3 | Mangosteen | 416 | 20 | 10 | 0 | 1 | 9 | 0,95 | 1,00 | 0,91 |
| 4 | Congratulations | 416 | 20 | 8 | 2 | 2 | 8 | 0,80 | 0,80 | 0,80 |
| 5 | Morning | 416 | 20 | 7 | 3 | 3 | 7 | 0,70 | 0,70 | 0,70 |


http://doi.org/10.26594/register.v11i2.4556

ISSN 2502-3357 (online) | ISSN 2503-0477 (print)
11 (2) 2025 29-40

The results from this table are presented in a graph as follows.

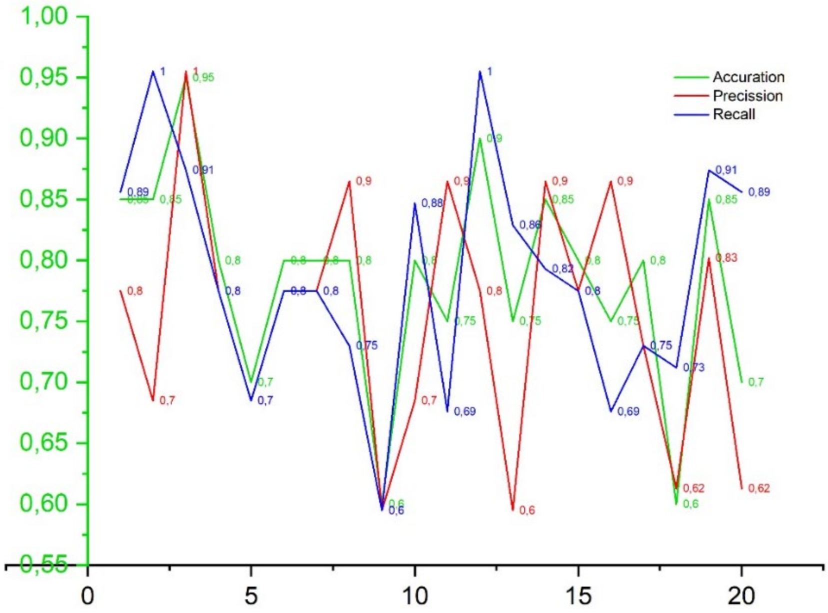


Fig 10. Confusion Matrix for data collection for the second scenario

Additionally, the model wore white clothing, and the background used was green. The testing process involved inputting 10 correct SIBI words and 10 incorrect SIBI words. The results of the analysis using Confusion Matrix can be seen in the Table 5. Figure 10 presents the results of Scenario 2, which yielded the most optimal performance. The accuracy, precision and recall values are notably high. The values obtained are as follows: Accuracy: 0.80, Precision: 0.79, and Recall: 0.82. Baed on these results, this scenario is selected as the basis for modeling SIBI recognition using CNN.

**3.2 Thorough Testing of System Accuracy, Precision, and Recall**

System testing was carried out by inputting two syllables into the system. These inputs correspond to SIBI movements demonstrated by a reenactor. Once the reenactor performs the designated SIBI gesture, the movements were automatically recorded using a PC camera. The system then recognized the recorded movement using a CNN-based recognition model. After the movement is recognized, the system automatically converts the image into text. The resulting text was then converted into Braille letters, enabling recognition by blind users. Before the system was tested on deaf and blind people, it was necessary to analyze the accuracy, precision, and recall using a Confusion Matric. The results of this evaluation are presented in the following system test table.

Table 6. System operation trial

| Id | Sentence | Status | |
|---|---|---|---|
| 1. | I eat mangosteen | success | TP |
| 2. | I ate chicken | success | TP |
| 3. | I play soccer | success | TP |
| 4. | Good morning | success | TP |
| 5. | Like mangosteen | success | TP |
| 6. | You eat | failed | FP |
| 7. | the day after tomorrow I bought mangosteen | success | TN |
| 8. | I'm leaving again | success | TN |
| 9. | I learn | failed | FN |
| 10. | good luck with the exam | success | TN |

Table 6 presents the results of system operation trials. The results of the system are categorized into 4 components: True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN). By utilizing Confusion Matrix, the data recorded in Table 5 can be processed so that accuracy, precision, and recall can be determined, providing a clearer and more detailed understanding of the system's performance. The results of the entire trial can be seen in the Figure 11.


