International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 08 | Aug 2022 www.irjet.net p-ISSN: 2395-0072

Real-Time Sign Language Detector

Akhilesh Shukla[1], Dr. Devesh Katiyar[2], Mr. Gaurav Goel[3]

[1]*Student*
[2]*Assistant Professor*
[3]*Assistant Professor, Dr. Shakuntala Misra National Rehabilitation University, Mohaan Road, Lucknow, India*

Abstract: *People who are deaf or hard of hearing, as well as others who are unable to communicate verbally, use sign language to interact with people and within their communities. A group of modern languages known as sign languages use a visual-manual modality to convey information. Real-time finger-spelling recognition in Sign Language presents a conundrum that is examined. Using webcam photos, we produced a dataset for the usual hand gestures used in ISL as well as a dataset for the identification of 36 different gestures (alphabets and numbers). Using a hand motion as input, the system instantly displays the recognized character on the monitor screen. This human-computer interaction (HCI) project aims to recognize several alphabets (a-z), digits (0-9) and a number of common ISL hand gestures. We employed a Pre-Trained SSD Mobile net V2 architecture trained on our own dataset to apply Transfer learning to the issue. We developed a solid model that reliably categorises Sign language in the vast majority of cases. The use of sensors (such glove sensors) and other image processing methods (like the edge detection approach, the Hough Transform, and others) has been used in numerous research in the past in this field, but these technologies are rather expensive, and many people cannot afford them. Different human-computer interaction methods for posture recognition were researched and assessed during the project. An assortment of image processing strategies with human movement categorization were shown to be the best answer. The technology can identify selected Sign Language signs with an accuracy of 70–80% even in low light and without a controlled background. As a result, we developed this free and user-friendly app to help such people. Apart from a small number of people, not everyone is conversant in sign language, thus they could need an interpreter, which could be troublesome and expensive. By developing algorithms that can instantly foresee alphanumeric hand motions in sign language, this research aims to close the communication gap. The major objective of this project is to develop an intelligent computer-based system that will enable deaf people to efficiently communicate with others by using hand gestures.*

Keywords: TensorFlow, OpenCV, Sign Language, HCI, Python, Machine Learning

1. INTRODUCTION

The main reason sign languages were developed was to help the dumb and deaf. They use a coordinated and accurate combination of hand motions, hand forms, and hand orientation to convey precise information. This HCI (Human Computer Interface) project aims to recognise several alphabets (a-z), digits (0-9) and several common ISL family hand gestures, including thank you, hello, and other gestures. ISL recognition is particularly challenging since it requires the use of both hands, which is a challenging problem for hand-gesture recognition.

Many research have used sensors (such glove sensors) and different image processing methods (like edge detection, Hough Transform, etc.) in the past, but these methods are rather expensive, and many individuals cannot afford them. In India, many people use hand gestures to communicate with others because they are deaf or hard of hearing. Apart from a limited group of people, not everyone is conversant in sign language, thus they could require an interpreter, which could be difficult and expensive. Building software that can instantly predict ISL alphanumeric hand movements is the aim of this project, which aims to close the communication gap.

2.PROBLEM STATEMENT

Hearing loss is the most prevalent sensory deficit in today's population. The prevalence of Significant Auditory Impairment in India is estimated by the WHO to be 6.3% of the population, affecting over 63 million people. According to the NSSO research, 291 out of every 100,000 people have severe to profound hearing loss (NSSO, 2001). Many of them are young children between the ages of 0 and 14. There is a tremendous loss of physical and financial output since there are so many young Indians who are hearing-impaired. The fundamental issue is that those with hearing impairments, such as the deaf and dumb, find it challenging to interact with hearing-impaired people since hearing-impaired people do not learn sign language.

The answer is to create a translator that can recognise sign language used by a disabled person, feed that sign into a transfer learning machine learning algorithm, which is then recognised by the neural network and translated on the screen so that a non-disabled person can understand what the sign is saying.

Thanks to translators and speech-to-text technology, it is now much simpler. But what about the people who are deaf or hard of hearing? The major objective of this project is to develop a programme that can help those who are deaf or hard of hearing. A very important problem is also the language barrier. Individuals who are unable to talk communicate using hand signals and gestures. The average person has problems understanding their own language. Therefore, a system that can recognise different signs and gestures and communicate information to common people is needed. It links people with physical disabilities and those without. We may use computer vision and neural networks to recognise the cues and produce the appropriate text output. Permit the individual to speak on his or her own. Make yourself accessible to people who are trying to save money. Everyone is welcome to use it, and it is totally free. There are numerous businesses developing products for the deaf and hard of hearing, but not everyone can afford Them. Some are prohibitively expensive for average middle-class people to transport.

3.RELATED WORK

The ways that humans and computers communicate have changed as a result of the ongoing advancements in information technology. A lot of research has been done in this area to improve communication between hearing and hearing-impaired people. Any attempt to recognise sign language falls under the umbrella of human computer interaction since sign language is a collection of hand movements and attitudes. Detecting sign language is divided into two categories.

