Scientific Journal of Informatics
Vol. 11, No. 2, May 2024
p-ISSN 2407-7658 https://journal.unnes.ac.id/journals/sji/index e-ISSN 2460-0040

Sign Language Detection System Using YOLOv5 Algorithm to Promote Communication Equality People with Disabilities

Maylinna Rahayu Ningsih[1*], Yopi Julia Nurriski[2], Fathimah Az Zahra Sanjani[3], M. Faris Al Hakim[4], Jumanto Unjung[5], Much Aziz Muslim[6]
[1,2,3,4,5]Department of Computer Science, Faculty of Mathematics and Natural Sciences, Universitas Negeri Semarang, Indonesia
[6]Faculty of Computer Science and Information Technology, Universiti Tun Hussein onn Malaysia, Malaysia

Abstract.
Purpose: Communication is an important asset in human interaction, but not everyone has equal access to this key asset. Some of us have limitations such as hearing or speech impairments, which require a different communicative approach, namely sign language. These limitations often present accessibility gaps in various sectors, including education and employment, in line with Sustainable Development Goals (SDGs) numbers 4, 8, and 10. This research responds to these challenges by proposing a BISINDO sign language detection system using YOLOv5-NAS-S. The research aims to develop a sign language detection model that is accurate and fast, meets the communicative needs of people with disabilities, and supports the SDGs in reducing the accessibility gap.
Methods: The research adopted a transfer learning approach with YOLOv5-NAS-S using BISINDO sign language data against a background of data diversity. Data pre-processing involved Super-Gradients and Roboflow augmentation, while model training was conducted with the Trainer of SuperGradients.
Result: The results show that the model achieves a mAP of 97,2% and Recall of 99.6% which indicates a solid ability in separating sign language image classes. This model not only identifies sign language classes but can also predict complex conditions consistently.
Novelty: The YOLOv5-NAS-S algorithm shows significant advantages compared to previous studies. The success of this performance is expected to make a positive contribution to efforts to create a more inclusive society, in accordance with the Sustainable Development Goals (SDGs). Further development related to predictive and real-time integration, as well as investigation of possible practical applications in various industries, are some suggestions for further research.

Keywords: Detection, Sign language, Yolov5, Disability, SDGs
Received May 2024 / **Revised** June 2024 / **Accepted** June 2024

INTRODUCTION

An important element in human interaction is communication, which serves as an important bridge that connects different communities around the world. All people use language systems to communicate with each other [1] and everyone uses a different language depending on the language they live in. uses a different language [2] depending on the language they live in. However, it is worth remembering that when it comes to communication, not everyone has equal access to this key capital. Some of us have limitations such as hearing impairment or speech impairment, which requires a different communicative approach, i.e. sign language. According to the WHO, there are 360 million people worldwide with hearing disabilities, 9% of them children [3].

In such a situation, it is important to realize that this issue impacts not only the individual, but also the achievement of global goals [4], especially those related to the Sustainable Development Goals (SDGs). One of the relevant SDGs is SDG number 10 on Reduced Inequality [5] where people with disabilities often face inequalities in access to healthcare, education, and economic opportunities [6], which hinders the achievement of this goal. The SDGs particularly address this issue, estimated to account for 1.3 billion people or 16% of the global population [7], [8]. In addition, according to the ILO and WHO, the number of people with disabilities in Indonesia reached 41 million out of 275 million people in 2022 and this is a

*Corresponding author.
Email addresses: maylinnarahayuningsih@students.unnes.ac.id (Ningsih)
DOI: 10.15294/sji.v11i2.6007

large number [9]. Indonesia's Central Statistics Agency (BPS) also explained that the number of Indonesian workers with disabilities in 2022 increased by 160.18% from the previous year, and in this case the fact that people with disabilities are often faced with the challenge or reality that they often face discrimination, marginalization, and exclusion in various fields of life [10], including access and participation in communication. Even though referring to the rights of persons with disabilities, it has been regulated in Indonesia in article 5 of Law No. 8 of 2016. For this reason, to make society more inclusive and reduce the accessibility gap, concrete actions are certainly needed.

