**HAND GESTURE AND VOICE CONVERSION FOR DEAF AND DUMB USING DEEP LEARNING**

**Abstract:**

Despite a wide social group that could profit from it, the idea of sign language recognition by technology is underutilized. Several technologies are available that can be helpful in establishing a link between this social group and the rest of the world. One of the main tools for enabling sign language users to communicate with the rest of society is understanding sign language. Computers can recognize sign language with the use of image categorization and machine learning, which can then be translated by humans. This study uses convolutional neural networks to identify sign language motions. The static sign language gestures were photographed using an RGB camera and comprise the image dataset used. The photos underwent preprocessing before being used as the input that had been cleaned. Inception v3 convolutional neural network model was used to retrain and test this dataset of sign language motions, and the findings are presented in this study. Several convolution filter inputs are processed on a single input in the model.

The resulting validation accuracy was greater than 90%. The numerous attempts at sign language detection using machine learning and image depth data are also reviewed in this work. It assesses the different difficulties involved in solving the issue at hand and also describes the problem's potential future.

**Key words:** Sign Language Recognition, Deep Learning, Inception V3, Image Processing.

**Introduction:**

A pattern recognition issue can be framed as comprehending human gestures. The required message can be conveyed if a computer can recognize and distinguish these human motion patterns reconstructed. Successful detection of static sign gestures for letters and numbers has been made. But, this system can also be expanded to recognize words and sentences. Here, American Sign Language (ASL) has been employed as the sign language whose gestures we are trying to recognize. This is also where many other sign languages have their roots. The bulk of people communicate with one another through spoken language. For a huge section of the population, it would not be conceivable without spoken language to exchange ideas. Nonetheless, despite the use of spoken language, certain people are unable to communicate with the majority of people. Those who are mute are unable to converse verbally. This segment of the community benefits from the use of sign language. Sign language provides the same instruments for communication as spoken language does, including facial emotions, static hand symbols, and hand gestures. There are numerous varieties of sign languages, just like there are in spoken languages. They have their roots in dialect and geography, just like spoken language. Polish Sign Language, American Sign Language, Indian Sign Language, etc. are some examples. It does have disadvantages of its own as a result of these variances. First of all, like all It is only appropriate for use among speakers of the language. Speech-impaired sign language users are unable to communicate with the general public, which is only fluent in spoken language. The distance between the two must be filled in order to enable greater and more effective communication. Human-computer interface is another challenge this group of people must overcome, particularly since sign language is their exclusive form of communication. There are between 250,000 and 500,000 ASL users. This is small compared to the size of the population. For the purpose of teaching and comprehending sign language, technology in the form of various software packages has been created.

Nonetheless, there has been good but limited progress in the use of contemporary technology for sign language recognition. A program that can effectively recognize and translate sign language is now required. Most importantly, it ought to serve as a link between sign language users and individuals who lack any immediate incentive to learn or comprehend the language. Our research makes a contribution to this endeavor by testing one such methodology in order to determine how well it recognizes sign language. In our experimentation, we have taken this into account and used American Sign Language (ASL).

**Existing system:**

Existing systems for sign language detection using deep learning. There are many other systems that use different deep learning architectures and techniques to recognize sign language gestures. This system uses a CNN to recognize sign language gestures. The CNN is trained on a large dataset of sign language gestures to learn to recognize the gestures.

**Disadvantages:**

* More Expensive.
* Difficult to scale up.
* Time consuming.

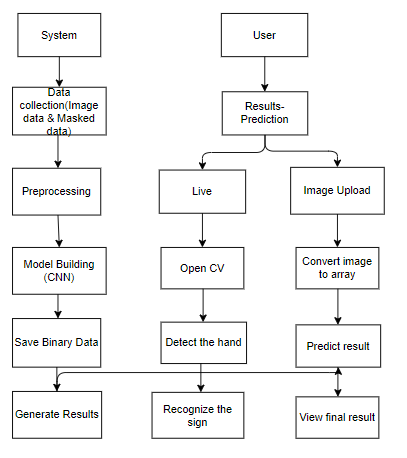
**Proposed system:**

One potential enhancement to this proposed system could be to incorporate real-time sign language detection using a video stream. This would involve modifying the feature extraction and model training steps to accommodate real-time processing, as well as implementing a mechanism for handling video input and output. Another enhancement could be to incorporate multiple modalities, such as audio and facial expressions, to improve the accuracy of sign language detection.

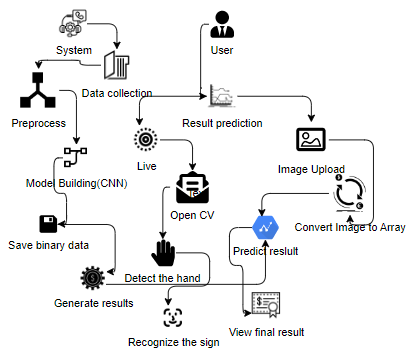
**Advantages:**

* Cheaper to operate.
* It can be scaled up quickly.
* Time minimising.

**Block Diagram:**

****

**Architecture:**

****

**SYSTEM SPECIFICATIONS:**

# H/W Specifications:

# Processor : I5/Intel Processor

# RAM : 8GB (min)

* Hard Disk : 128 GB

**S/W Specifications:**

* Operating System : Windows 10
* Server-side Script : Python 3.6
* IDE : PyCharm, Jupyter notebook
* Libraries Used : Numpy, IO, OS, Flask, Keras, pandas, tensorflow, Segmentation

**Modules:**

**System:**

**Data collection:** First, System will collect the user data through image or masked data.

**Preprocessing:** System will preprocess the data which was collected by the system.

**Model Building:** CNN algorithm is used in this project.

**Save Binary data:** The system will save the binary data.

**Generate results:** Finally, it will generate the results.

**User:**

**Result prediction:** User, will predict the result which was saved by the system.

**Live**: The predicted result will show the live result to the user.

**Image upload:** The user will upload the image and the system will show the exact result.

**Open CV:** Here, the cam will open to detect the Gesture.

**Recognize the sign:** The cam will recognize the sign of the hand.