2023 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (ViTECoN) | 979-8-3503-4798-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/VITECON58111.2023.10157784

Real-Time Sign Language Detection

Sangeeta Kurundkar
Department of Electronics and Telecommunication
Vishwakarma Institute of Technology
Pune, India
sangeeta.kurundkar@vit.edu

Arya Joshi
Department of Electronics and Telecommunication
Vishwakarma Institute of Technology
Pune, India
arya.joshi21@vit.edu

Aryan Thaploo
Department of Electronics and Telecommunication
Vishwakarma Institute of Technology
Pune, India
aryan.thaploo21@vit.edu

Sarthak Auti
Department of Electronics and Telecommunication
Vishwakarma Institute of Technology
Pune, India
sarthak.auti21@vit.edu

Anish Awalgaonkar
Department of Electronics and Telecommunication
Vishwakarma Institute of Technology
Pune, India
anish.awalgaonkar21@vit.edu

Abstract— **Lack of communication or miscommunication brought on by linguistic problems can lead to awkward situations in today's culture. Those who are deaf or have trouble hearing use sign language, a visual form of communication, to communicate. It communicates meaning through hand gestures and body language. Yet not everyone is able to interpret sign language, which might result in miscommunication. A system was developed employing cutting-edge technologies to address this problem, including deep learning, machine learning, convolutional neural networks, computer vision, TensorFlow, and Python. This technology is made to accurately detect and identify sign language motions in real-time. The system develops a real-time sign language recognition tool using OpenCV. The 26 letters in American Sign were categorised using a CNN classifier. The use of technology to bridge communication gaps and create inclusive environments is crucial in our society. This system's high accuracy rate is an excellent indication of its reliability, providing a promising solution for individuals who use sign language to communicate. It would be interesting to learn more about the specific CNN classifier used in the project and how it was trained to recognize ASL gestures. Overall, the implementation of this technology can create more inclusive and accessible environments, ensuring that everyone can communicate effectively, regardless of their abilities or differences.**

***Keywords*— Sign Language, Real-time, OpenCV, American Sign Language, CNN, Deep learning**

I. INTRODUCTION

Language is a crucial tool for communication, allowing individuals to express themselves, connect with others, and lead a more harmonious life. With over 7000 languages spoken worldwide, the power of language cannot be underestimated. Visual languages such as sign languages make use of hand, body, and eye motions. For people with disabilities to have a means of communication, sign languages are crucial. Specially abled people can express themselves, communicate with others, and share their feelings using sign languages. The disadvantage is that not everyone is familiar with sign languages, which restricts communication between those with disabilities and average people. Automated solutions for sign language translation that can easily translate signs could help to get around this restriction. In fact, more than 100 different sign languages, including Indian Sign Language, American Sign Language (ASL), Italian Sign Language, German Sign Language, and many others, are used worldwide to assist the deaf and speech-impaired. Almost 70 million deaf people use sign language as their first language worldwide, and many hearing people also use it as their first language. Despite the fact that American Sign Language (ASL) is a highly common sign language, not everyone is familiar with it. To address this issue, a system has been developed that accurately predicts ASL letters in real time. This pattern recognition system is designed to understand human gestures, and if machines can recognize and differentiate these patterns, the desired message can be recreated. Currently, the system is capable enough to detect static drawing gestures for letters and numbers, but it has the potential to recognize words and phrases as well. The technology has advanced significantly in recent years, and extensive research has been conducted to help the deaf community. Furthermore, the system can be extended to create automated editors, allowing users to write with gestures and communicate more accurately, such as in legal documents. ASL, like spoken language, has similar features, and it is a common language used for communication through hand and facial gestures. The American Sign Language alphabet is illustrated in Fig.