Reviewing another point of SDGs number 4 regarding quality education is closely related to this issue. Due to the lack of adequate means and support and support, people with disabilities often find it difficult to receive a high-quality education. high-quality education. In situations like this, greater efforts need to be made [11] to ensure that the education system can become more accessible and acceptable to all [12], including those with disabilities. Moreover, point 8 SDGs [13] on decent work and economic growth [14], is also relevant because people with disabilities often face discrimination in the workplace. Barriers include physical accessibility restrictions and lack of awareness of the potential of people with disabilities. Hence the need for inclusion in the workplace must be improved to ensure that everyone, regardless of physical ability, can participate actively and effectively in the workplace, regardless of their physical abilities, can actively participate in and obtain decent gain gainful employment.

In the context of the SDGs, creating inclusive and sustainable societies [15], [16] means bridging the gap in accessibility of communication, including paying special attention to the development of resources and technologies for the interpretation of sign language. Therefore, efforts are needed to improve the communication equality of people with disabilities as one way to support the SDGs. Languages Not only does sign language serve as an alternative way to communicate for people with disabilities, its complex linguistic system allows for interaction [17] and intense expression. Hand gestures, facial expressions and body expressions are an important part of sign language to convey meaning and form a rich and meaningful communication system. Sign language plays an important role in daily life for people with disabilities [18], as it is the main way they receive and transmit information. However, those who are able to interpret using sign language professionally proficiently are scarce, which hinders communication between people with disabilities and ordinary people. between people with disabilities and ordinary people.

Recently, there has been a lot of research on the development of systems that can categorize the gestures of various sign languages according to certain classes to develop sign language detection systems by identifying hand gestures, facial expressions, and body expressions [19], facial expressions, and body expressions of people with disabilities [20]. One of the multidisciplinary challenges that has yet to be fully solved is the automatic recognition of human gestures. Currently, there are several techniques that can be applied, including the use of machine learning techniques to identify sign language [21]. Sign language recognition is becoming increasingly interesting as Deep Learning (DL) techniques are developed.

In identifying hand gestures various models either deep learning, classification or conditional random fields and others have been widely performed [22],[23]. However, categorization of gestures from different subjects to be difficult to predict under changing lighting conditions is still a problem with many solutions being developed [24]. Viewing hand gestures and displaying the results is a natural way to create interfaces. By using a camera, the device can record this activity [25]. Deep Learning (DL) algorithms can identify hand gestures in recorded images. Many algorithms have been used to detect and recognize hand gestures, including [26] YOLOv5. including YOLOv5, which stands out for its ability to provide fast and accurate provide fast and accurate results. YOLOv5 is particularly helpful in addressing the issue of gesture categorization, even under changing lighting conditions or different subjects [27].

Relevant research related to the YOLOv5 algorithm includes improving multi-scale object recognition and target perception capabilities [28], using a YOLOv5-based updated broken egg detection model that combines BiFPN and CBAM. The model achieved an average accuracy of 92.4%, surpassing the original network and other models, and can detect egg processing paths in real-time, although the omission rate may increase with higher transport speeds. Using the YOLOv5 model, another study that used deep learning techniques to detect and count agricultural pests [29] utilized data augmentation and transfer learning, which achieved an 84% improvement in accuracy and an 84% improvement in transport speed. achieved an accuracy improvement of 84% and precision improvements of 15%, 18%, and 7%, respectively.

Relevant research on the use of YOLOv5, such as in [30], discusses the use of three deep learning-based models, including YOLOv5x, and attention methods to improve sign language recognition, with accuracy results of 98.9% on the MU HandImages ASL dataset and 97.6% on the kkhorNama: BdSL dataset, as well as a lightweight and fast model for real-time use. Another study by analyzing three different CNN architectures, [31] tried to develop an Indonesian Sign Language (BISINDO) hand gesture recognition system, with Densenet121 showing the best accuracy. In another study [32], a machine learning approach was used to develop a BISINDO speech translation system, which was successful with an accuracy value of 98% using the Support Vector Machine (SVM) method. However, when tested directly on the user, the accuracy of the model drastically decreased to 78% as it exceeded the system's effective range but still able to detect both static and dynamic gestures in real-time. and dynamic gestures in real-time

Although YOLOv5 has been used in previous research to improve sign language and hand gesture recognition, further research opportunities can be expanded by focusing on more predictive and real-time integration. In the meantime, our research tries to advance this contribution by exploring the application of live video or hand gestures via the camera. In our research, we propose and emphasize the advantages of the YOLOv5 algorithm with the YOLOv5-NAS-S architecture and transfer learning COCO in fast and accurate object detection using the learning COCO in fast and accurate object detection using the BISINDO sign language, having knowledge of the BISINDO sign language and has strong feature knowledge which has an impact on the improved sign language recognition.

