Electronic Integrated Computer Algorithm Journal
Vol. 3, No. 1, October 2025, pp 24~27

ISSN: 3031-0350, DOI: 10.62123/enigma.v3i1.108

Implementation of Machine Learning For Indonesian Sign Language Recognition Using Convolutional Neural Network Model

Umi Salamah [1], Al-Khowarizmi[1*]

[1] *Department of Information Technology, Universitas Muhammadiyah Sumatera Utara*, Indonesia.

DOI : 10.62123/enigma.v3i1.108

Received : September 06, 2025
Revised : October 05, 2025
Accepted : October 12, 2025

Keywords:
Indonesian Sign Language, Machine Learning , CNN, *Gesture Recognition*

ABSTRACT

Sign language is the primary means of communication for people with hearing impairments. However, the public's limited understanding of Indonesian Sign Language (BISINDO) remains a communication barrier. This study implemented machine learning with a Convolutional Neural Network (CNN) model to automatically recognize BISINDO gestures. The dataset consists of 2,600 manually captured hand images representing the letters A–Z. The training process was carried out through data pre-processing, image augmentation, and CNN parameter optimization. Test results showed that the system was able to recognize BISINDO letters with high accuracy and could combine letters into simple words such as "HAI", "SAYA", and "UMI" in real-time. This study demonstrates that CNN is effective in supporting a computer-based sign language translation system, thus becoming an inclusive communication solution for people with hearing impairments.

1. INTRODUCTION

Artificial intelligence (AI) technology is developing rapidly and has had a significant impact on various fields [1], including industry, healthcare, education, and communications [2]. One of the most prominent branches of AI is machine learning [3][4], which allows computer systems to learn from data without having to be explicitly programmed. In the realm of image processing [5], machine learning is widely used through deep learning approaches [6]. One algorithm that has proven effective in recognizing complex visual patterns is the Convolutional Neural Network (CNN) [7][8][9].

Communication is a fundamental aspect of human life. For people with hearing impairments, sign language is the primary means of conveying messages, emotions, and information. In Indonesia, Indonesian Sign Language (BISINDO) [10][11] is a primary form of communication for the deaf community [12]. However, public understanding of BISINDO remains low, resulting in deaf people often experiencing communication barriers. The limited number of sign language interpreters also exacerbates the situation, leading to limited access to education, employment, and public services.

Numerous studies have demonstrated the application of machine learning technology to sign language recognition. While international research has largely focused on American Sign Language (ASL) [13][14], research related to BISINDO remains limited [15][16]. This opens up opportunities to develop a sign language recognition system specifically for the Indonesian context. This research aims to build a CNN model to recognize the BISINDO letters, test its performance on a hand image dataset, and implement the system into a web-based application that can be used in real time. This research aims to create a system that can help reduce the communication gap between deaf people and the general public.

2. METHODOLOGY

2.1 Research Design

This research uses a quantitative approach with an image-based observation method. Data were obtained through manual recording using a laptop camera. Each letter of the BISINDO alphabet is represented in 100 images [17], resulting in a total of 2,600 images as the research dataset.

2.2 Problem Identification and Literature Study

The research phase began with problem identification, namely communication barriers due to the limited understanding of BISINDO among the general public [18]. Next, a literature review was conducted on CNN, computer vision, and previous research related to sign language recognition.

2.3 Data Collection and Pre-Processing

*Corresponding Author Email: alkhowarizmi@umsu.ac.id

Data collection was performed by recording hand gestures using Python and OpenCV. After data collection, pre-processing was performed, including image conversion to grayscale, pixel value normalization to the 0–1 range, and data augmentation through rotation, flipping, and lighting variations. These steps aim to improve the model's generalizability to real-world conditions.