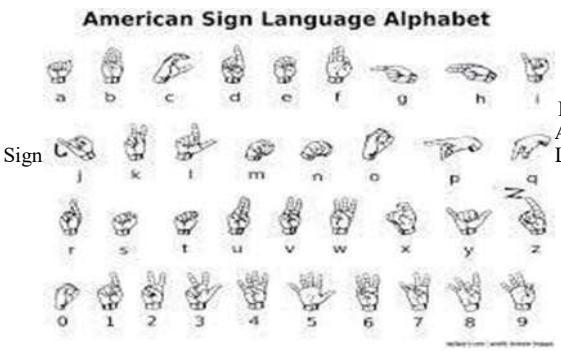

Figure.1. Sign American Language Alphabets[16]

In this project, we take a dataset from Kaggle and use the Python programming language to train images and classify them into their respective types, alphabets, using a CNN classifier. Real- time algorithms then classify the gestures into different categories. We use several Python libraries to achieve this goal.

II. Literature Review

Author introduced a novel architecture for a sign language identification system that utilizes recognized characteristics and extraction techniques for hand movements using the image processing algorithm Speeded Up Robust Features. This work was published in Volume 7 in 2017.[2]

Manisha Ingle proposed a vision-based method for recognizing hand gestures. In their approach, they converted RGB to grayscale throughout the segmentation procedure to avoid the possibility of pointless detections. They presented a method using MATLAB called the improvised Scale Invariant Feature Transform.[3]

The paper describes a promising approach for a sign language detection system that relies on the Hidden Markov Model. The input system uses feature vectors that replicate hand signs. A camera captures a picture, which is then processed using various algorithms. The symbol is recognized, and the text is displayed using a template matching algorithm. To implement the system, several different libraries such as Python and OpenCV are used.[4]

III. Methodology

A. Technological tools used:

- VS Code IDE
- Python Programming Language
- CNN (Convolution neural network).

B. Methodology:

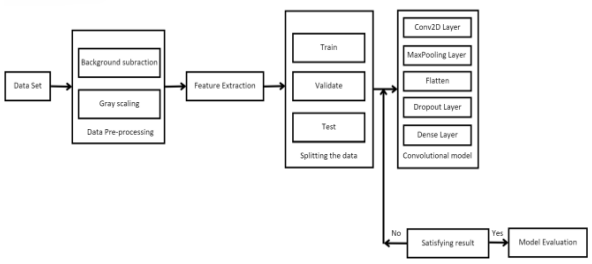

Figure.2. Flow of the proposed model [13]

1. Collecting the dataset:
The collection of data is a critical aspect of any research or investigation. It is crucial to gather pertinent information that satisfies the study requirements, including extreme cases, to produce thorough and realistic observations. The data must be carefully collected to ensure correctness and precision because it forms the basis for the research.

2. Loading and Pre-processing of Dataset:
The initial step in the process is to import the dataset and make necessary modifications to enable the implementation of various algorithms for accurate prediction of sign language alphabet labels. The dataset contains grayscale images, which are conducive to real-time detection, and has been split into separate training and testing subsets. Corresponding labels have been assigned to each image, allowing for efficient classification by algorithms.

3. Model Building:
The machine learning model was implemented using the Keras library and included convolutional layers, max pooling, and flattening techniques. To prevent overfitting, the model was further regularized using batch normalization and dropout layers. Additionally, a call back function was used to monitor the accuracy of the validation set and halt the training process if the accuracy failed to improve over a certain number of consecutive epochs. Compiling and training the model.The latest technique for detecting objects using deep learning is based on convolutional neural networks (CNN). This method involves passing the input image through a CNN to produce a convolutional feature map. CNNs are a type of neural network commonly used in deep learning for image recognition and processing of pixel data. While there are other types of neural networks, CNNs are the preferred architecture for recognition and object detection tasks. The basic architecture of a CNN involves several layers of convolution and pooling operations, followed by fully connected layers for classification.:

Explaining 5 Layers of Convolutional Neural Network

- Convolutional Layer.
- Pooling Layer.
- Fully Connected Layer.
- Dropout.
- Activation Functions

A user-friendly yet reliable deep learning library made specifically for Python is called Keras. The purpose of this project is to train a fundamental convolutional neural network (CNN) using Keras and then use it to solve a real-world problem. The output layer will make use of the cross-entropy loss function, with categorical cross entropy being one sort of cross-entropy loss that Keras specifically recognises. (>2 classes), hence the latter will be used. a list of stats. We only allow Keras to report accuracy numbers because this is a classification problem. In Keras, training a model is as simple as using fit and supplying a few parameters. There are numerous available parameters.

