Journal of Applied Engineering and Technological Science
Vol 4(1) 2022 : 215-222

DETECTION OF THE DEAF SIGNAL LANGUAGE USING THE SINGLE SHOT DETECTION (SSD) METHOD

Dadang Mulyana Iskandar[1], Mesra Betty Yel[2], Aldi Sitohang[3*]
Informatics Engineering Study Program, Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika, Indonesia[123]
aldisitohang20@gmail.com

Received : 26 August 2022, Revised: 03 September 2022, Accepted : 03 September 2022
**Corresponding Author*

ABSTRACT
Sign Language is a language that prioritizes manual communication, body language, and lip movements, instead of sound, to communicate. Deaf people are the main group who use this language, usually by combining hand shape, orientation and movement of the hands, arms, and body, and facial expressions to express their thoughts. Therefore, the researcher created an image recognition program in sign language using the Single Shot Detection (SSD) method, which is a convolution activity by combining several layers of preparation, by utilizing several components that move together and are motivated by a biological sensory system. The letters used in making sign language programs use the letters of the alphabet (az). This sign language detection programming that runs on the Google Collaboratory application.
Keywords : *Detection, Sign Language, SSD, Google Collaboratory*

1. Introduction

The development of information technology in the latest era, the world has been in the development of the industrial revolution 4.0 with transformations in all aspects, especially the use of the development of Artificial Intelligence (AI), one of which is machine learning(Abidi et al., 2022; Sujatha et al., 2022). Almost all remote areas can now feel the technology that is developing today. Information technology is used to process data, including processing, obtaining, compiling, storing, manipulating data in various ways to produce quality information, namely relevant, accurate and timely information used for personal, business, government and strategic information for decision making(Agarwal & Raj, 2022). The development of information technology has had a major impact on various fields of people's lives, both in terms of social, economic, educational, development, as well as tourism. Indonesia has 2 standard sign language systems that are commonly encountered, namely SIBI (Indonesian Sign Language System) and BISINDO (Indonesian Sign Language)(Hosea & Wirawan, 2022; Aditama et al., 2022). Based on the Regulation of the Ministry of Education and Culture of the Republic of Indonesia Number 061/U/1994 concerning Standardization of the Indonesian Sign Language System for the Deaf, SIBI is a sign language that has been inaugurated by the government in accordance with the applicable curriculum and as a teaching standard in Special Schools (SLB). In making this deaf sign language detection application using the Single Shot Detection (SSD) method for making detection applications, hand movements are used to detect the letters that you want to use in order to know the letters you want to say(Gupta et al., 2022).

As for classifying function values from all possible input values, a learning technique is needed or what is usually known as supervised learning or a label is given to the algorithm as the basis for truth(Leqi et al., 2022). According to experts, namely Kerlinger, stating that survey research methodology is research conducted on large or small populations, but the data studied are data from samples taken from the population, to find relative events, distributions, and relationships between variables. sociological and psychological. The RQ Formulation of effectiveness in research should focus on 5 elements known as PICOC:

1. Population (F): target group for investigation (e.g. people, software, etc.)
2. Intervention (I): determine the aspect of the investigation or problem that is of interest to the researcher
3. Comparation (C): The aspect of the investigation to which the Intervention (I) will be compared.

4. Outcomes (O): Effects and outcomes of the Intervention
5. Context (C):background or research environment

Table 1 - Review of PICOC

Survey Theme Title: Detection of Deaf Sign Language Using Single Shot Detection (SSD) Method	
Population	Detection, Sign Language, and Single Shot Detection (SSD)
Intervention	a. How to design a sign language detection support system for the general public. b. How to apply the Single Shot Detection (SSD) method in the process of detecting fingers so that they can be understood as text.
Comparison	N/A
Outcomes	a. To help translate people with special needs with general people who cannot speak sign language so that they can communicate. b. Make it easier to learn sign language with the help of a video camera to detect the letters we want to learn.
Context	Public Dataset

Programming Language

According to Bahar et al., (2022) "Programming language is a syntax for defining computer programs, this language allows a programmer to create an application program, for example: pyhton, javam borland delphi". As for other research according to Cox et al., (2022) suggests that "Programming language is a computer language used in writing programs."

Google Collaboration

Google Colab or Google Collaboratory, is an executable document that can be used to store, write, and share programs that have been written via Google Drive. The explanation was adapted from Tutorials Point, this software is basically similar to the free Jupyter Notebook in the form of a cloud that is run using a browser, such as Mozilla Firefox and Google Chrome(Sampathila et al. 2022).

