p-ISSN 2715-2871
e-ISSN 2714-5247

Tensorflow Based AI Training for Translating Indonesian Sign Language (BISINDO)

Akbar Sujiwa
Department of Physics
Universitas Pembangunan Nasional "Veteran" Jawa Timur
Surabaya, Indonesia
akbarsujiwa.ft@upnjatim.ac.id

Nailul Hasan
Department of Physics
Universitas Pembangunan Nasional "Veteran" Jawa Timur
Surabaya, Indonesia
nailul.hasan.fisika@upnjatim.ac.id

Nenni Mona Aruan
Department of Physics
Universitas Pembangunan Nasional "Veteran" Jawa Timur
Surabaya, Indonesia
nenni.aruan.ft@upnjatim.ac.id

Abstract — **Communication is an important aspect of social life, including for deaf people who rely on sign language as their primary means of interaction. However, the limited understanding of sign language in society, especially BISINDO (Bahasa Isyarat Indonesia), is often a barrier to effective communication. This study aims to develop a simple sign language translator system using TensorFlow and Keras-based machine learning technology, as well as image processing support from OpenCV. The system is trained to recognize six basic signs in BISINDO, namely "Halo", "Nama", "Saya, "A", "D", and "I", using hand images as training data. Each word is represented by 20 images with variations in horizontal position. The test results show that the system is able to recognize signs with fairly good accuracy under certain conditions, but still has difficulty when faced with variations in hand positions that are different from the training data. This study shows the potential of sign language recognition technology in supporting inclusive communication, as well as the importance of data enrichment and variation in viewpoints in the model training process so that the system can function more optimally in various real conditions.**

Keywords—BISINDO, Sign Language, TensorFlow, AI, Translating

I. INTRODUCTION

Communication is a fundamental aspect of human life, and sign language is the primary means for deaf people to interact with their surroundings. Deaf people face unique challenges in communicating in various aspects of life, from social interactions to work and education environments [1]. However, the limited understanding of Indonesian sign language by the general public is still a barrier to realizing inclusive communication. Sign language is a unique language because it differs in each country. In Indonesia, there are two categories of sign language development, namely SIBI (Indonesian Sign Language System) and BISINDO (Indonesian Sign Language). This is a sign language that can help deaf people communicate [2]. Indonesian Sign Language (BISINDO) was created naturally by deaf people to communicate what deaf people want, feel, know, think, and do in everyday life [3]. Therefore, a technological solution is needed that is able to bridge the communication gap between deaf people and the wider community.

The development of artificial intelligence (AI), especially in the fields of image recognition and natural language processing, has opened up great opportunities in the development of automatic sign language translation systems [4]. One of the frameworks widely used in the development of AI models is TensorFlow, an open-source software library developed by Google for numerical computing, which is now widely used by many large companies. Tensorflow provides an interface for expressing machine learning algorithms and applications for executing these algorithms [5], [6]. TensorFlow is also very powerful for building and training neural networks (such as CNN, RNN, LSTM). Supports automatic differentiation for backpropagation. TensorFlow is now tightly integrated with Keras as a high-level API, which makes the process of training and evaluating models easier and more intuitive [7], [8]. An open-source library that supports efficient training and deployment of deep learning models. Using TensorFlow, AI models can be trained to recognize hand gestures from images or videos, then translate them into text or voice [9][10], [11].

In recent years, various studies have been conducted to develop artificial intelligence (AI) systems capable of translating sign languages, especially English sign languages such as American Sign Language (ASL) [12], [13]. Researchers have utilized technologies such as computer vision, machine learning, and deep learning to build systems that can recognize hand gestures and facial expressions, then translate them into text or voice in real-time [14], [15][16]. The results of these studies indicate that AI has great potential in bridging communication between deaf people and the wider community. This innovation not only increases the accessibility of communication but also opens up new opportunities in the fields of education and inclusive public services. However, most of the development of this technology is still focused on international sign languages, especially in the context of English, while development for local sign languages such as Bahasa Isyarat Indonesia (BISINDO) is still very limited.

This study aims to develop and train a TensorFlow-based AI model that is able to translate Indonesian Sign Language in real-time [17]. The main focus of this study is the process of training the model using the BISINDO hand gesture dataset, as well as evaluating the model's performance in recognizing and translating gestures with high accuracy. It is hoped that the results of this study can be the first step in the development of sign language translator technology that can be implemented on mobile or wearable devices, so that it can facilitate daily interactions for the deaf.