METHODS

This In this research, an object detection approach is proposed to implement a real-time BISINDO sign language detection system. In general, the development flow of this sign language detection system includes four steps, namely data collection, data pre-processing, model training and evaluation, and system deployment, as illustrated in Figure 1.

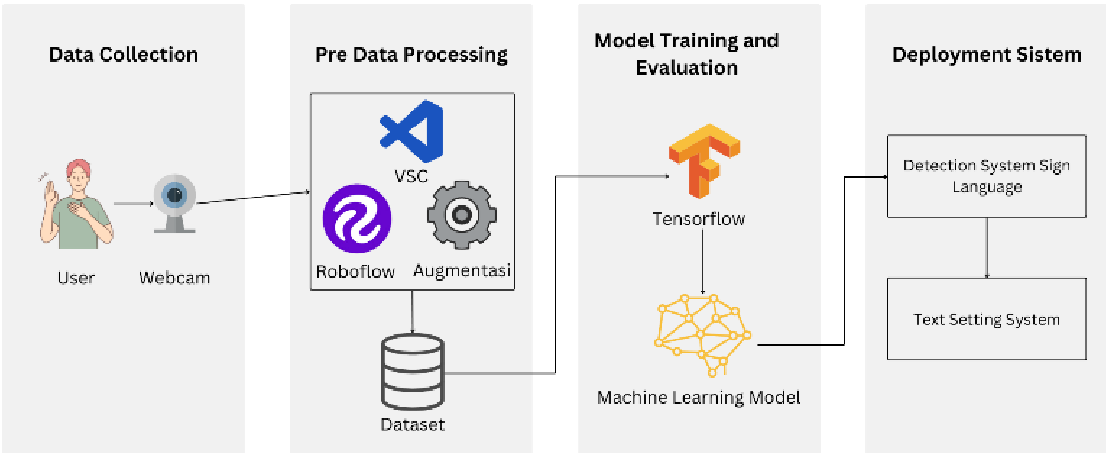

Figure 1. Research method

Model development

In the image data processing stage, Roboflow and Super-Gradients are used. In addition, the TensorFlow library and the YOLOv5 algorithm model as object detection are used in model creation and evaluation. The results obtained are expected to be in the form of sign language representations from images, videos, and real-time test recordings with text based on the probability of similarity. To achieve this goal, the development and evaluation of the model were carried out as illustrated in the flow diagram provided in Figure 2.