1. The training data, also known as X and Y, which consists of labels and images.
2. The total number of epochs (i.e., dataset iterations)to train for.
3. The validation data (also known as test data), which is used in training to periodically assess how well the network performs against fresh data.

IV. RESULTS

In this project, signs were detected and translated to aid people with disabilities in simple social interaction. This system utilised deep learning methods and produced effective results. The results achieved are discussed. It has been accomplished in this paper utilising the TensorFlow object detection API. The technology instantly recognises sign language. Images were taken for data acquisition utilising the laptop's webcam with Python and OpenCV. The dataset that the system was trained on is modest, despite the fact that it has a high accuracy rate of 99%. Following are the graphs of the results obtained:

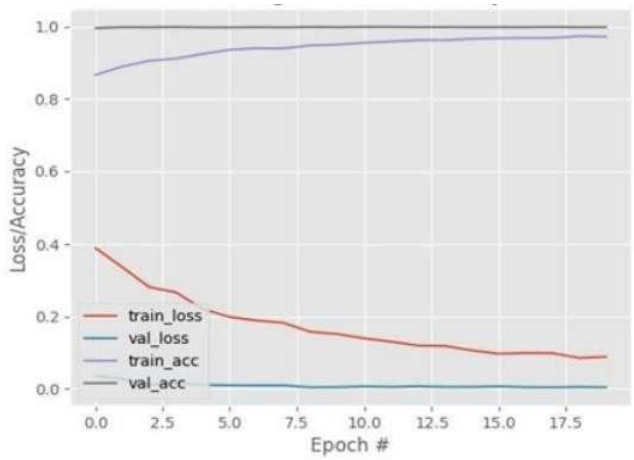

Figure.5 accuracy

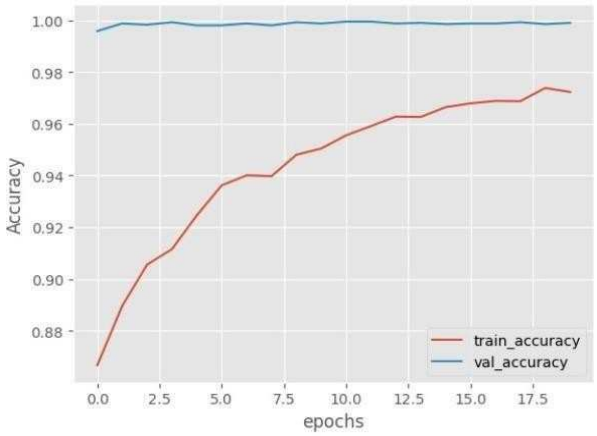

Figure.4 Epochs vs Accuracy

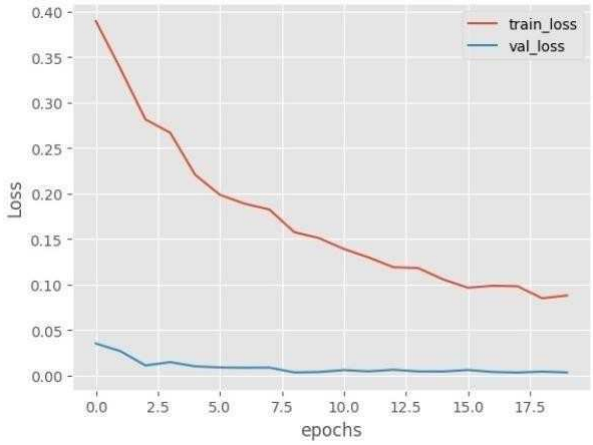

Figure.5. Epochs vs Loss

V. CHALLENGES

Finding the appropriate dataset proved to be the hardest obstacle for us. Some datasets simply had a few symbols, while others just contained a few photos. Choosing the right dataset-model combination is difficult because some models tend to overfit a particular dataset. Some of the areas that will be addressed in the future are implementing our model for use with different sign languages and improving the model's ability to recognize words and phrases

VI. CONCLUSION

Real-time sign language detection is a rapidly advancing field of research that has the potential to revolutionize communication for the deaf and hard-of-hearing community. By using machine learning and computer vision technologies, it is now possible to accurately recognize and translate sign language gestures in real-time. However, there are still challenges that need to be addressed, such as improving the accuracy of recognition algorithms, developing systems that can work in various lighting and background conditions, and making the technology more affordable and accessible to a wider range of people. Despite these challenges, the progress that has been made so far is promising, and it is likely that real- time sign language detection will continue to evolve and become more widely used in the future.