Sign Language

Sign Language is a language that prioritizes manual communication, body language, and lip movements, instead of sound, to communicate(Abu-Jamie & Abu-Naser, 2022). Deaf people are the main group who use this language, usually by combining hand shape, orientation and movement of the hands, arms, and body, and facial expressions to express their thoughts. Contrary to the opinion of many, in reality no international sign language has been successfully implemented. Sign language is unique in its kind in each country(Chen et al., 2022). Sign language may differ in countries that speak the same language. For Indonesia, there is a sign language, namely Indonesian Sign Language (BISINDO) whose development is supported by a donor agency from Japan involving the Chinese University of Hong Kong and the University of Indonesia(Nurpiena et al., 2021).

Detect

In general, detection is an attempt to find and determine the existence, assumption, or reality. From the detection results, it can be seen who the perpetrator of the crime was by accessing information without special cooperation from the sender. In the history of radio communication, the term "detector" was first used for a device that detects the presence or absence of a radio signal, as all communications use Morse code(Hassija et al., 2021).

Single Shot Detector (SSD)

Single Shot Detector (SSD) is a method for recognizing or detecting an object in an image using a single deep neural network and one of the most popular object detection algorithms due to its ease of implementation, good accuracy vs. ratio required computation. Single Shot Detector (SSD) only needs to take a single shot to detect multiple objects in the image(Magalhães et al., 2021).

The Single Shot Detector (SSD) method is included in object detection in real time, although it is more intuitive than its counterparts such as R-CNN, Fast R-CNN Faster R-CNN and You Only Look Once (YOLO), Single Shot Detector (SSD).) is a very powerful algorithm. Being simple in design, the implementation is more direct from GPU and deep learning framework point of view and thus does the heavy lifting of detection at lightning speed(Nithin & Jaisharma, 2022).

2. Research Methods

In this study, the image of the deaf sign language spread from various sources on the internet. As for the sample in this study had 26 images of hand shapes, During the shooting process, four images were obtained for each background, which means 12 images for each alphabet. The total images from letters A to Z obtained are 312 images. The following are 26 images used in this study and their explanations:

Table 2 - Alphabet Image Dataset

No.	Variable	Train	Test	Variable Operational Definition
1	A	8	4	Image of hand shape A
2	B	8	4	Image of a B bentuk shape hand
3	C	8	4	C shaped hand image
4	D	8	4	D shaped hand image
5	E	8	4	Image of hand shape E
6	F	8	4	The image of a hand in the shape of F
7	G	8	4	G shaped hand image
8	H	8	4	The image of a H .-shaped hand
9	I	8	4	I shape hand image
10	J	8	4	Image of a J bentuk-shaped hand
11	K	8	4	K shape hand image
12	L	8	4	L shaped hand image
13	M	8	4	Image of a M-shaped hand
14	N	8	4	N shape hand image
15	O	8	4	Image of an O .-shaped hand
16	P	8	4	P shaped hand image
17	Q	8	4	Image of a Q-shaped hand
18	R	8	4	Image of an R .-shaped hand
19	S	8	4	S shape hand image
20	T	8	4	Image of a T . shaped hand
21	U	8	4	U shaped hand image
22	V	8	4	V shaped hand image
23	W	8	4	W shaped hand image
24	X	8	4	The image of an X-shaped hand
25	Y	8	4	Y shaped hand image
26	Z	8	4	Z shaped hand image

Data Collection Techniques

Data collection techniques were used to collect data according to research procedures in order to obtain the required data. Data collection techniques are the most strategic step in research because the main purpose of research is to collect data. Data collection techniques in this study using the Documentation Technique. Documentation is a method of reviewing and processing data from pre-existing documents and supporting research data. Documentation method is used to collect sign language data.