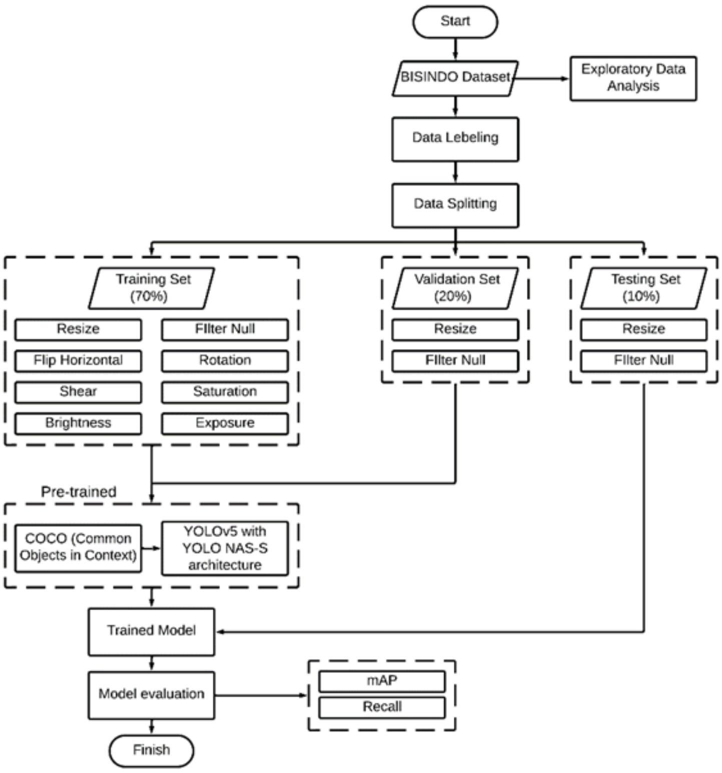

Figure 2. Diagram of model development

Dataset

The dataset used in this research is a type of BISINDO Sign Language. The data collection process is performed by taking data independently using a mobile phone camera in the form of 47 videos with a duration of 50 minutes with a resolution of 1280 x 720 pixels. Finalization of the dataset is done by capturing video into images using Visual Studio Code and Python library. The total number of images generated is 2388 datasets covering 47 classes involving letters of the BISINDO sign language alphabet and several types of words. Each dataset contains a collection of images taken under various environmental conditions with variations in lighting and background. Sign language BISINDO sign language with 47 classes implemented in this system can be seen in Table 1.

Table 1. BISINDO sign language

Alphabet Letter	Word
A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z.	SUKA, RUMAH, BAIK, BANTU, JANGAN, MAAF, KERJA, MINUM, APA, DIA, KEREN, SENANG, MARAH, BERMAIN, KAMU, KAPAN, AKU, AYAH, SEDIH, SABAR, KAKAK

Data pre-processing

After preparing the dataset, data pre-processing is done by entering the created dataset into Roboflow. Preprocessing starts with data labeling and null filtering to filter and eliminate datasets without samples, then resizing the image to 640 x 640 pixels to fit the input shape in the model. image to 640 x 640 pixels to fit the input shape in the model. Furthermore, the dataset is divided into 3 parts, namely, training dataset (70%), validation (20%), and testing (10%). (20%), and testing (10%).

Data augmentation was applied to the training dataset to increase dataset variation that can positively affect model performance. Augmentations performed include Horizontal flip, Rotation between -15° and +15°, Shear between ±15° Horizontal and ±15° Vertical, Saturation between -48% and +48%, Brightness between -20% and +20%, and Exposure between -9% and +9%. After augmentation of the dataset, the number of new datasets becomes 5518 BISINDO sign language datasets which are more diverse. An example of the augmented image can be seen in Figure 3.

Figure 3. Example of image augmentation

Modelling

In the model initialization step, the YOLO (You Only Look Once) object detection model was developed using the YOLO NASS (Neural Architecture Search-Small) architecture as the basis. The model initialization process is done through the models.get() function call, which retrieves the YOLO NAS-S model from an existing model library. The number of classes for object detection is determined by setting the num_classes parameter, which is adjusted to the number of classes in the dataset used. In the initialization stage, the model weights are initialized using pre-trained weights that have previously been trained on the COCO (Common Objects in Context) dataset. The use of pre-trained weights from COCO is done with the aim of improving the measurement method (Lin et al., 2014) in object detection tasks, considering that the COCO dataset has a wide coverage and a variety of objects.

Hyperparameter tuning

Hyperparameters are parameters used to control the model training process. In the table below, hyperparameters along with their respective values will be presented. This includes parameters such as silent_mode which determines whether log messages should be displayed during training, optimizer which specifies the optimization algorithm used, and max_epochs which determines the maximum number of training iterations. Each hyperparameter has predefined values, and their settings can affect the overall performance and behavior of the model. Table 2 will provide a clearer overview of these values and how they impact the training process.

Table 2. Hyperparameter configuration table

Parameter	**Concept Typical**	**Tuning Range**	**Explanation of Tuning Results**
silent_mode	Runs training in silent mode	True (usually not tuned)	Set to True to reduce console output noise
average_best_models	Averages weights of the best models	True (usually not tuned)	Set to True to improve model stability and performance
warmup_mode	Method for warmup	linear_epoch_step, constant	Tested linear_epoch_step for gradual increase in LR
warmup_initial_lr	Initial learning rate during warmup	1e-7 to 1e-4	Found 1e-6 to work well for smooth transition
lr_warmup_epochs	Number of epochs for learning rate warmup	1 to 10	3 epochs provided a good balance
initial_lr	Starting learning rate	1e-4, 5e-4, 1e-3	5e-4 selected for optimal learning speed
lr_mode	Schedule for adjusting learning rate	cosine, step, constant	cosine mode for gradual reduction of LR
cosine_final_lr_ratio	Final learning rate ratio in cosine annealing	0.01 to 0.1	0.1 ratio for effective LR decay
optimizer	Optimization algorithm	Adam, SGD, RMSprop	Adam chosen for efficient convergence
optimizer_params.weight_decay	Weight decay for the optimizer	0 to 0.01	0.0001 for regularization without over-penalizing
zero_weight_decay_on_bias_and_bn	Excludes biases and batch norm layers from weight decay	True (usually not tuned)	Set to True to prevent overregularization
ema	Uses exponential moving average of model weights	True (usually not tuned)	Enabled for stable weight updates
ema_params.decay	Decay rate for EMA	0.9 to 0.999	0.9 chosen for balanced EMA

