

# Convolutional Implementation Neural Network (CNN) on the System and Introduction Sign Language to Support Communication with People with Disabilities

Mashandy<sup>1</sup>, Imran Lubis<sup>2</sup>, Magdalena Simanjuntak<sup>3</sup>

<sup>1,2,3</sup>STMIK Kaputama

[m.mashandy12345@gmail.com](mailto:m.mashandy12345@gmail.com), [imran.loebis.medan@gmail.com](mailto:imran.loebis.medan@gmail.com), [magdalena.simanjuntak84@gmail.com](mailto:magdalena.simanjuntak84@gmail.com)

## Abstract

The results of this study are based on verifying communication solutions for deaf and hard of hearing participants in Indonesia, the majority of whom use the Indonesian Sign Language System (SIBI). The goal is to design an accurate and efficient system for detecting and recognizing sign language using Convolutional Neural Network (CNN) and accessible in real-time on mobile devices to support more inclusive communication. This study uses a quantitative approach and uses deep learning techniques by utilizing a SIBI image dataset totaling 5,280 images (24 classes without the letters J and Z) with a division of 80% for training data and 20% for validation data. This research process starts from image preprocessing, designing a CNN architecture consisting of convolution, pooling, and fully connected layers, implementing it systematically with Python–TensorFlow, and integrating it into the Flutter application through APIs. Evaluation was conducted using Intersection over Union (IoU) metrics and measured algorithm performance with precision, recall, and mean Average Precision (mAP). The model is capable of achieving a precision of 0.97, a recall of 1.00 and a high mAP in a wide range of backgrounds and lighting.

**Keywords:** Sign Language, SIBI, CNN, Deep Learning, Image Detection, Image Processing, TensorFlow, Flutter, mAP

## 1. Introduction

Communication is certainly the main foundation in human interaction, the exchange of ideas, information, and emotions. However, for people with hearing and speech disabilities or commonly known as the deaf and speech-impaired, communication is often a complex challenge that hinders their participation in social, educational, and professional life.

According to data from the World Health Organization (WHO) in 2019, it is estimated that there are 466 million people in the world with hearing loss, of which 34 million are children. As many as 360 million or around 5.3% of the world's population are deaf. Based on BPS data in 2023, the number of people with disabilities in Indonesia is 22.97 million people, or around 8.5% of the total population. which depends on the Indonesian Sign System (SIBI) or local sign language. Although sign language has become a vital solution, the lack of understanding of it by the general public still creates a significant communication gap [1].

In the application of Deep Learning with the CNN method or Convolutional Neural Network to detect hand movements in the sign language of the SIBI alphabet and translate it [2]. It is hoped that the results of this research can be the basis for the development of an optimized sign language recognition application specifically for SIBI, and can support people with disabilities and the general public to communicate more effectively.

## 2. Literatur Riview

### 2.1. Sign Language

Sign language is a communication language for deaf people in conveying and exchanging information. However, the communication process is carried out manually with the help of hand, lip and body movements. Sign language is a unique language, as it differs in each country. In Indonesia, there are two categories of sign language development, namely, SIBI (Indonesian Sign System) sign language and BISINDO (Indonesian Sign Language). BISINDO is formed and develops naturally in a deaf environment. The language develops according to the understanding of deaf people from various backgrounds. So that this language became a very early language for them to the point that it was also called the mother tongue for the deaf. BISINDO does not use the Indonesian suffix structure like [3].

## 2.2. Sibi (Indonesian Sign System)

SIBI (Indonesian Sign Language) is a sign language that adapts ASL (American Sign Language) movements. The SIBI alphabet language has 26 movements representing 26 alphabets, consisting of 24 demonstrations in the form of passive movements, and 2 other demonstrations in the form of active movements (J and Z) using one hand. SIBI is not something new in formal communication, because its application is also used in the teaching and learning process at SLB (Extraordinary School). Regulation of the Minister of Education of the Republic of Indonesia number 0161/U/1994 regulates the use of SIBI in SLB (Extraordinary School) [3].

## 2.3. Machine Learning

Machine learning is useful for mimicking or replacing human behavior to automate or solve problems. Machine learning seeks to mimic the way intelligent beings or human processes learn and generalize. Machine learning has at least two main functions: Classification and prediction are an integral part of machine learning.

## 2.4. Flutter

Flutter is a cross-platform framework that is meant to create incredibly powerful mobile apps. Flutter was developed by Google, first and introduced to the public in 2016. Google chose flutter as the application framework for the next-generation operating system. In addition, flutter is compatible with the Fuchsia operating system, which is supported by apps designed for android and iOS operating systems.