VII. ACKNOWLEDGMENT

We would like to thank the open-source community on the Internet, especially the TensorFlow mentors, whose help has been an integral part of our learning process.

VIII. REFERENCES

[1] Suharjito, Ricky Anderson, Fanny Wiryana, Meita Chandra Ariesta, Gede Putra Kusuma “Sign Language Recognition Application Systems for Deaf-Mute People: A Review Based on Input-Process-Output”.

[2] Kamal Preet Kour, Dr. Lini Mathew" Sign Language Recognition Using Image Processing", International Journals of Advanced Research in Computer Science and Software Engineering, Vol. 7, Issue 8, 2017.

[3] Neelam K. Gilorkar, Manisha M. Ingle, “Real Time Detection And Recognition Of Indian And American Sign Language Using Sift”, International Journal of Electronics and Communication Engineering & Technology, Vol. 5, Issue 5, 2014.

[4] B. Bauer,H. Hienz "Relevant features for video- based continuous sign language recognition", IEEE International Conference on Automatic Face and Gesture Recognition, 2002.

[5] S. Shrenika, M.M. Bala “Sign Language Recognition Using Template Matching Technique”, IEEE conference 2020.

[6] Mahesh Kumar N B “Conversion of Sign Language into Text” International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 9 (2018) pp. 7154-7161.

[7] Sanket Bankar1, Tushar Kadam2, Vedant Korhale3, Mrs.

[8] A. A. Kulkarni4 “Real Time Sign Language Recognition Using Deep Learning” International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 09 Issue: 04 | Apr 2022.

[9] "Machine Learning-based Hand Sign Recognition" by Ms. Greeshma Pala, Ms. Jagruti Bhagwan Jethwani, Mr. Satish Shivaji Kumbhar, and Ms. Shruti Dilip Patil. Proceedings of the International Conference on Artificial Intelligence and Smart Systems (ICAIS-2021) IEEE Xplore Part Number: CFP21OAB-ART; ISBN: 978-1-7281-9537-7.

[10] Amrutha K, Prabhu P “ML Based Sign Language Recognition System” 2021 International Conference on Innovative Trends in Information Technology (ICITIIT) | 978- 1-6654-0467-9/21/$31.00 ©2021 IEEE.

[11] Subhashini Yadav, Shreyashi Raj, Kashish Awasthi, Rohit Bidwan, Lokesh Jain “Real-Time Sign Language Detection” ISSN: 2321-9653; IC Value: 45.98; Volume 10 Issue V May 2022 International Journal for Research in Applied Science & Engineering Technology (IJRASET).

[12] Muthu Mariappan H, Dr Gomathi V “Real-Time Recognition of Indian Sign Language” Second International Conference on Computational Intelligence in Data Science (ICCIDS-2019).

[13] Muskan Dhiman, Dr G.N. Rathna “SIGN LANGUAGE RECOGNITION” Summer Research Fellowship Programme of India's Science Academy.

[14] Kai Zhao1, Kejun Zhang2, Yu Zhai2, Daotong Wang2, Jianbo Su1,2 “Real-Time Sign Language Recognition Based on Video Stream” Proceedings of the 39th Chinese Control Conference July 27-29, 2020, Shenyang, China.

[15] Dr. Dhananjay Kalbande, Aditya Das, Shantanu Gawde, and Khyati Suratwala, "Sign Language Recognition Using Deep Learning on Specially Processed Static Gesture Pictures."

[16] Kaustubh jadhav1 , abhishek jaiswal2 , abbas munshi3 , mayuresh yerendekar4, “sign language recognition using neural network”.

[17] https://www.pinterest.com/pin/sign-languages-on-twitter--460070918189217920/