ema_params.decay_type	Type of EMA decay	"threshold" (usually not tuned)	Threshold decay type used
max_epochs	Maximum number of training epochs	10 to 100	25 epochs for sufficient training without overfitting
mixed_precision	Uses mixed precision training	True (usually not tuned)	Enabled for better computational efficiency
loss.num_classes	Number of classes in the dataset	Depends on the dataset (usually not tuned)	Set according to the dataset's number of classes
loss.reg_max	Maximum regression value for the loss function	Specific to the task, e.g., 16	16 used based on task requirements
valid_metrics_list.score_thres	Score threshold for validation metrics	0.05 to 0.5	0.1 for balanced precision and recall.
valid_metrics_list.top_k_predictions	Top K predictions to consider during validation	100 to 300	300 for comprehensive validation
metric_to_watch	Primary metric for determining the best model	mAP@0.50, Precision@0.50, Recall@0.50, F1@0.50	mAP@0.50 chosen to monitor overall detection performance

Model evaluation

To ensure optimal quality and performance, model evaluation is conducted as an important stage in the model development process. Mean Average Precision (mAP) and Recall are the two-evaluation metrics used in this study. mAP measures the average precision accuracy of the model across multiple identified classes or objects, providing a complete picture of the extent to which the model can recognize and separate objects. The formula of mAP can be seen in Formula (1).

$$mAP = \frac{1}{N}\sum_{1=1}^{N} AP_1 \quad (1)$$

Then, Recall measures the extent to which the model is able to recognize and separate all true positive instances where the calculation formula can be seen in Formula (2).

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

RESULTS AND DISCUSSIONS

The results stage is based on the testing scheme described in the research method. At this stage, the proposed model will be tested using the YOLO-NAS-S architecture model of SuperGradients by determining the number of classes and using pre-trained weights from the COCO dataset. The model is trained using Trainer with data loaders on the training dataset and validation dataset. To see the architecture of the model architecture used, can be seen in Figure 4.

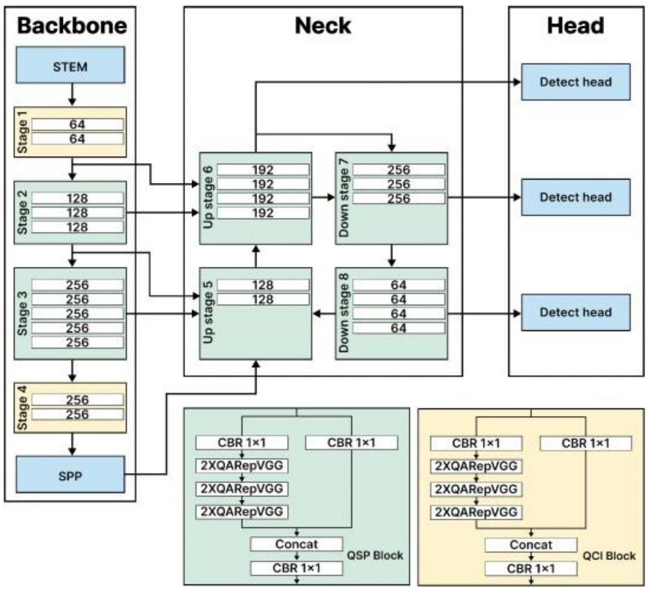

Figure 4. YOLO-NAS-S architecture [33]

Before discussing the results, it is crucial to understand the impact of preprocessing on the dataset. Initially, the dataset consisted of raw BISINDO sign language images with varying resolutions, orientations, and lighting conditions. After preprocessing with Roboflow, which included data labeling, null filtering, resizing images to 640 x 640 pixels, and augmentation, the dataset quality improved significantly. This preprocessing step was essential to ensure uniformity in the dataset, making it suitable for input into the YOLO-NAS-S model. The preprocessing made the dataset more varied in terms of lighting, viewing angles, and sizes, thereby enhancing the dataset's quality, which can lead to better model performance. The comparison of the dataset before and after preprocessing is shown in Figure 5 and Figure 6.

