**A MAJOR PROJECT REPORT ON**

**SIGN LANGUAGE TO TEXT AND SPEECH TRANSLATION IN REAL TIME USING CNN**

Submitted to Jawaharlal Nehru Technology University, Hyderabad,

Is partial fulfilment of requirement for the award of the degree of

**BACHELOR OF TECHNOLOGY**

**IN**

**INFORMATION TECHNOLOGY**

**SUBMITTED BY**

**MUDUGULA AASHISH MUDIRAJ (18J21A1233)**

**KARKA BHARGAVI** **(18J21A1219)**

​ **SAHIL MANDAVIA (18J21A1243)**

**Under the guidance of**

**Mr. P.THIMMA REDDY**

Associate Professor



Department of Information Technology

**JOGINPALLY B.R. ENGINEERING COLLEGE**

Accredited by NAAC with B++ Grade, Recognized under Sec. 2(f) of UGC Act. 1956

Approved by AICTE and Affiliated to Jawaharlal Nehru Technological University,

Hyderabad

Bhaskar Nagar, Yenkapally, Moinabad,

RangaReddy, Hyderabad, Telangana - 500075.

**2018-2022**

**JOGINPALLY B.R ENGINEERING COLLEGE**

Accredited by NAAC with B++ Grade, Recognized under Sec. 2(f) of UGC Act. 1956

Approved by AICTE and Affiliated to Jawaharlal Nehru Technological University, Hyderabad

Bhaskar Nagar, Yenkapally, Moinabad,

RangaReddy, Hyderabad, Telangana - 500075.



**CERTIFICATE**

This is to certify that the project entitled “**SIGN LANGUAGE TO TEXT AND SPEECH TRANSLATION IN REAL TIME USING CNN**“ is the bonafide work carried out by **MUDUGULA AASHISH MUDIRAJ** **(18J21A1233), KARKA BHARGAVI (18J21A1219), SAHIL MANDAVIA (18J21A1243)** of B.Tech **(INFORMATION TECHNOLOGY)** under our guidance and supervision. The Project Report is submitted to JNTU Hyderabad in partial fulfillment of requirements of the award of B.Tech degree during the academic year 2021-2022.

**INTERNAL GUIDE HEAD OF THE DEPARTMENT**

Mr. P.THIMMA REDDY Mr. P.SRINIVAS

Associate ProfessorAssociate Professor

**External Examiner**

**ABSTRACT**

Creating a desktop application that uses a computer’s webcam to capture a person signing gestures for American Sign Language (ASL), and translate it into corresponding text and speech in real time. The translated sign language gesture will be acquired in text which is farther converted into audio. In this manner we are implementing a finger spelling sign language translator. To enable the detection of gestures, we are making use of a Convolutional neural network (CNN). A CNN is highly efficient in tackling computer vision problems and is capable of detecting the desired features with a high degree of accuracy upon sufficient training.

1. **INTRODUCTION**

American Sign Language (ASL) is natural syntax that has the same etymological homes as being speaking languages, having completely different grammar, ASL can be express with destiny of actions of the body. In native America, people who are deaf or cant see, its a reliable source of absurdity. There is not any formal or familiar form of sign language. Different signal languages are speculating in particular areas. For a case, British Sign Language (BSL) is an entirely different language from an ASL, and USA people who familiarise with ASL would not easily understand BSL. Some nations adopt capabilities of ASL of their sign languages. Sign language is a way of verbal exchange via human beings diminished by speech and listening to loss. Around 360 million human beings globally be afflicted via unable to hearing loss out of which 328000000 are adults and 32000000 children. hearing impairment extra than 40 decibels in the better listening to ear is referred as disabling listening to loss. Thus, with growing range of people with deafness, there is moreover a rise in demand for translators. Minimizing the verbal exchange gap among listening to impaired and regular humans turns into a want to make certain effective conversation among all. Sign language translation is one of the amongst most growing line of research nowadays and its miles the maximum natural manner of communication for the humans with hearing impairments. A hand gesture recognition gadget can offer an opportunity for deaf people to talk with vocal humans without the need of an interpreter. The system is built for the automated conversion of ASL into textual content and speech.

1. **EXISTING SYSTEM**

American Sign Language (ASL) is natural syntax that has the same etymological homes as being speaking languages, having completely different grammar; ASL can be express with destiny of actions of the body. In native America, people who are deaf or can’t see, it’s a reliable source of absurdity. There is not any formal or familiar form of sign language. Different signal languages are speculating in particular areas. For a case, British Sign Language (BSL) is an entirely different language from an ASL, and USA people who familiarize with ASL would not easily understand BSL. Some nations adopt capabilities of ASL of their sign languages. Sign language is a way of verbal exchange via human beings diminished by speech and listening to loss.

**3.1 Disadvantage of existing system**

1. Less efficiency.

2. Utilizes Complex Algorithms

3. Multiple Sensors

4. Recognizes Hand gestures

**3. PROPOSED SYSTEM**

In this manner we are implementing a finger spelling sign language translator. To enable the detection of gestures, we are making use of a Convolutional neural network (CNN). A CNN is highly efficient in tackling computer vision problems and is capable of detecting the desired features with a high degree of accuracy upon sufficient training.

**4.1 Advantages of Proposed system**

1. More efficiency.

2. Eliminates the need for an interpreter for communication between sign language and speech language

3. Real time translation

4. Does not require any additional hardware

**4. MODULES**

**1.Image Acquisition**

The gestures are captured through the web camera. This OpenCV video stream is used to capture the entire signing duration. The frames are extracted from the stream and are processed as grayscale images with the dimension of 50\*50. This dimension is consistent throughout the project as the entire dataset is sized exactly the same.

**2.Hand Region Segmentation & Hand Detection and Tracking**

The captured images are scanned for hand gestures. This is a part of preprocessing before the image is fed to the model to obtain the prediction. The segments containing gestures are made more pronounced. This increases the chances of prediction by many folds.

**3.Hand Posture Recognition**

The preprocessed images are fed to the keras CNN model. The model that has already been trained generates the predicted label. All the gesture labels are assigned with a probability. The label with the highest probability is treated to be the predicted label.

**4.Display as Text & Speech**

The model accumulates the recognized gesture to words. The recognized words are converted into the corresponding speech using the pyttsx3 library. The text to speech result is a simple work around but is an invaluable feature as it gives a feel of an actual verbal conversation.

**5. SOFTWARE REQUIREMENT SPECIFICATION**

**SOFTWARE REQUIRMENTS :**

* Operating System **:** Microsoft Windows 10
* Programming Language **:** Python
* Python IDE **:** Spyder
* Back End Tools **:** Python 3.6

**HARDWARE REQUIRMENTS :**

* Processor with speed  **:** Intel [i3] (2.0 GHz)
* Hard Disk **:**  40 GB
* RAM  **:** 4 GB