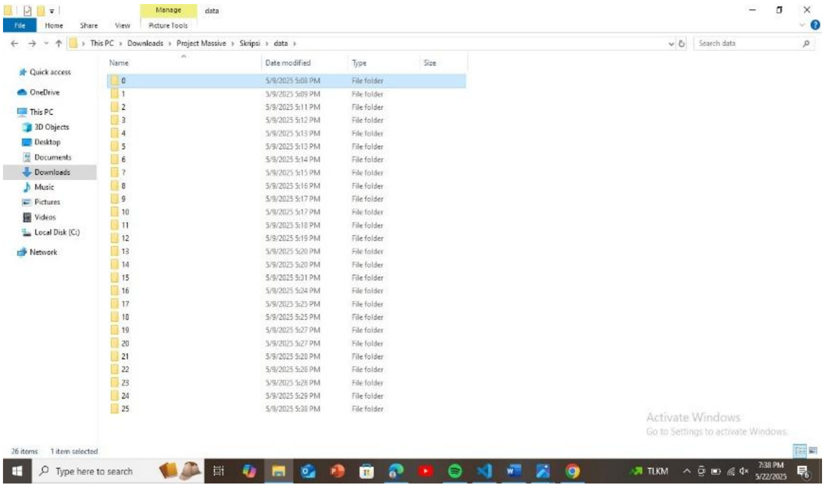

Figure 1. Research Dataset Structure Consisting of 26 Folders

The dataset is organized into 26 folders, each representing a letter of the BISINDO alphabet from A to Z. Each folder contains 100 images captured manually using a laptop camera, bringing the total dataset to 2,600 images. Variations in lighting conditions, shooting angles, and hand positions were included to enrich the training data and improve the model's generalization capabilities. The well-organized dataset structure facilitates pre-processing, training data sharing, validation, and testing, while ensuring that each letter has sufficient representation to build an accurate and stable CNN model

.

2.4 Detection Results

The designed CNN model consists of three convolutional layers with a 3×3 kernel and a ReLU activation function [19], followed by a 2×2 max pooling layer, and two fully connected layers with 128 and 64 neurons, respectively. The output layer uses a Softmax activation function, which produces 26 classes, corresponding to the number of letters in the BISINDO alphabet.

2.5 Model Training and Evaluation

The model training process was carried out by dividing the dataset into 80% training data, 10% validation data, and 10% test data. The optimizer used was Adam with a learning rate of 0.001, while the loss function used was categorical cross-entropy. Model evaluation was carried out using accuracy, precision, recall, and F1-score metrics.

2.6 System Implementation

In the final stage, the model was integrated into a web-based application using Streamlit. MediaPipe was used to detect hand landmarks in real-time, allowing users to perform BISINDO letter gestures in front of the camera, and the system automatically displayed the predicted results in text.

3. RESULTS AND DISCUSSION

3.1 Data Collection Results

The dataset consists of 2,600 images of the letters A–Z BISINDO. Image variations include different backgrounds, varying lighting, and varying hand positions. These variations aim to improve the model's generalizability to real-world conditions.

3.2 Model Training Results

Model training demonstrated that the CNN achieved over 90% accuracy on the test data. Training and validation curves demonstrated stable model performance without significant overfitting. This demonstrates that data augmentation techniques successfully improved the model's generalization capabilities.

3.3 System Implementation

The web-based system was successfully implemented. The application interface allows users to directly activate the camera, display real-time imagery, and recognize BISINDO gestures. The system can recognize letters with a response time of less than two seconds. The recognized letters can then be combined to form simple words such as "HAI," "SAYA," and "UMI".

3.4 Realtime Test

Real-time testing demonstrated that the system performed quite stably, despite limitations in dim lighting conditions or busy backgrounds. This system has proven effective as a first step in developing sign language translation technology in Indonesia. The next challenge is to improve accuracy in more diverse environments and expand the system's capabilities to recognize more complex sentence sequences.