Figure 5. Dataset before preprocessing

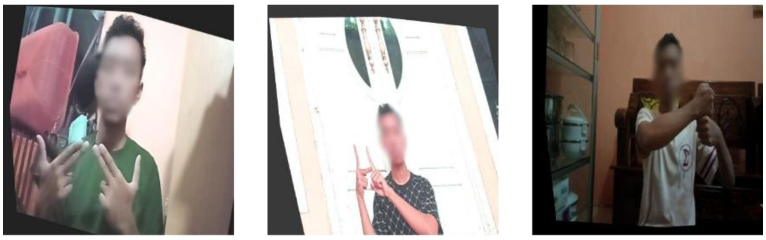

Figure 6. Dataset after preprocessing dataset after preprocessing

In addition, when comparing the model's prediction results to make it better, fine tuning is done by selecting the best model based on the performance on the validation dataset. Then, the model is tested on the data loader of the testing dataset, and the model is evaluated using the model evaluation using the object detection evaluation matrix. Model evaluation results using mAP and Recall before and after fine tuning can be seen in Table 3.

Table 3. Evaluation results

	Before Fine Tuning	**After Fine Tuning**
mAP	0,8698	0,9726
Recall	0,9963	0,9968

The results show the success of the model in separating sign language image classes very well, supported by the value of the model evaluation matrix mAP (mean Average Precision) and Recall. The mAP evaluation result of 0.9726 shows that the model can provide predictions with a high level of average precision accuracy. While the Recall value of 0.9968 confirms that the model is able to identify most or even all of the actual sign language objects. sign language objects. In seeing the results of images that are able to recognize image can be seen in Figure 7.

Figure 7. Example of model prediction results

This success gives confidence that the model is not only capable of recognize a class of sign language, but can also make consistent predictions on images, videos and real-time test recordings, even in the consistently on images, videos and real-time test recordings, even in situations that may be complex and full of variations. that may be complex and full of variations. The model demonstrated superiority in recognizing BISINDO sign language on images created with diversity of various contexts, such as different environmental conditions, different levels of lighting levels, and complex backgrounds. In support of Before Fine Tuning After Fine Tuning mAP 0.8698 0.9726 Recall 0.9963 0.9968 the success of the proposed model, a comparison with several existing methods which can be seen in Table 4.

Table 4. Model comparison

Author	Dataset	Method	Result
Attia et al., [30]	MU HandImages ASL dan OkkhorNama: BdSL	YOLOv5	98.9% on MU dataset HandImages ASL dataset and 97.6% on the BdSL dataset
Fauzi., [32]	Sign Language BISINDO	SVM	98%, when tested to 78%
Handhika et al., [33]	BISINDO Centre Dataset Interpreter Services Indonesian Sign Language Interpreter Service Centre, Jakarta	Generalized Learning Vector Quantization	94,3 %
Purpose Method	BISINDO Sign Language	YOLO-NAS-S+Fine Tuning	99,6%

CONCLUSION

In this research, an object detection approach is used to develop a real-time Indonesian Sign Language (BISINDO) detection system. The development process involves four main steps: Data Collection, Data Preprocessing, Model Training and Evaluation, and System Deployment. Firstly, researchers collected a comprehensive dataset of BISINDO sign movements, which were then preprocessed to enhance their quality and suitability for training. Tasks such as data cleaning, normalization, and augmentation were performed to ensure a diverse and representative dataset. The YOLOv5-NAS-S algorithm was then trained using the collected and preprocessed data. After model training, rigorous evaluation was conducted to assess performance metrics such as Average Precision (mAP) and Recall, achieving high values such as a reported mAP of 97.2% and Recall of 99.6%, demonstrating the effectiveness and reliability of the model. Furthermore, the developed system can be accessed via real-time webcam, showing the potential of object detection techniques to contribute to creating a more inclusive society, in line with the Sustainable Development Goals (SDGs).

REFERENCES

[1] M. Alsulaiman *et al.*, “Facilitating the communication with deaf people : Building a largest Saudi sign language dataset,” *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 35, no. 8, p. 101642, 2023, doi: 10.1016/j.jksuci.2023.101642.

[2] N. Hidayat and M. F. Al Hakim, “Halal Food Restaurant Classification Based on Restaurant Review in Indonesian Language Using Machine Learning,” *Sci. J. Informatics*, vol. 8, no. 2, pp. 314–319, 2021, doi: 10.15294/sji.v8i2.25356.