II. METHODS

The research methods used in this study include data collection techniques, as well as the stages of development and testing of the designed system. This study uses a quantitative approach with software engineering methods, which aims to build and evaluate an artificial intelligence (AI) system in translating sign language. In addition, a model training process was also carried out using sign language image data, as well as testing the accuracy of the system in recognizing and translating hand gestures into text. Details regarding tools, materials, implementation stages, and analysis techniques will be explained systematically in the following sub-chapters.

A. Software Design

Software development in this study was carried out by utilizing the TensorFlow and Keras frameworks as the main tools in building machine learning models. TensorFlow is an open-source platform developed by Google and is widely used for numerical computation and the development of machine learning and deep learning models [18]. Keras itself is a high-level API that runs on TensorFlow and is designed to facilitate the process of creating and training neural network models with simpler and more intuitive syntax [19]. In addition, OpenCV (Open-Source Computer Vision Library) is also used to support the image processing process. The data used is a collection of photos taken directly by researchers while performing sign language movements, which are then processed into a digital image dataset. OpenCV is used to perform various stages of image pre-processing, such as color conversion, normalization, edge detection, and hand area segmentation, so that the data is more ready to be analyzed by machine learning models [20]. Each photo represents one letter or word in sign language and is labeled according to the movement shown. The model is then trained to recognize and differentiate these gestures, so that it can automatically translate sign language into text. This approach is expected to be a technological solution that supports inclusive communication for the deaf.

In recognizing sign language, the system is designed to detect hand patterns formed from hand movements or positions when performing a gesture. These patterns include finger shape, hand direction, wrist orientation, and relative position to the body. The system utilizes image processing techniques to extract visual features from input images, such as hand contours, object edges, and key points. With the help of libraries such as OpenCV, this process includes segmenting the hand from the background, converting colors, and normalizing the size and position of the image to be consistent. Once the hand features are identified, a pre-trained machine learning model will compare the pattern to known data to determine what gesture is displayed. The system's ability to accurately recognize hand patterns is key in the process of classifying sign language into text.

The system that is built utilizes hand landmarks as the main component in the sign language recognition process. Hand landmarks are key points on the hand that accurately represent the position and anatomical structure of the hand, such as fingertips, finger joints, and wrists. The system detects and generates a total of 21 landmark points on one hand, which are then connected by lines to form a hand skeleton. These points are obtained through image processing using libraries such as OpenCV, and can be strengthened with the help of hand pose detection models such as MediaPipe. Information from hand landmarks is then processed by the machine learning model to recognize the displayed gesture patterns [21].

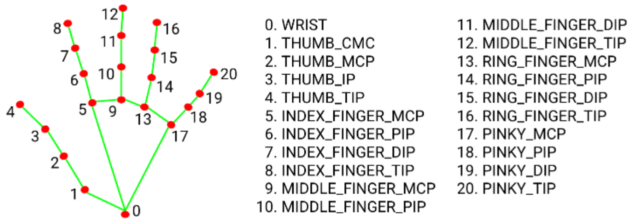

Figure 1. Hand Landmark [21]

B. Model Training

In training the system built, a number of sample images are needed that represent hand movements or articulations according to BISINDO signs. These images serve as training data that are the basis for the model to learn the visual patterns of each form of hand gesture. Each image is taken with adequate position and lighting to ensure good visual quality, so that important features of the hand can be accurately recognized by the system. The gestures used include certain letters and sentences in BISINDO, and each is recorded in the form of images repeatedly with various hand positions. This is done to increase data diversity and allow the machine learning model to recognize patterns more generally and robustly to differences in environmental conditions. In its application in this study, three words and three letters were chosen to represent everyday conversations. The three words taken as research objects are “Halo”, "Nama", and "Saya", when translate to english become “Hallo”, “Name” and “My”. The selection of these two words is based on the consideration that both are part of the basic expressions in self-introduction in everyday communication. While the letters are taken as a representation of the name of the sign language user, so for testing the name with a fairly short number of letters was chosen, namely "A", "D", and "I". For each sign word, 20 data images were taken, so that each word or letter in the sign language taken for model training could total 120 image data.