## 2.5. Python

Python programming language is the best programming language for the field of Machine Learning and Deep Learning because Python has several advantages, namely, ease of writing syntax, large community support, having a complete library and also being popular. Using python and using its libraries will increase the productivity of machine learning and deep learning developers. And for beginners who want to get into the world of Machine Learning and Deep Learning, don't worry because there are so many learning resources provided, especially on the internet [4].

## 2.6. Dart

Dart is a multi-functional programming language that is ideal for scripting, programming, and markup languages. Dart can be used for different types of application development.

## 2.7. TensorFlow

Tensorflow was developed by researchers from the Google team and is one of the deep learning platforms and libraries for data science that is open source. This software framework was created to make the application of machine learning and deep learning concepts easier, TensorFlow deals with computational algebra optimization techniques to make the calculation of mathematical expressions easier.

## 2.8. Mediapipe

Mediapipe is an open source framework for building a multimodal application or in the form of video, images, audio, and other data. This framework was created by Google.Inc with the aim of making it easier for developers to build an application in the field of computer vision such as artificial intelligence and image processing. The framework can be run on Android, iOS, JavaScript, and Python platforms.

## 2.9. Ngrok

Ngrok is an application that opens a local web server to the internet and is a proxy server with a tunneling method. Ngrok works by bypassing the local server's web access to the internet via a proxy server and being given a connection from the Ngrok cloud service so that it can be accessed from anywhere, Ngrok creates a public HTTPS URL that forwards traffic to the localhost 5000 model accessible Flutter without a production server.

# 3. Analysis and Design

## 3.1. Research Methodology

The research method used is **experimental research with a Research and Development (R&D) approach**. The research stage began with problem identification and collection of supporting theories, followed by the collection of 5,280 SIBI sign language image imagesets of 5,280 images in 28x28 pixels format which were divided into training data (80%) and validation data (20%). Furthermore, the system was designed using UML modeling and Flutter-based application interface design. The implementation is carried out by building a Convolutional Neural Network (CNN) model using Python and TensorFlow, then integrating it through the Flask REST API so that it can run in real-time on mobile devices. The final stage is testing, including black box testing for application functionality and model performance evaluation using accuracy, precision, recall, and mean average precision (mAP) metrics.

## 3.2. System Requirement Analysis

a. Evaluation Method

Intersection over Union (IOU) is a measurement used to measure how much area overlaps between the predicted bounding box and the actual bounding box (ground-truth). IOU is calculated by dividing the area of overlap between the two bounding boxes by the combined area of the two. IOU is typically used to determine whether a detection is considered true or false in the evaluation of object detection algorithms.

Precision is the ability of a model to identify objects that are relevant or important only. Precision is calculated by dividing the number of true-positive object detections by the total number of positive detections performed by the model. Precision measures how accurate the model is in classifying objects as positive, i.e. objects that are relevant or important. The higher the precision value, the fewer objects are misclassified as positive.

Recall is the ability of a model to find all relevant or important objects. Recall is calculated by dividing the number of true-positive object detections by the total number of objects that are actually present in the image or video. Recall measures how well the model finds all relevant or important objects, regardless of whether the model classifies those objects correctly or incorrectly. The higher the recall value, the more objects the model manages to find. Mean Average Precision (mAP) is an evaluation metric to measure the accuracy of an object detection model across all classes of objects in a dataset. mAP is calculated by taking the average of the AP (Average Precision) value for each class of objects. AP is calculated by calculating the area under the Precision-Recall (P-R) curve on each object class. The higher the mAP value, the better the performance of the object detection model.

### 3.3. Model Implementation

Here are two flowchart flows where there is a training flowchart and a trial flowchart:

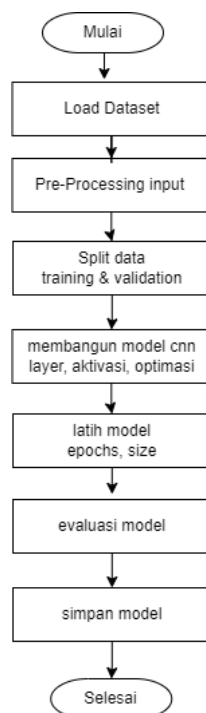


Fig. 1 : Flowchart Training

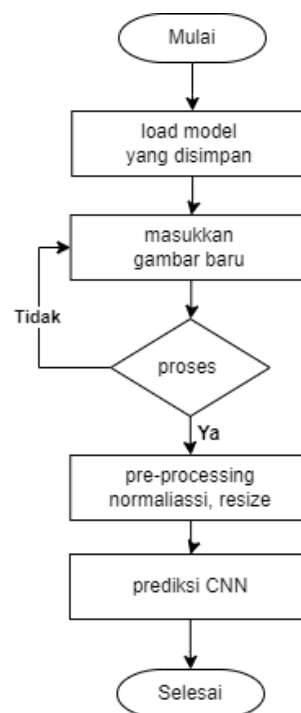


Fig. 2 : Trial Flowchart

**Fig. 1** The following is a description of the Flowchart training:

1. Load the letter image dataset first
2. Pixel normalization, augmentation (rotation, flip, and resize)
3. Divide it into training data (80%) and validation (20%)
4. Add a CNN layer, pooling
5. Train incoming models
6. Monitor accuracy and loss on data validation
7. Save the model for testing