[3] M. Pellegrini *et al.*, “Changes in Weight and Nutritional Habits in Adults with Obesity during the ‘ Lockdown ’ Period Caused by the COVID-19 Virus Emergency,” pp. 1–11, 2020.

[4] C. Allen, S. Malekpour, and T. Bennich, “Recurring patterns of SDG interlinkages and how they can advance the 2030 Agenda,” pp. 1465–1476, 2023, doi: 10.1016/j.oneear.2023.10.008.

[5] F. Úbeda, F. Javier, E. Aracil, and A. Mendez, “How sustainable banking fosters the SDG 10 in weak institutional environments,” *J. Bus. Res.*, vol. 146, no. March, pp. 277–287, 2022, doi: 10.1016/j.jbusres.2022.03.065.

[6] V. Vuong and M. Palmer, “Love Thy Neighbour ? Social Attitudes Towards Persons With Disabilities,” *World Dev.*, vol. 174, no. November 2023, p. 106464, 2024, doi: 10.1016/j.worlddev.2023.106464.

[7] R. A. Baksh, S. E. Pape, J. Smith, A. Strydom, and J. Smith, “Understanding inequalities in 19 outcomes following hospital admission for people with intellectual disability compared to the general population : a matched cohort study in the UK,” 2021, doi: 10.1136/bmjopen-2021-052482.

[8] WHO, “Disability,” 2023.

[9] M. G. H. K. K., “Melibatkan Disabilitas,” 2023.

[10] T. Mladenov and C. S. Brennan, “The global COVID-19 Disability Rights Monitor : implementation , findings , disability studies response The global COVID-19 Disability Rights