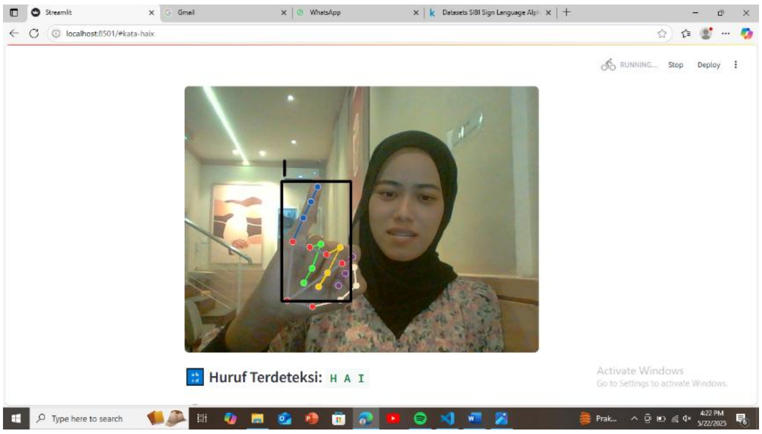

Figure 2. Letter Detection Process in the System

MediaPipe is used to detect hand landmarks, which are displayed as colored dots connected according to the anatomical structure of the hand. After these dots are extracted, a CNN performs classification to determine the letter that corresponds to the gesture. In this example, the system successfully recognized the letters "H," "A," and "I," which were then arranged into the word "HAI." These results demonstrate the system's ability to process visual data quickly, with a detection time of approximately two seconds for each letter. The CNN architecture bolsters the recognition accuracy, enabling the system to function practically as a communication tool.

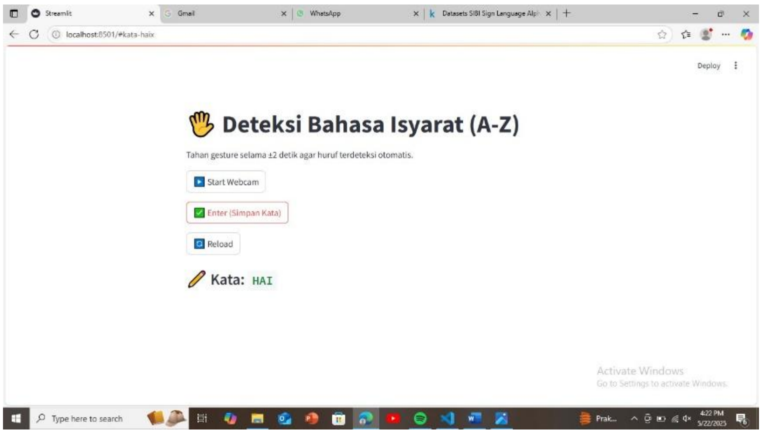

Figure 3. Display of the web-based BISINDO detection application interface

On the main page, users can activate the webcam by pressing the Start Webcam button, then perform a hand gesture that will be processed by the system. The Enter (Save Word) button functions to save the detected letters and arrange them into a word. Meanwhile, the Reload button is used to clear the previous detection results to allow for new input. At the bottom, the system displays the resulting word formed from the sequence of detected letters. This display demonstrates the successful integration between the CNN model and the web-based user interface, allowing the system to be used interactively without the need for additional applications.

4. CONCLUSSION

This research successfully implemented a Convolutional Neural Network (CNN) model for recognizing Indonesian Sign Language (BISINDO) letters. The model achieved over 90% accuracy on test data and accurately recognized letters. The web-based system implementation enabled real-time letter recognition and simple word construction, thus supporting basic communication for people with hearing impairments.

The results of this study demonstrate the effectiveness of CNN for image-based sign language recognition. For further research, it is recommended to expand the dataset with varying lighting and users, develop the system to construct sentences, and add text-to-speech features for more inclusive communication.

REFERENCES

[1] Y. Xu *et al.*, "Artificial intelligence: A powerful paradigm for scientific research," *The Innovation*, vol. 2, no. 4, Nov. 2021, doi: 10.1016/j.xinn.2021.100179.

[2] R. Raj and A. Kos, "Artificial intelligence: Evolution, developments, applications, and future scope," *Przegląd Elektrotechniczny*, vol. 99, 2023.

[3] I. S. O. An, "AI & Machine Learning," 2020.

[4] H. Nozari, J. Ghahremani-Nahr, and A. Szmelter-Jarosz, "AI and machine learning for real-world problems," in *Advances in computers*, vol. 134, Elsevier, 2024, pp. 1–12.