**Fig. 2** The following is the description of the trial flowchart :

1. Pre-trained cnn model inputs
2. Insert a new letter image
3. The system will start to adjust the image format to the model (size, normalization), if it is correct then it will continue with classification, if not then the user will start to enter the image
4. The model starts to classify letters
5. Predicted results displayed

### 3.3.1. Research Supporting Data

Table 1: Data Inputan

Abjad	
A	N
B	O
C	P
D	Q
E	R
F	S
G	T
H	U
I	V
K	W
L	X
M	Y

The following data will be used as supporting data in the creation of applications for the Implementation of CNN on the Sign Language Detection and Recognition System to Support Communication for Persons with Disabilities.

### 3.4. Design

The design made in this application is made as attractive as possible and the appearance of how to use it is very young to make it easier for users of this application.

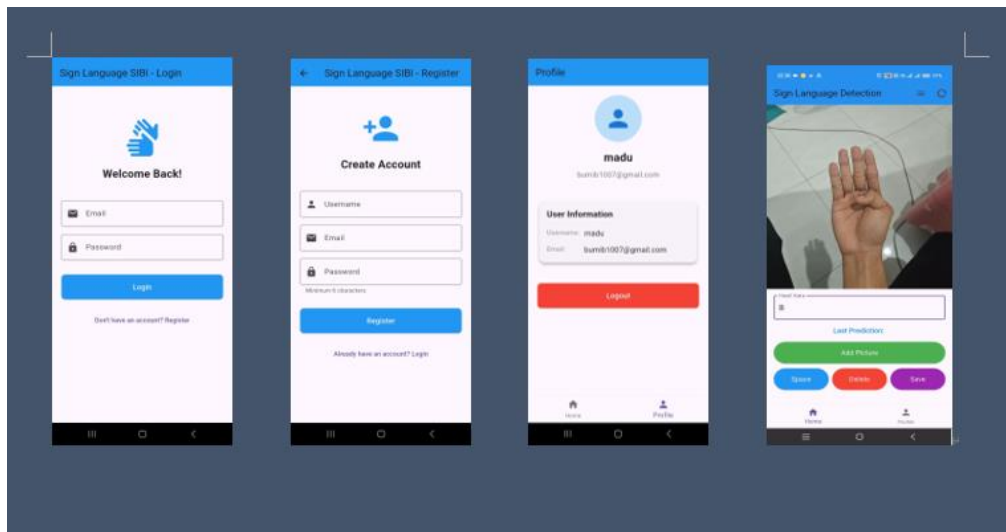


Fig. 3 : Application Design

## 4. Discussion and Implementation

Trial is a form of implementation of the SIBI system that is made with the aim of ensuring that the system is already in a ready-to-use condition. After the system has been analyzed and designed in detail, it will move to the implementation stage. System implementation is the stage of laying the system so that it is ready to operate. Testing using the Black Box testing method is a testing stage that focuses on the functional requirements of the software.

### 4.1. Evaluation Value in the System

Table 2: Evaluasi Pada Sistem

Epoch	Trainig loss	Trainng Accuracy	Validation loss	Validation accuracy
25	0.2964	91.02%	0.3658	91.25%
50	0.2431	92.84%	0.1875	93.80%
100	0.2115	94.41%	0.1732	95.91%
120	0.1989	95.02%	0.1654	96.34%

From the table above, it is explained that the data entered has epoch 120 which has increased considerably, so epoch 25, 50, 100, and 120 are used as epoch that affects the testing of data at the time of training.

## 4.2. Chart Visualization Results

Here is a graph of the training data with 120 epochs. The accuracy validation value obtained during the prediction process was 96.34%.