Figure 2 and Figure 3 show examples of training data used in the system training process, which show hand patterns when performing the BISINDO gesture. The images were taken directly from the data collection process using a camera, then processed with the help of OpenCV to detect and extract hand features. It can be seen that the system reads hand patterns based on points (keypoints) and connecting lines that represent the position of finger joints and hand joints. These points describe important parts of the hand, such as the fingertips, joints, and base of the palm,

p-ISSN 2715-2871
e-ISSN 2714-5247

while the lines between them form the hand skeleton that the system uses to recognize the shape and direction of the gesture. This pattern is then used as the basis for gesture classification, where each combination of the position and orientation of these points is associated with a certain word label. This approach allows the system to understand the structure of hand movements in more detail and precisely.

Figure 2. Training data from the word "Halo"

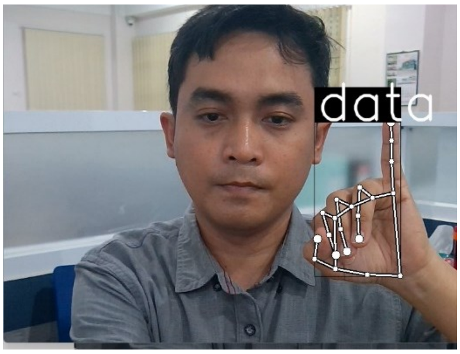

Figure 3. Training data from the letter "I"

C. System Accuracy Testing

System accuracy is tested directly with user hand movements. In this stage, the user performs sign language movements in real-time in front of the camera, and the system is asked to recognize and display the classification results. This test aims to observe whether the output produced by the system matches the gesture movements performed. If the system displays the correct label according to the gesture given, then the recognition is considered successful. Conversely, if the results do not match, then it is recorded as a classification error. Through this observation, the level of accuracy of the system in real conditions can be determined, as well as how stable the system is in recognizing hand gestures from various positions and individual variations. The results of this test also serve as the basis for evaluating and improving the system to make it more reliable and responsive in its use.

III. Results and Discussion

After testing, the results showed that the system was able to recognize hand gestures for the word and the letter signs with a fairly good level of accuracy. Most of the input given was successfully classified correctly by the system, according to the previously trained gesture labels. However, there were several cases where the system experienced errors in recognizing gestures, especially when the hand position was incorrect and when only one hand moves even though two hands are needed to translate, the system immediately translates, so this could be a translation error. Overall, the system showed adequate performance in recognizing simple sign language, indicating that the machine learning-based approach with OpenCV and TensorFlow support has the potential to be further developed. These results are the basis for evaluating and improving the system to be able to handle more variations of gestures and more complex usage conditions.

The results of hand movement recognition can be seen in the results Figure 4-9. It can be observed that the system has been able to correctly translate signals using either one hand or two hands.

Figure 4. System translates the word “Halo”

Figure 5. System translates the word “Nama”

Figure 6. System translates the word “Saya”

p-ISSN 2715-2871
e-ISSN 2714-5247

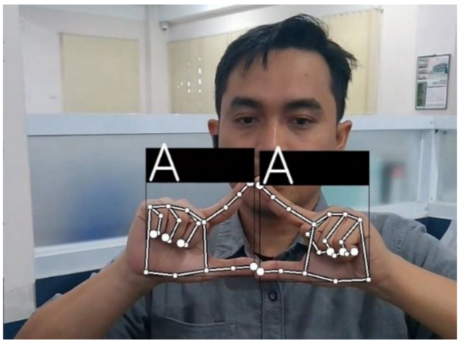

Figure 7. System translates the letter "A"

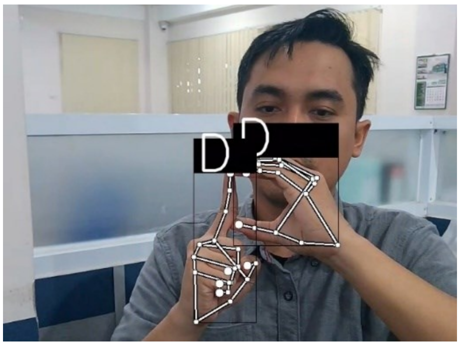

Figure 8. System translates the letter "D"

Figure 9. System translates the letter "I"

However, the system that was built still has several shortcomings that need to be considered. In a number of tests, it was found that the system failed to translate certain gestures, one of which occurred when the user showed the word "Saya" with a vertical hand position. The system was unable to recognize the gesture correctly because during the training process, the data used only included hand positions in a horizontal orientation. This causes the model to not have sufficient generalization ability to recognize variations in positions that are different from those previously learned. This failure shows that in order for the system to recognize sign language more naturally and accurately in various conditions, it is necessary to add variations in training data, especially in terms of angle and orientation. By training the model using more diverse data, it is hoped that the system can adapt to the way users sign in real situations, so that the translation results become more reliable and inclusive.