The Data Glove methodology falls under the first category. In this method, the user dons a glove with electromechanical devices connected to digitise hand and finger motion into actionable data. This approach has the drawbacks of being less accurate and requiring constant additional gear. The second group, computer-vision-based techniques, simply need a camera, enabling natural contact between people and computers without the need for any additional hardware. In addition to various advancements in the ASL field, Indians began working in ISL.

Comparing a new image's key point to the key points of standard images organised alphabetically in a database, for example, in order to identify the new image with the label of the closest match. Similar to this, other efforts have been made to effectively recognise edges. One such concept was to employ bilateral filtering in combination with colour data in depth images to correct edges.

People are using deep learning and neural networks to improve detection systems as they get more advanced. The Histogram technique, the Hough transform, OTSU's segmentation algorithm, and a neural network are just a few examples of the feature extraction and machine learning methods used in reference to identify the ASL. The process of processing photographs with a computer involves gathering, processing, analysing, and comprehending the findings. Low-level image processing is required for image quality improvement (such as eliminating noise and boosting contrast) while higher-level pattern recognition and image understanding are required to recognise characteristics in the image.

4.HAND GESTURE AND SIGN LANGUAGE RECOGNITION TECHNIQUE

In order to foresee and string hands into sentences that are semantically correct and intelligible, techniques like segmenting hands from the backdrop and recognising hand motion trajectories for distinctive indications are applied. Additionally, problems with gesture detection include motion modelling, motion analysis, pattern identification, and machine learning. SLR models employ manually constructed or automatically adjusted parameters. The background and surroundings of the model, such as the lighting in the room and the speed of the motions, have an impact on the model's capacity to do the categorization. The gesture seems unique in 2D space as a result of variations in viewpoints. Gestures can be recognised using a variety of techniques, including sensor-based and vision-based systems. In the sensor-based approach, devices with sensors capture a wide range of characteristics, including the trajectory, location, and velocity of the hand. The use of photographs or video recordings of hand motions is used in vision-based techniques, on the other hand. The procedures used to achieve sign language recognition are as follows:

4.1 The camera utilized by the system for recognizing sign language

The frame taken by a web camera on a laptop or PC is the foundation of the suggested sign language detection system. Image processing is carried out using the OpenCV computer library in Python.

Capturing Images:

In order to improve accuracy using a huge dataset, many pictures of distinct sign language symbols were captured from various angles and under varied lighting circumstances.

4.2 Segmentation:

A specific area of the full image that contains the sign language symbol that needs to be anticipated is then chosen as the capture stage is completed. For the sign to be seen, bounding boxes are enclosed. These boxes ought to be tightly encircling the area that has to be identified from the image. The labelled hand movements had specific names assigned to them. The labelling portion was performed using Google's Teachable Machine.

Image selection for training and test purposes

4.3 Creating TF Records

The numerous training and test photos were converted into record files.

4.4 Classification

Both supervised and unsupervised learning strategies are acceptable categories. In order to train a system to recognise patterns in incoming data and anticipate future data, one method is known as supervised machine learning. Inferring a function using supervised machine learning involves applying a set of known training data to labelled training data.

5.TECHNOLOGY UTILIZED

5.1 Interface

PyCharm for inserting python libraries in a script format, it is typically a python code where we can easily estimate our data sets model in one single notebook.

5.2 Operating System

Windows 11

5.3 Software

Python, IDE(PyCharm), NumPy, cv2 (OpenCV), TensorFlow, MediaPipe, Google's Teachable Machine

5.2 Hardware Environment

RAM- 8GB, GRAPHIC CARD, ROM-512GB SSD

6.TOOLS USED

6.1 TensorFlow

TensorFlow is an end-to-end open source platform for machine learning. It has a vast, adaptable ecosystem of tools, libraries, and community resources that enables academics to advance the state-of-the-art in machine learning and developers to quickly create and deploy ML-powered apps.

6.2 Keras

The Keras API was created with people in mind, not with computers. Best practises for lowering cognitive load are followed by Keras, which provides consistent & simple APIs, reduces the amount of user activities necessary for typical use cases, and offers clear & actionable error signals. Additionally, it contains a wealth of developer instructions and documentation.

6.3 OpenCV

An open-source, highly efficient Python library called OpenCV was created to address computer vision-related problems. Its main emphasis is on real-time applications that offer computational efficiency for handling enormous amounts of data. To identify objects, people, and even human handwriting, it processes images and videos.

6.4 Teachable Machine:

Teachable Machine is a web platform that enables quick and simple machine learning model creation for your projects without the need for coding. Export your model for use on websites, applications, and more after teaching a computer to recognise your sights, sounds, and poses.