http://doi.org/10.26594/register.v11i2.4556

ISSN 2502-3357 (online) | ISSN 2503-0477 (print)
11 (2) 2025 29-40

| N=20 | True | False |
|---|---|---|
| True | TP=9 | FP=1 |
| False | FN=2 | TN= 8 |

Fig 11. Confusion Matrix Data

Figure 11 visualizes of the results from the system operation trials. Similar to Table 6, the results are divided into True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN). By utilizing Confusion Matrix, the data recorded in Figure 6. can be processed so that accuracy, precision, and recall can be ensured, making the results of the trial clearer and more detailed. The complete trial results are illustrated in the figure below.

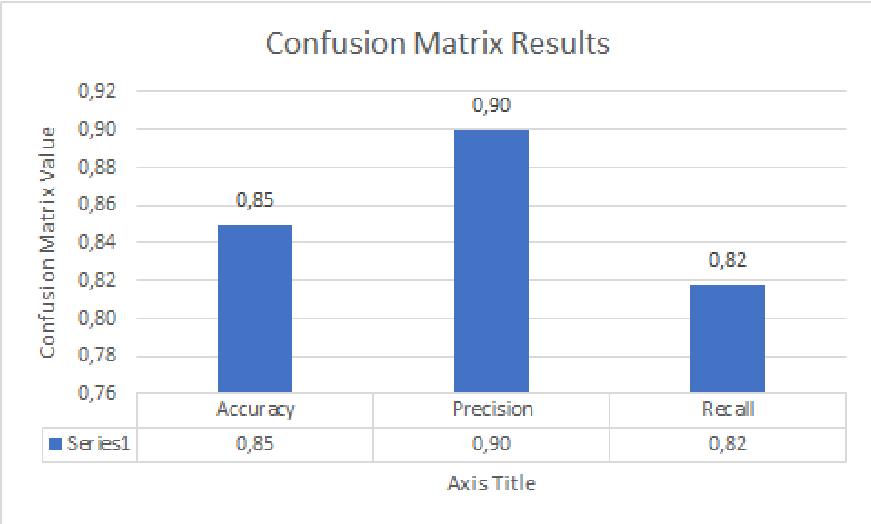


Fig 12. Confusion Matrix results for the entire system

Figure 12 presents the calculation results with Confusion Matrix, obtaining an accuracy value of 85%, precision of 90%, and recall of 82%. Based on the recall value, it can be concluded that the system is capable of accurately recognizing each SIBI word and successfully converting sentences into Braille and audio output.

## 4. Conclusion

The One-Way Communication System Using CNN for Interaction between Deaf and Blind People is designed to bridge communication between people with hearing and visual impairments. The system takes SIBI as input and produces Braille text as output. The system uses CNN to recognize SIBI gestres. In this study, the CNN model is divided into 3 stages, namely the Training Process, SIBI Model DataSet, and Detection Process. The camera is used to record SIBI gestures, which are then processed by the CNN. The SIBI gestures were then converted into the words (text). Then the resulting text is converted into Braille. System performance was analyzed using the Confusion Matrix, showing an overall accuracy of 85%, precision of 90%, and recall of 82%. Because the recall value is still below 85%, the system still needs improvement.

For future development, the system is very suitable if implemented in special education schools, especially those with heterogeneous student populations including both deaf and blind individuals. It is expected that this system can help facilitate more effective communication among people with disabilities.

## Author Contributions

J. Sulaksono: Conceptualization, methodology, validation, and writing – review & editing. I. A. D. Giriantari: Funding acquisition, investigation, project administration,resources, and visualization. M. Sudarma: Data curation, software, and writing – original draft. I. B. A. Swarmardika Formal analysis, supervision, and writing – review & editing.


http://doi.org/10.26594/register.v11i2.4556

ISSN 2502-3357 (online) | ISSN 2503-0477 (print)
11 (2) 2025 29-40

**Acknowledgment**

We would like to express our gratitude to University Nusantara PGRI Kediri and Udayana University for their valuable support in completing this research, in terms of material, mental, and moral support.

**Declaration of Competing Interest**

We declare that we have no conflict of interest.

## References

[1] Y. Xu, D. Wang, P. Collins, H. Lee, and M. Warschauer, "Same benefits, different communication patterns: Comparing Children's reading with a conversational agent vs. a human partner," *Comput. Educ.*, vol. 161, p. 104059, 2021, doi: https://doi.org/10.1016/j.compedu.2020.104059.