IV. CONCLUSION

This research has successfully developed a simple BISINDO sign language translator system using TensorFlow and Keras-based machine learning technology, with image processing support from OpenCV. The system is able to recognize six basic gestures, namely "Halo", "Nama", "Saya, "A", "D", and "I", with a fairly good level of accuracy under certain conditions. Testing shows that the system works effectively when the hand position matches the trained data, but still has difficulty recognizing variations in hand positions that have never been recognized before, such as vertical orientation. This shows the importance of diverse training data in building a more responsive and natural system. Overall, this system has the potential to be further developed to support inclusive communication for the deaf, especially if the number of vocabularies, movement variations, and accuracy are added through further training with more complete and representative data.

REFERENCES

[1] Jecho Gustian Althanio, Rasianna BR. Saragih, and Yuliati, "Komunikasi Interpersonal Pada Penyandang Tunarungu Dalam Interaksi Sosial di Kota Bengkulu," *JURNAL KAGANGA*, vol. 8, no. 2, pp. 126–135, Oct. 2024.

[2] A. S. Nugraheni, A. P. Husain, and H. Unayah, "OPTIMALISASI PENGGUNAAN BAHASA ISYARAT DENGAN SIBI DAN BISINDO PADA MAHASISWA DIFABEL TUNARUNGU DI PRODI PGMI UIN SUNAN KALIJAGA," *Jurnal Holistika*, vol. 5, no. 1, pp. 28–33, Feb. 2023, doi: 10.24853/holistika.5.1.28-33.

[3] Nisria, Mustafa, and Hadis, "IMPLEMENTASI BISINDO DALAM BERKOMUNIKASI PADA SESAMA ANAK TUNARUNGU," *PINISI JOURNAL OF EDUCATION*, vol. 2, no. 5, pp. 1–10, Sep. 2022.

[4] J. L. J. Boobala A., C. Charan Kesava Reddy, C. A. Reddy, and C. Bala Venkata Sai Rohith, "Real-Time Sign Language and Audio Conversion Using AI," *2024 International Conference on Communication, Control, and Intelligent Systems (CCIS)*, pp. 1–6, Dec. 2024, doi: 10.1109/CCIS63231.2024.10932061.

[5] M. Abadi *et al.*, "TensorFlow: A system for large-scale machine learning," *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation, OSDI 2016*, pp. 265–283, May 2016, Accessed: May 11, 2025. [Online]. Available: https://arxiv.org/pdf/1605.08695

[6] F. Ertam, "Data classification with deep learning using tensorflow," *2nd International Conference on Computer Science and Engineering, UBMK 2017*, pp. 755–758, Oct. 2017, doi: 10.1109/UBMK.2017.8093521.

[7] M. Amsaprabhaa, H. S. Sree, Jayashre, K. Muthamizhvalavan, N. Gummaraju, and S.

Padmajaa, "American Sign Language Real Time Detection Using TensorFlow and Keras in Python," *2024 3rd International Conference for Innovation in Technology, INOCON 2024*, 2024, doi: 10.1109/INOCON60754.2024.10511469.

[8] D. H. Noronha, K. Gibson, B. Salehpour, and S. J. E. Wilton, "LeFlow: Automatic Compilation of TensorFlow Machine Learning Applications to FPGAS," *Proceedings - 2018 International Conference on Field-Programmable Technology, FPT 2018*, pp. 396–399, Dec. 2018, doi: 10.1109/FPT.2018.00082.

[9] A. O. Hashi, S. Z. M. Hashim, and A. B. Asamah, "A Systematic Review of Hand Gesture Recognition: An Update From 2018 to 2024," *IEEE Access*, 2024, doi: 10.1109/ACCESS.2024.3421992.

[10] E. Krsak and T. Kello, "Improving teoretic train driving time with AI and TensorFlow," *ISIA 2020 - Proceedings, 4th International Symposium on Informatics and its Applications*, Dec. 2020, doi: 10.1109/ISIA51297.2020.9416531.