7.ALGORITHM USED

7.1 Convolutional Neural Network

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning system that can take an input image and give various elements and objects in the image importance (learnable weights and biases), as well as distinguish between them. Compared to other classification algorithms, a ConvNet requires far less pre-processing. In contrast to basic approaches, which need hand-engineering of filters, ConvNets can learn these filters and properties with sufficient training. ConvNets are multilayer artificial neural networks that can process input from either 2D or 3D data. Each layer in the network is composed of a number of 2D or 3D planes, and each plane is made up of a large number of

independent neurons. Neighboring layer neurons are connected, but those in the same layer are not.

By using the right filters, a ConvNet may capture the spatial and temporal features of a picture. Furthermore, the architecture performed better fitting to the image collection as a result of lowering the number of parameters involved and reusing weights. ConvNet's main objective is to simplify image processing by the extraction of pertinent characteristics from images while keeping essential data required for successful prediction. This is quite helpful for creating an architecture that can manage huge amounts of data in addition to gathering and learning characteristics.

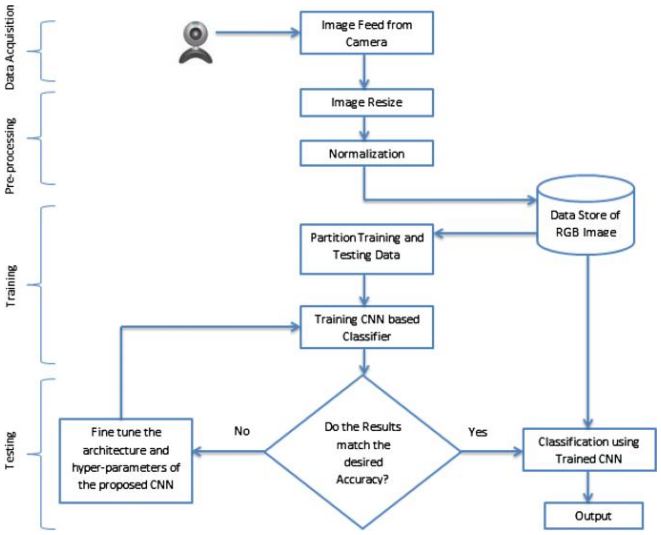

Figure 1: Working Architecture

8.APPLICATION AND FUTURE WORKS

8.1 Application

1) The dataset is easily expandable and adaptable to the needs of the user, and it may prove to be a significant step in closing the communication gap for dumb and deaf people.

2) The worth of the disabled people's labour can be acknowledged at worldwide gatherings by using the sign detection model, which makes them easier to interpret.

3) The paradigm is accessible to everyone and may be utilised by anyone with a basic understanding of technology.

4) This strategy can be used in primary schools to introduce sign language to children as early as possible.

8.2 Future Scope

1) Application of our concept to other sign languages, such as American or Indian sign language.

2) More practice with a big dataset will help you detect symbols more quickly.

3) Enhancing the expression recognition capability of the model

9.CONCLUSION

A sign language detection system's main objective is to give deaf and hearing people a useful way to communicate through hand gestures. The suggested system will be utilised with a webcam or any other built-in camera that recognises objects by detecting and processing their indicators. We can infer from the model's conclusions that the suggested system can generate trustworthy results when light and intensity are controlled. Additionally, new motions can be easily added, and the model will be more accurate with more photographs taken from different perspectives and frames. As a result, the model may simply be scaled up to a huge size by expanding the dataset. Due to environmental factors like low light levels and an uncontrolled background, the model's detection accuracy is limited. As a result, we'll try to fix these issues and increase the dataset to get more accurate results.

10.REFERENCES

[1] Martin, D. S. (Ed.). (1985). Cognition, Education, and Deafness: Directions for Research and Instruction. Gallaudet University Press. https://doi.org/10.2307/j.ctv2rh28x6

[2] Harshith, C., Shastry, K.R., Ravindran, M., Srikanth, M.V., & Lakshmikhanth, N. (2010). Survey on Various Gesture Recognition Techniques for Interfacing Machines Based on Ambient Intelligence. ArXiv, abs/1012.0084.

[3] Ekbote, J., & Joshi, M.J. (2017). Indian sign language recognition using ANN and SVM classifiers. 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 1-5.

[4] By great learning team ,"Real-Time Object Detection Using TensorFlow", december 25 ,2021 https://www.mygreatlearning.com/blog/object-detection-usingtensorflow/

[5] OpenCV - Wikipedia

[6] Google's Teachable Machine, https://teachablemachine.withgoogle.com/

[7] Sakshi Goyal, Ishita Sharma, Shanu Sharma, 2013, Sign Language Recognition System For Deaf And Dumb People, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 02, Issue 04 (April 2013),

[8] Shirin Tikoo |COMPUTER VISION| Develop Sign Language Translator with Python | Skyfi Labs

https://www.skyfilabs.com/project-ideas/sign-language-translator-using-python"

[9] TensorFlow https://www.tensorflow.org/

[10] Keras https://keras.io/