[2] K. L. Manikanta, N. Shyam, and S. S, "Real-Time Hand Gesture Recognition and Sentence Generation for Deaf and Non-Verbal Communication," in *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)*, 2024, pp. 1–6. doi: 10.1109/ADICS58448.2024.10533502.

[3] R. S. C. K. Dharmarathne, K. A. Medagedara, N. B. W. N. Madhubashinee, P. T. Maitipe, D. Sriyaratna, and K. Abeywardena, "STEP UP: Systematically Motivating the Children with Low Psychological Maturity Level and Disabled Children using Gamification and Human Computer Interaction," in *2022 IEEE 7th International conference for Convergence in Technology (I2CT)*, 2022, pp. 1–6. doi: 10.1109/I2CT54291.2022.9824182.

[4] S. J. Khalid and I. A. Ali, "Mind Controlled Educational Robotic Toys for Physically Disabled Children: A Survey," in *2022 International Conference on Computer Science and Software Engineering (CSASE)*, 2022, pp. 348–354. doi: 10.1109/CSASE51777.2022.9759670.

[5] A. Young, R. Oram, and J. Napier, "Hearing people perceiving deaf people through sign language interpreters at work: on the loss of self through interpreted communication," *J. Appl. Commun. Res.*, vol. 47, no. 1, 2019.

[6] I. W. Leigh, J. F. Andrews, C. A. Miller, and J.-L. A. Wolsey, "Deaf People and Society : Psychological, Sociological, and Educational Perspectives," *Deaf People Soc.*, Dec. 2022, doi: 10.4324/9781003183686.

[7] P. Pryandi, M. B. Dewantara, H. L. H. S. Warnars, A. Ramadhan, N. Noordin, and F. H. A. Razak, "Smartphone Application for the Deaf and the Deaf Caring Community," in *2023 International Conference on Inventive Computation Technologies (ICICT)*, 2023, pp. 767–773. doi: 10.1109/ICICT57646.2023.10134458.

[8] R. A. Ramadhani, I. K. G. D. Putra, M. Sudarma, and I. A. D. Giriantari, "Database of Indonesian Sign Systems," *Int. Conf. Smart Green Technol. Electr. Inf. Syst.*, 2018.

[9] H. A. Nugroho, M. Yusro, and W. Djatmiko, "Electronic dictionary of Indonesian sign systems (ED-ISS) based on Raspberry Pi as learning media for deaf students," *AIP Conf. Proc.*, vol. 2331, no. 1, p. 030004, Apr. 2021, doi: 10.1063/5.0041698.

[10] R. A. Ramadhani, "A new technology on translating Indonesian spoken language into Indonesian sign language system | Ramadhani | International Journal of Electrical and Computer Engineering (IJECE)." https://ijece.iaescore.com/index.php/IJECE/article/view/23386 (accessed Jul. 08, 2022).

[11] J. C. Hailianto and A. A. Purwita, "Wonder Reader: Development of a Refreshable Braille Display Using a Carriage System and Self-Locking Braille Cells," in *2023 IEEE 13th International Conference on Consumer Electronics - Berlin (ICCE-Berlin)*, 2023, pp. 41–46. doi: 10.1109/ICCE-Berlin58801.2023.10375659.

[12] A. N. Sihananto, E. M. Safitri, Y. Maulana, and F. Fakhruddin, "INDONESIAN SIGN LANGUAGE IMAGE DETECTION USING CONVOLUTIONAL NEURAL NETWORK ( CNN ) METHOD," pp. 13–21, 2023.

[13] I. Ketut, G. Dharma Putra, and R. A. Ramadhani, "Utilization of Fuzzy Ontology for the Meaning of Homonymous and Homophones Ambiguous Sentences," *Qubahan Acad. J.*, vol. 3, no. 4, pp. 234–244, Nov. 2023, doi: 10.48161/QAJ.V3N4A176.

[14] G. Wimmer, M. Häfner, and A. Uhl, "Improving CNN training on endoscopic image data by extracting additionally training data from endoscopic videos," *Comput. Med. Imaging Graph.*, vol.

ISSN 2502-3357 (online) | ISSN 2503-0477 (print)
11 (2) 2025 29-40

86, p. 101798, 2020, doi: https://doi.org/10.1016/j.compmedimag.2020.101798.