Application of the Methodology

The following are the stages of the methodology used in the research:

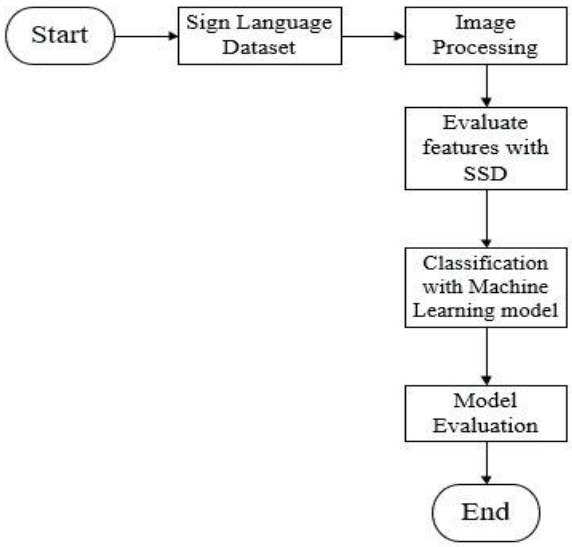

Fig 1. Stages of Methodology Application

Data Collection

The process of collecting data in this study used the scraping method. The scraping method used is to download images related to the object in this study, namely the image of the deaf sign language hand shape in the variables of this study which was obtained on the kaggle website.

Dataset Creation

The image data in this study are downloaded from the kaggle website. There are three backgrounds used in taking pictures, namely a plain white shirt, a white wall, and a white shirt with dots. The angle of view of the image taken is the front view with the object distance from the lens of approximately 70 cm. In this study, two separate datasets were used, each containing training data and test data.

Test Design

The test design is the stage to find out how the performance of the system that has been built. This test is carried out in two stages, namely the training stage and the testing stage. The design of this test is to be able to find out the results of the analysis and the design carried out is in accordance with what is expected. At this stage, the researcher will design an application design based on Single Shot Detector (SSD) for Extracting Text on Images to be created.

3. Results and Discussions

Implementation Methodology

In this study, to detect sign language in the deaf using the scraping method from various sources on the internet. The scraping method used is by downloading images from the kaggle website using 4 attributes consisting of variables, training, testing, and Variable Operational Definitions. In making the model, it is necessary to sample data in the form of images to detect sign language in the form of hands in each alphabet. In this case the image to be detected is in the form of 26 hand-shaped images, in shooting, four images are obtained for each background, which means 12 images for each alphabet. The total images from letters A to Z obtained are 312 images. The initial stage in making a model for detection is done by training to classify sign language models with the help of a video camera to detect letters to be easily understood by children with special needs. In the training model there are training data (train) and test data (test). Train data is used to train the model and test data is used to validate the training data. Model training is done using Jupyter Programming(Chandel et al., 2022).

Data collection

This data collection uses 26 variable images obtained from the Kaggle website. The images used in this study are in the form of hand drawings based on the alphabet in each movement.

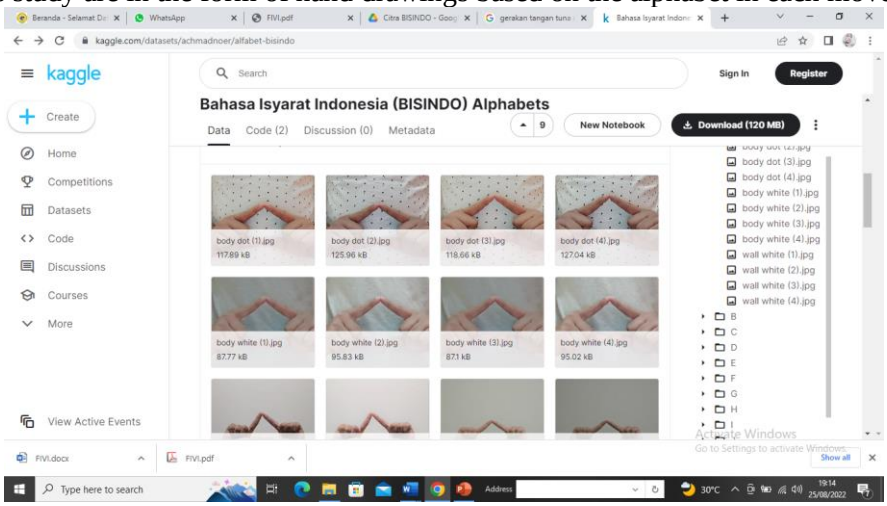

Fig 2. Image of Alphabet Hand Shapes

Then on the Kaggle website, data from 26 variable images of hand movements, before being saved, the image can be selected first which images can be used in this study. The stored images will be used to carry out the training and testing process.