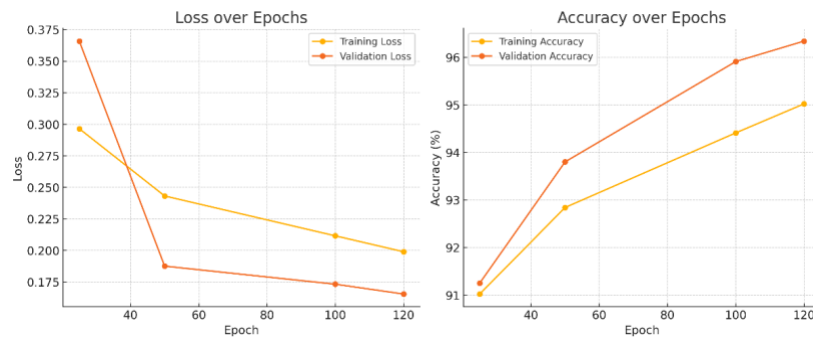


Fig. 4 : Training Results Graph

From the image above, it can be explained that the left graph shows a **decrease in losses** as the epoch increases, and the right graph shows that there is an **increase in accuracy**, both in training and validation consistently.

## 5. Conclusion and Recommendations

### 5.1. Conclusion

Based on the results of the research that has been carried out by the author, the author can provide several conclusions, including the following:

1. The CNN system is designed to run well without errors and overfitting, and it can be used in the real world by users (friends with disabilities and those who want to learn sign language).
2. Second, the system created can be well integrated in real time, which can be accessed anywhere and anytime using a smartphone with a mobile device.
3. Third, the system created can evaluate the performance results of each class that is entered into the system to be trained, starting from the accuracy, and responsiveness of the system in recognizing incoming images.

The system developed has gone through the process of testing the interface and black box. This system is able to run well on various devices such as laptops/PCs and Android smartphones, and has a user-friendly interface for image capture and storage. Testing shows all functions are running as planned without any significant problems. In addition, the model's performance evaluation showed excellent results with an accuracy of 96.34% and an mAP value above 0.87, indicating accurate and reliable letter detection capabilities. Overall, the system shows optimal performance and is suitable for image-based letter recognition.

### 5.2. Recommendations

After going through various processes in completing research on the Implementation of CNN on Sign Language Detection and Recognition System to Support Communication with Disabilities, the researcher realized the need for suggestions and inputs so that this research could be continued to the next researcher who is interested in conducting further research. Here are some related suggestions :

1. It is recommended to expand the dataset by adding image and movement variations to improve the accuracy of the system in recognizing sign language more robustly.
2. Ideally, in the future, system testing with various lighting conditions and backgrounds can be carried out to ensure the reliability of the system in various real-world situations.
3. It is recommended that the researcher further add features such as real-time letter recognition, automatic correction features, to increase the added value of the system.

## References

- [1] Christina, Maria, Karen Paruntu, H Anis, and Elko L Mamesah. 2023. XII NOMOR 8 TAHUN 2016 TENTANG PENYANDANG DISABILITAS DI INDONESIA 1. <https://www.un.org/development/desa/disabilities/conventi>.
- [2] Trisianto, Didik, and Michael Arthur Limantara. 2024a. 6 Surabaya Jurnal Sistem Cerdas dan Rekayasa (JSCR) SISTEM PEMBELAJARAN ISYARAT BAHASA INDONESIA (SIBI) MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK (CNN).
- [3] Abdillah, Gunawan, and Ridwan Ilyas. 2024b. 5 Deteksi Objek Bahasa Isyarat Huruf Bisindo Menggunakan SSD-MobileNet.
- [4] Riziq sirfatullah Alfarizi, M, Muhamad Zidan Al-farish, Muhamad Taufiqurrahman, Ginan Ardiansah, and Muhamad Elgar. 2023. 2 Karimah Tauhid PENGGUNAAN PYTHON SEBAGAI BAHASA PEMROGRAMAN UNTUK MACHINE LEARNING DAN DEEP LEARNING.
- [5] N. Z. Munantri, H. Sofyan, and M. Y. Florestiyanto, "Digital Image Processing Application for Tree Age Identification," *Ict*, vol. 16, no. 2, p. 97, 2020, doi: 10.31315/telematics.v16i2.3183.
- [6] Mr. Reza et al., "Artificial Intelligence: Image Processing & Application with Python," *Sem. Nas. Service. Mass. LPPM UMJPP*. 1–8, 2022, [Online]. Available: <http://jurnal.umj.ac.id/index.php/semnaskat>
- [7] Jinan, B. H. Hayadi, and U. P. Utama, "Classification of Rice Plant Diseases Using Convolutional Neural Network Method through Leaf Imagery (Multilayer Perceptron)," *J. Comput. Eng. Sci.*, vol. 1, no. 2, pp. 37–44, 2022.
- [8] N. Syahrudin and T. Kurniawan, "Inputs and Outputs in the Python Programming Language," *J. STMIK Python Programming Basics*, no. June 2018, pp. 1–7, 2018, [Online]. Available: <https://www.researchgate.net/publication/338385483>