[15] R. A. Ramadhani, I. K. G. D. Putra, M. Sudarma, and I. A. D. Giriantari, "Detecting Indonesian ambiguous sentences using Boyer-Moore algorithm," *TELKOMNIKA (Telecommunication Comput. Electron. Control.*, vol. 18, no. 5, pp. 2480–2487, Oct. 2020, doi: 10.12928/TELKOMNIKA.V18I5.14027.

[16] R. A. Ramadani, I. K. G. D. Putra, M. Sudarma, and I. A. D. Giriantari, "A new technology on translating Indonesian spoken language into Indonesian sign language system," *Int. J. Electr. Comput. Eng.*, vol. 11, no. 4, pp. 3338–3346, Aug. 2021, doi: 10.11591/IJECE.V11I4.PP3338-3346.

[17] A. Luque, M. Mazzoleni, A. Carrasco, and A. Ferramosca, "Visualizing Classification Results: Confusion Star and Confusion Gear," *IEEE Access*, vol. 10, pp. 1659–1677, 2022, doi: 10.1109/ACCESS.2021.3137630.

[18] Suharjito, N. Thiracitta, and H. Gunawan, "SIBI Sign Language Recognition Using Convolutional Neural Network Combined with Transfer Learning and non-trainable Parameters," *Procedia Comput. Sci.*, vol. 179, pp. 72–80, 2021, doi: https://doi.org/10.1016/j.procs.2020.12.011.

[19] S. B. Abdullahi, K. Chamnongthai, V. Bolon-Canedo, and B. Cancela, "Spatial–temporal feature-based End-to-end Fourier network for 3D sign language recognition," *Expert Syst. Appl.*, vol. 248, p. 123258, 2024, doi: https://doi.org/10.1016/j.eswa.2024.123258.

[20] J. Liu, W. Xue, K. Zhang, T. Yuan, and S. Chen, "TB-Net: Intra- and inter-video correlation learning for continuous sign language recognition," *Inf. Fusion*, vol. 109, p. 102438, 2024, doi: https://doi.org/10.1016/j.inffus.2024.102438.

[21] A. García-Pérez *et al.*, "CNN-based in situ tool wear detection: A study on model training and data augmentation in turning inserts," *J. Manuf. Syst.*, vol. 68, pp. 85–98, 2023, doi: https://doi.org/10.1016/j.jmsy.2023.03.005.

[22] X. Guo *et al.*, "Multimodel ensemble estimation of Landsat-like global terrestrial latent heat flux using a generalized deep CNN-LSTM integration algorithm," *Agric. For. Meteorol.*, vol. 349, p. 109962, Apr. 2024, doi: 10.1016/J.AGRFORMET.2024.109962.

[23] H. Xia, R. Zhu, H. Yuan, and C. Song, "Rapid quantitative analysis of cotton-polyester blended fabrics using near-infrared spectroscopy combined with CNN-LSTM," *Microchem. J.*, vol. 200, p. 110391, May 2024, doi: 10.1016/J.MICROC.2024.110391.

[24] J. Teuwen and N. Moriakov, "Convolutional neural networks," *Handb. Med. Image Comput. Comput. Assist. Interv.*, pp. 481–501, Jan. 2020, doi: 10.1016/B978-0-12-816176-0.00025-9.

[25] F. Hawlader, F. Robinet, and R. Frank, "Leveraging the edge and cloud for V2X-based real-time object detection in autonomous driving," *Comput. Commun.*, vol. 213, pp. 372–381, 2024, doi: https://doi.org/10.1016/j.comcom.2023.11.025.

[26] J. Sulaksono, I. A. D. Girinatari, M. Sudarma, and I. B. A. Swarmardika, "SIBI Syllable Recognition System With LSTM," in *2023 International Conference on Smart-Green Technology in Electrical and Information Systems (ICSGTEIS)*, 2023, pp. 94–97. doi: 10.1109/ICSGTEIS60500.2023.10424084.

[27] G. Phillips *et al.*, "Setting nutrient boundaries to protect aquatic communities: The importance of comparing observed and predicted classifications using measures derived from a confusion matrix," *Sci. Total Environ.*, vol. 912, p. 168872, 2024, doi: https://doi.org/10.1016/j.scitotenv.2023.168872.