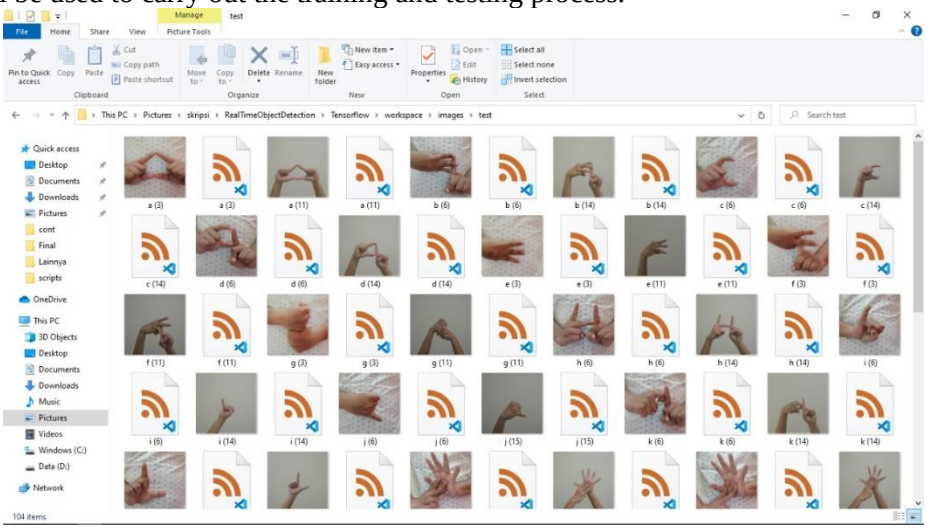

Fig 3. Image Download Results

Dataset Creation

After the image data is downloaded, then create a dataset folder containing the train folder and the test folder. In the train and test folders, folders are created based on the image variables in each Alphabet consisting of AZ. The dataset is divided into 208 training data and 104 testing data. Before making the dataset for the detection process, the image files that we have collected are put into a folder. Then the creation of the dataset is done using the jupyter programming application with several stages carried out. Before making the dataset for the detection process, the image files that have been collected into a folder earlier, are entered into the jupyter program directory first. Then the creation of the dataset is done using the jupyter application with several stages carried out.

1. Create Label Map

```
In [2]: labels = [{'name':'a', 'id':1},
                  {'name':'b', 'id':2},
                  {'name':'c', 'id':3},
                  {'name':'d', 'id':4},
                  {'name':'e', 'id':5},
                  {'name':'f', 'id':6},
                  {'name':'g', 'id':7},
                  {'name':'h', 'id':8},
                  {'name':'i', 'id':9},
                  {'name':'j', 'id':10},
                  {'name':'k', 'id':11},
                  {'name':'l', 'id':12},
                  {'name':'m', 'id':13},
                  {'name':'n', 'id':14},
                  {'name':'o', 'id':15},
                  {'name':'p', 'id':16},
                  {'name':'q', 'id':17},
                  {'name':'r', 'id':18},
                  {'name':'s', 'id':19},
                  {'name':'t', 'id':20},
                  {'name':'t', 'id':21},
                  {'name':'v', 'id':22},
                  {'name':'w', 'id':23},
                  {'name':'x', 'id':24},
                  {'name':'y', 'id':25},
                  {'name':'z', 'id':26},
                ]

with open(ANNOTATION_PATH + '/label_map.pbtxt', 'w') as f:
    for label in labels:
        f.write('item { \n')
        f.write('\tname:\'{}\'\n'.format(label['name']))
        f.write('\tid:{}\n'.format(label['id']))
        f.write('}\n')
```

Fig 3. Creating a Label

The next process is to load image data and label data from the dataset folder. determine the location (path) of training data (train) used as training data and test data (test) as validation data.

Test

Tests are carried out to obtain accuracy results from each test model that was previously made. The image data used in this research is 26 variable image data with 208 training data and 104 test images. Next, do image data training, with the results of the accuracy of the test, where the correct number of datasets is divided by 100. After training the model, then evaluate the performance of the model on the test set. Evaluation is carried out to see the possibility of failure of the image object that is read in the detection process, it will also get the accuracy value and with the highest probability that will be obtained from the entire test model.