[5] B. J. Erickson, "Basic artificial intelligence techniques: machine learning and deep learning," *Radiologic Clinics*, vol. 59, no. 6, pp. 933–940, 2021.

[6] I. Papastratis, C. Chatzikonstantinou, D. Konstantinidis, K. Dimitropoulos, and P. Daras, "Artificial intelligence technologies for sign language," *Sensors*, vol. 21, no. 17, p. 5843, 2021.

[7] I. M. Pramono, Z. Niswati, and A. Agustina, "MODEL PENERJEMAH BAHASA ISYARAT INDONESIA DENGAN METODE CONVOLUTIONAL NEURAL NETWORK (CNN)," in *Semnas Ristek (Seminar Nasional Riset dan Inovasi Teknologi)*, 2024.

[8] M. B. Septyono, F. T. Anggraeny, and R. Mumpuni, "Pengenalan Ekspresi Wajah dengan LBP dan Multi-Level CNN," *Jurnal Informatika Polinema*, vol. 11, no. 4, pp. 469–476, 2025.

[9] I. I. Arifah, F. N. Fajri, and G. Q. O. Pratamasunu, "Deteksi Tangan Otomatis Pada Video Percakapan Bahasa Isyarat Indonesia Menggunakan Metode YOLO Dan CNN," *Journal of Applied Informatics and Computing*, vol. 6, no. 2, pp. 171–176, 2022.

[10] S. A. Sanjaya and H. F. Ilone, "BISINDO Sign Language Recognition: A Systematic Literature Review of Deep Learning Techniques for Image Processing," *The Indonesian Journal of Computer Science*, vol. 12, no. 6, 2023.

[11] R. A. Gani and T. A. B. Wirayuda, "Recognizing Indonesian Sign Language (BISINDO) Alphabet Using Optimized Deep Learning," in *2025 International Conference on Advancement in Data Science, E-learning and Information System (ICADEIS)*, IEEE, 2025, pp. 1–8.

[12] A. Nugroho, R. Setiawan, and A. Harris, "Deteksi Bahasa Isyarat Bisindo Menggunakan Metode Machine Learning," *Jurnal PROCESSOR*, vol. 18, no. 2, 2023.

[13] T.-W. Chong and B.-G. Lee, "American sign language recognition using leap motion controller with machine learning approach," *Sensors*, vol. 18, no. 10, p. 3554, 2018.

[14] B. Alsharif, A. S. Altaher, A. Altaher, M. Ilyas, and E. Alalwany, "Deep learning technology to recognize american sign language alphabet," *Sensors*, vol. 23, no. 18, p. 7970, 2023.

[15] S. Nur Budiman, S. Lestanti, and H. Yuana, "SIBI (Sistem Bahasa Isyarat Indonesia) berbasis Machine Learning dan Computer Vision untuk Membantu Komunikasi Tuna Rungu dan Tuna Wicara," *Jurnal Teknologi dan Manajemen Informatika*, 2023.

[16] I. Amri, "Implementasi Algoritma Convolutional Neural Network Untuk Menerjemahkan Bahasa Isyarat," *Jurnal Multidisiplin Saintek Volume*, vol. 2, no. 9, pp. 70–87, 2024.

[17] R. A. Gani and T. A. B. Wirayuda, "Recognizing Indonesian Sign Language (BISINDO) Alphabet Using Optimized Deep Learning," in *2025 International Conference on Advancement in Data Science, E-learning and Information System (ICADEIS)*, IEEE, 2025, pp. 1–8.

[18] L. Arisandi and B. Satya, "Sistem klarifikasi bahasa isyarat indonesia (Bisindo) dengan menggunakan algoritma convolutional neural network," *Jurnal Sistem Cerdas*, vol. 5, no. 3, pp. 135–146, 2022.

[19] Y. Jiang, J. Xie, and D. Zhang, "An adaptive offset activation function for CNN image classification tasks," *Electronics (Basel)*, vol. 11, no. 22, p. 3799, 2022.