Monitor : implementation , findings , disability studies,” *Disabil. Soc.*, vol. 36, no. 7–8, pp. 1356–1361, 2021, doi: 10.1080/09687599.2021.1920371.
[11] M. Antoninis, “SDG 4 baselines , midpoints and targets : Faraway , so close ?,” *Int. J. Educ. Dev.*, vol. 103, p. 102924, 2023, doi: 10.1016/j.ijedudev.2023.102924.
[12] F. Menashy and Z. Zakharia, “Partnerships for education in emergencies : The intersecting promises and challenges of SDG 4 and SDG 17,” *Int. J. Educ. Dev.*, vol. 103, no. November, p. 102934, 2023, doi: 10.1016/j.ijedudev.2023.102934.
[13] S. M. Rai, B. D. Brown, and K. N. Ruwanpura, “SDG 8 : Decent work and economic growth – A gendered analysis,” *World Dev.*, vol. 113, pp. 368–380, 2019, doi: 10.1016/j.worlddev.2018.09.006.
[14] S. Teresa and F. Su, “Sustainable Technology and Entrepreneurship Measuring business impacts on the SDGs : a systematic literature review,” vol. 2, no. July, 2023, doi: 10.1016/j.stae.2023.100044.
[15] A. Stauropoulou, D. Ph, E. Sardianou, G. Malindretos, K. Evangelinos, and I. Nikolaou, “World Development Sustainability The effects of economic , environmentally and socially related SDGs strategies of banking institutions on their customers ’ behavior,” *World Dev. Sustain.*, vol. 2, no. December 2022, p. 100051, 2023, doi: 10.1016/j.wds.2023.100051.
[16] M. R. Ningsih, K. A. H. Wibowo, A. U. Dullah, and J. Jumanto, “Global recession sentiment analysis utilizing VADER and ensemble learning method with word embedding,” *J. Soft Comput. Explor.*, vol. 4, no. 3, 2023, doi: https://doi.org/10.52465/joscex.v4i3.193.
[17] M. M. Kamruzzaman, “Arabic Sign Language Recognition and Generating Arabic Speech Using Convolutional Neural Network,” vol. 2020, 2020.
[18] M. Alaghband, H. Reza, and I. Garibay, “Machine Learning with Applications A survey on sign language literature,” *Mach. Learn. with Appl.*, vol. 14, no. June, p. 100504, 2023, doi: 10.1016/j.mlwa.2023.100504.
[19] S. Katoch, V. Singh, and U. Shanker, “Indian Sign Language recognition system using SURF with SVM and CNN,” *Array*, vol. 14, no. January, p. 100141, 2022, doi: 10.1016/j.array.2022.100141.
[20] H. Alsolai, L. Alsolai, F. N. Al-wesabi, M. Othman, M. Rizwanullah, and A. A. Abdelmageed, “Automated Sign Language Detection and Classification using Reptile Search Algorithm with Hybrid Deep Learning,” *HELIYON*, p. e23252, 2024, doi: 10.1016/j.heliyon.2023.e23252.
[21] A. Mannan, A. Abbasi, A. R. Javed, A. Ahsan, T. R. Gadekallu, and Q. Xin, “Hypertuned Deep Convolutional Neural Network for Sign Language Recognition,” vol. 2022, 2023.
[22] B. Navaneeth and M. Suchetha, “PSO optimized 1-D CNN-SVM architecture for real-time detection and classi fi cation applications,” *Comput. Biol. Med.*, vol. 108, no. September 2018, pp. 85–92, 2019, doi: 10.1016/j.compbiomed.2019.03.017.
[23] N. Kefyalew, T. Menore, T. Ayodeji, and O. Salau, “Recognition of Amharic sign language with Amharic alphabet signs using ANN and SVM,” *Vis. Comput.*, vol. 38, no. 5, pp. 1703–1718, 2022, doi: 10.1007/s00371-021-02099-1.
[24] R. Elakkiya, “Machine learning based sign language recognition: a review and its research frontier,” *J. Ambient Intell. Humaniz. Comput.*, vol. 12, no. 7, pp. 7205–7224, 2021, doi: 10.1007/s12652-020-02396-y.
[25] D. Li, C. R. Opazo, X. Yu, and H. Li, “Word-level Deep Sign Language Recognition from Video : A New Large-scale Dataset and Methods Comparison,” pp. 1459–1469, 2020.
[26] A. A. Hakim, E. Juanara, and R. Rispandi, “Mask Detection System with Computer Vision-Based on CNN and YOLO Method Using Nvidia Jetson Nano,” *J. Inf. Syst. Explor. Res.*, vol. 1, no. 2, Jul. 2023, doi: 10.52465/joiser.v1i2.175.
[27] Y. Chen, L. Li, W. Li, Q. Guo, Z. Du, and Z. Xu, “AI programming language for AI computing systems,” *AI Comput. Syst.*, 2023, doi: 10.1016/B978-0-32-395399-3.00014-7.
[28] Y. Luo, Y. Huang, Q. Wang, K. Yuan, Z. Zhao, and Y. Li, “An improved YOLOv5 model : Application to leaky eggs detection,” *LWT*, vol. 187, no. June, p. 115313, 2023, doi: 10.1016/j.lwt.2023.115313.
[29] S. O. Slim, I. A. Abdelnaby, M. S. Moustafa, M. B. Zahran, H. F. Dahi, and M. S. Yones, “Smart insect monitoring based on YOLOV5 case study : Mediterranean fruit fly Ceratitis capitata and Peach fruit fly Bactrocera zonata,” *Egypt. J. Remote Sens. Sp. Sci.*, vol. 26, no. 4, pp. 881–891, 2023, doi: 10.1016/j.ejrs.2023.10.001.
[30] N. F. Attia, M. T. Faheem, S. Ahmed, and M. A. M. Alshewimy, “Efficient deep learning models based on tension techniques for sign language recognition,” *Intell. Syst. with Appl.*, vol. 20, no. June, p. 200284, 2023, doi: 10.1016/j.iswa.2023.200284.

[31] S. Yap, B. N. Panggiri, and G. Darian, “Enhancing BISINDO Recognition Accuracy through Comparative Analysis of Three CNN Architecture Models,” *2023 Int. Conf. Inf. Manag. Technol.*, no. 2019, pp. 732–737, 2023, doi: 10.1109/ICIMTech59029.2023.10277780.
[32] M. Z. Fauzi, “Recognition of Real-Time BISINDO Sign Language-to-Speech using Machine Learning Methods,” *2023 Int. Conf. Comput. Sci. Inf. Technol. Eng.*, pp. 986–991, 2023, doi: 10.1109/ICCoSITE57641.2023.10127743.
[33] T. Handhika and I. Sari, “The Generalized Learning Vector Quantization Model to Recognize Indonesian Sign Language (BISINDO),” *Third Int. Conf. Informatics Comput.*, pp. 1–6, 2018.