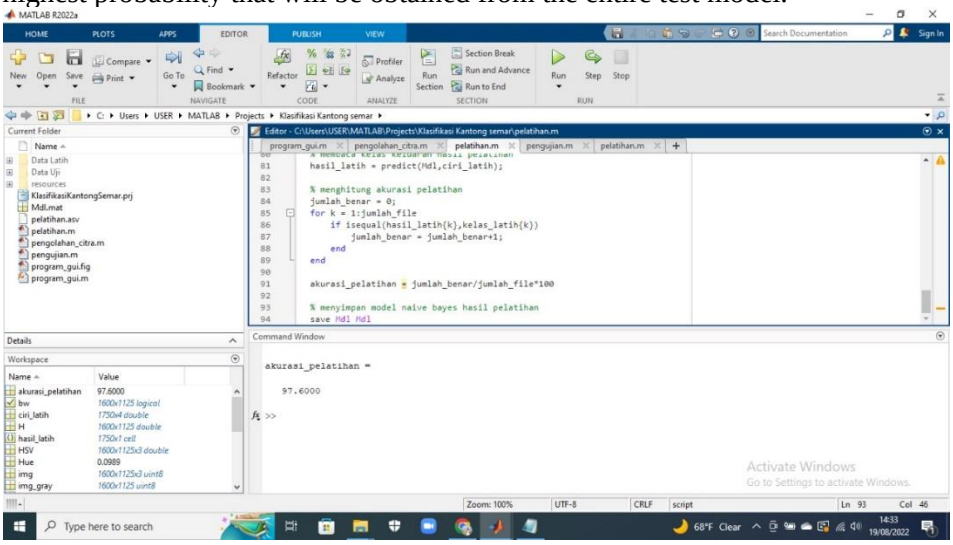

Fig 4. Training Evaluation

The picture above is the code used to see the evaluation of the performance of the Single Shot Detector (SSD), from the picture it can be seen the evaluation of the accuracy generated from the train data with an accuracy value of 90.00%.

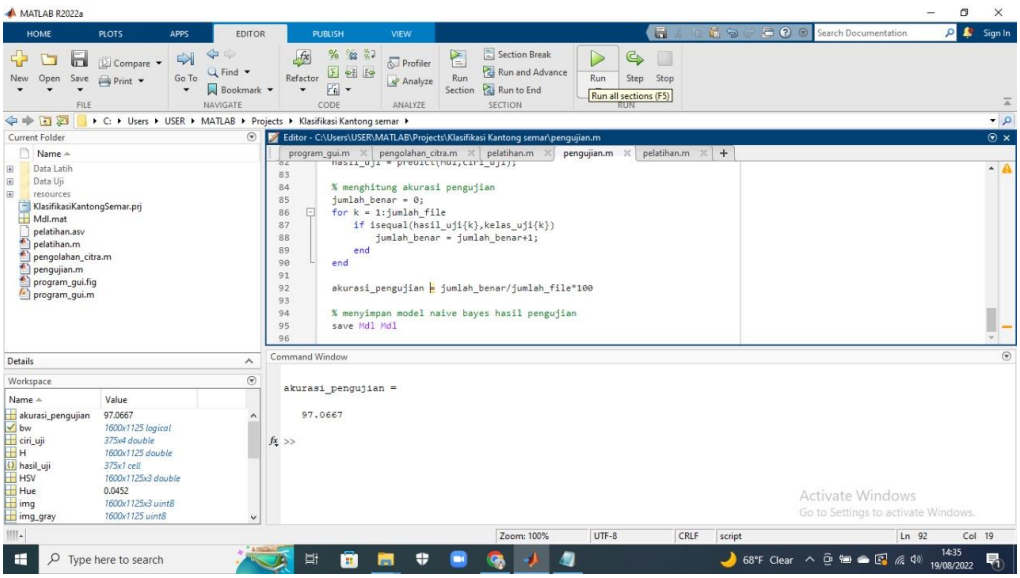

Fig 5. Test Evaluation

The picture above is the code used to see the evaluation of the performance of the Single Shot Detector (SSD), from the picture it can be seen the evaluation of the accuracy generated from the test data with an accuracy value of 90.00%.

Final Test Results

The evaluation of the detection results from the Single Shot Detector (SSD) method, which was trained using train data as many as 208 images, obtained an evaluation value generated from the test data as many as 104 images by testing the accuracy values obtained as follows.

Table 1 - Evaluation of Accuracy Value

Single Shot Detector (SSD) Detection	
Training Accuracy	*Testing Accuracy*
90.00%	90.00%

4. Conclusion

Based on the research and the results of the application of the Single Shot Detection (SSD) method in classifying 26 types of images, the following conclusions can be drawn: Accuracy will be achieved better if larger data trains are used. This is evidenced by testing the amount of training data as much as 80% and validation data as much as 20% in each test model. And then by evaluating the Single Shot Detection (SSD), the accuracy value is obtained based on train data of 90.00% and test data of 90.00% accuracy value.