[11] W. F. Bastari, A. Sujiwa, and R. Setyobudi, "Penerapan Internet Of Things Pada Aplikasi Alat Deteksi Dan Monitoring Tekanan Darah," *Seminar Nasional Hasil Riset dan Pengabdian*, vol. 5, pp. 609–621, Aug. 2023, Accessed: May 16, 2025. [Online]. Available: https://snhrp.unipasby.ac.id/prosiding/index.php/snhrp/article/view/601

[12] A. S. Karche, A. V. Kamble, K. A. Maru, S. S. Kedari, and D. D. Sarpate, "American Sign Language Recognition Application," *2025 International Conference on Emerging Smart Computing and Informatics (ESCI)*, pp. 1–6, Mar. 2025, doi: 10.1109/ESCI63694.2025.10988218.

[13] P. Venkadesh, G. Jelin Taric, P. Mary mariyal, S. V. Divya, S. Danumithra, and K. Anuranjani, "SignChatAI-Generative AI for Deaf-and-Mute Community using Indian Sign Language(ISL)," *International Conference on Computing and Intelligent Reality Technologies, Proceedings of ICCIRT 2024*, pp. 84–88, 2024, doi: 10.1109/ICCIRT59484.2024.10921817.

[14] G. Serrano and D. Kwak, "Real-time Sign Language Recognition Using Computer Vision and AI," *Proceedings - 2023 International Conference on Computational Science and Computational Intelligence, CSCI 2023*, pp. 1214–1220, 2023, doi: 10.1109/CSCI62032.2023.00198.

[15] I. S. M. Dissanayake, P. J. Wickramanayake, M. A. S. Mudunkotuwa, and P. W. N. Fernando, "Utalk: Sri Lankan sign language converter mobile app using image processing and machine learning," *ICAC 2020 - 2nd International Conference on Advancements in Computing, Proceedings*, pp. 31–36, Dec. 2020, doi: 10.1109/ICAC51239.2020.9357300.

[16] A. Sujiwa and R. R. Dianto, "INFUSION MONITORING SYSTEM FOR PATIENTS BASED ON THE INTERNET OF THINGS (IOT) WITH ANDROID NOTIFICATION SYSTEM," *BEST : Journal of Applied Electrical, Science, & Technology*, vol. 4, no. 2, pp. 41–46, Nov. 2022, doi: 10.36456/BEST.VOL4.NO2.6172.

[17] R. R. Milinda, Sara, B. Ramsiej, R. Yasmin, V. D. Ambeth Kumar, and K. Sekar, "Sign Language Detection in Real-Time Applications," *International Conference on Computing and Intelligent Reality Technologies, Proceedings of ICCIRT 2024*, pp. 1–4, 2024, doi: 10.1109/ICCIRT59484.2024.10921810.

[18] C. Mattmann, *Machine Learning with TensorFlow, Second Edition*. Manning, 2020. [Online]. Available: http://ieeexplore.ieee.org/document/10280532

[19] D. Pietro Pau, T. A. Naramo, and M. D. Randriatsimiovalaza, "Coding Mel Spectrogram using Keras and Tensorflow for Home Appliances Tiny Classification," *Digest of Technical Papers - IEEE International Conference on Consumer Electronics*, vol. 2023-January, 2023, doi: 10.1109/ICCE56470.2023.10043378.

[20] A. Kumari Sirivarshitha, K. Sravani, K. S. Priya, and V. Bhavani, "An approach for Face Detection and Face Recognition using OpenCV and Face Recognition Libraries in Python," *2023 9th International Conference on Advanced Computing and Communication Systems, ICACCS 2023*, pp. 1274–1278, 2023, doi: 10.1109/ICACCS57279.2023.10113066.

[21] "GitHub - Kazuhito00/hand-gesture-recognition-using-mediapipe: MediaPipe(Python版)を用いて手の姿勢推定を行い、検出したキーポイントを用いて、簡易なMLPでハンドサインとフィンガージェスチャーを認識するサンプルプログラムです。（Estimate hand pose using MediaPipe(Python version). This is a sample program that recognizes hand signs and finger gestures with a simple MLP using the detected key points.）." Accessed: May 19, 2025. [Online]. Available: https://github.com/Kazuhito00/hand-gesture-recognition-using-mediapipe