References

Abidi, M. H., Mohammed, M. K., & Alkhalefah, H. (2022). Predictive Maintenance Planning for Industry 4.0 Using Machine Learning for Sustainable Manufacturing. *Sustainability*, *14*(6), 3387.

Abu-Jamie, T. N., & Abu-Naser, S. S. (2022). Classification of Sign-language Using VGG16. *International Journal of Academic Engineering Research (IJAER)*, *6*(6).

Aditama, P. W., Anggara, I. G. A. S., & Jayanegara, I. N. (2022). Implementation of Sibi And Bisindo Letters Recognition Using Augmented Reality During Pandemic. *IJISTECH (International Journal of Information System and Technology)*, *5*(5), 602-611.

Agarwal, H., & Raj, V. J. P. (2022). Adoption of Human Resource Analytics in Information Technology and Information Technology Enabled Services Industry in India. *Journal of Data Science, Informetrics, and Citation Studies*, *1*(1), 50-57.

Bahar, A. Y., Shorman, S. M., Khder, M. A., Quadir, A. M., & Almosawi, S. A. (2022, June). Survey on Features and Comparisons of Programming Languages (PYTHON, JAVA,

AND C#). In *2022 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS)* (pp. 154-163). IEEE.

Chandel, S., Clement, C. B., Serrato, G., & Sundaresan, N. (2022). Training and evaluating a jupyter notebook data science assistant. *arXiv preprint arXiv:2201*.12901.

Chen, Y., Wei, F., Sun, X., Wu, Z., & Lin, S. (2022). A simple multi-modality transfer learning baseline for sign language translation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 5120-5130).

Cox, R., Griesemer, R., Pike, R., Taylor, I. L., & Thompson, K. (2022). The Go programming language and environment. *Communications of the ACM*, *65*(5), 70-78.

Gupta, A. M., Koltharkar, S. S., Patel, H. D., & Naik, S. (2022, March). DRISHYAM: An Interpreter for Deaf and Mute using Single Shot Detector Model. In *2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS)* (Vol. 1, pp. 365-371). IEEE.

Hassija, V., Chamola, V., Agrawal, A., Goyal, A., Luong, N. C., Niyato, D., ... & Guizani, M. (2021). Fast, reliable, and secure drone communication: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, *23*(4), 2802-2832.

Hosea, W. K., & Wirawan, I. (2022). The Impact of Implementing the Gamification Method in Learning Indonesian Sign Language with Bisindo Vocabulary. *International Journal of Open Information Technologies*, *10*(9), 62-69.

Leqi, L., Huang, A., Lipton, Z., & Azizzadenesheli, K. (2022, June). Supervised Learning with General Risk Functionals. In *International Conference on Machine Learning (pp. 12570-12592)*. PMLR.

Magalhães, S. A., Castro, L., Moreira, G., Dos Santos, F. N., Cunha, M., Dias, J., & Moreira, A. P. (2021). Evaluating the single-shot multibox detector and YOLO deep learning models for the detection of tomatoes in a greenhouse. *Sensors*, *21*(10), 3569.

Nithin, A., & Jaisharma, K. (2022, February). A Deep Learning Based Novel Approach for Detection of Face Mask Wearing using Enhanced Single Shot Detector (SSD) over Convolutional Neural Network (CNN) with Improved Accuracy. *In 2022 International Conference on Business Analytics for Technology and Security (ICBATS)* (pp. 1-5). IEEE.

Nurpiena, S. A., Wihidayat, E. S., & Budianto, A. (2021). Developing Indonesia Sign Language (BISINDO) Application with Android Based for Learning Sign Language. *Journal of Informatics and Vocational Education*, *4*(1).

Sampathila, N., Chadaga, K., Goswami, N., Chadaga, R. P., Pandya, M., Prabhu, S., ... & Upadya, S. P. (2022, September). Customized Deep Learning Classifier for Detection of Acute Lymphoblastic Leukemia Using Blood Smear Images. In *Healthcare* (Vol. 10, No. 10, p. 1812). MDPI.

Sujatha, M., Priya, N., Beno, A., Blesslin Sheeba, T., Manikandan, M., Tresa, I. M., ... & Thimothy, S. P. (2022). IoT and Machine Learning-Based Smart Automation System for Industry 4.0 Using Robotics and Sensors. *Journal of Nanomaterials*.