**PROJECT**

**ON**

**SIGN DETECTION AND RECOGNITION**

****

**SRI BALAJI UNIVERSITY**

**SCHOOL OF COMPUTER STUDIES**

**Program: Masters of Computer Application 2022 Batch**

|  |  |
| --- | --- |
|  |  |
| 1.Aditya Patel (MCA2302114)  2.Khushi Sahita (MCA2302073)  3.Shivangi Bhosle(MCA2302067)  4.Tejas Patil (MCA2302080)  **Class**-- MCA(A) SEM-I |  |
|  |  |

**Submitted by:**

**CERTIFICATE**

Certified that the Project Report entitled **“SIGN DETECTION AND RECOGNITION”**, submitted by **Aditya Patel, Khushi Sahita, Shivangi Bhosle, Tejas Patil** of **MCA**, is their own work and has been carried out under my supervision. It is recommended that the candidates may now be evaluated for their work by the University.

|  |  |
| --- | --- |
| **Name 1 : Aditya Patel** | **Roll no. :** MCA2302114 |
| **Name 2 : Khushi Sahita** | **Roll no. :** MCA2302073 |
| **Name3 : Shivangi Bhosle** | **Roll no. :** MCA2302067 |
| **Name 4 : Tejas Patil** | **Roll no. :** MCA2302080 |

|  |
| --- |
| **Date : 30/11/20** |
|  |

Prof. Prachi Soni

(Assistant Professor)

**ACKNOWLEDGEMENT**

We feel immense pleasure to introduce “**SIGN DETECTION AND RECOGNITION**” as our major project.

I express my sincere thanks to our instructor **Dr. Richa Purohit** and **Asst. Prof. Prachi Soni** who guided us to the successful completion of this project report. We take this opportunity to express our deep sense of gratitude for their individual guidance, constant encouragement and immense motivation which have sustained our efforts at all stages of this project report. We are grateful and appreciate all the staff members of the School of Computer Application for their cooperation and support.

I extend my sincere thanks to our principal **Dr.G.Y.Shitole** for his support and for all the facilities provided for the preparation of this project report.

Also, we wish to thank our parents & friends who helped us a lot in collecting data, pictures and continuous help and support.

Finally, we would wish to thank everyone involved in this project time.

**ABSTRACT**

A real time sign language detector is a significant step forward in improving communication between the deaf and the general population. We are pleased to showcase the creation and implementation of sign language recognition model based on a Convolutional Neural Network(CNN).We utilized a Pre-Trained SSD Mobile net V2 architecture trained on our own dataset in order to apply Transfer learning to the task. We developed a robust model that consistently classifies Sign language in majority of cases. Additionally, this strategy will be extremely beneficial to sign language learners in terms of practising sign language. Various human-computer interface methodologies for posture recognition were explored and assessed during the project. A series of image processing techniques with Human movement classification was identified as the best approach. The system is able to recognize selected Sign Language signs with the accuracy of 70-80% without a controlled background with small light.

**TABLE OF CONTENT**

**Chapter 1: INTRODUCTION TO THE LANGUAGES USED Page.No**

* WHY PYTHON? …………………………………………………………………………………………..6
* INTRODUCTION TO CNN, TENSORFLOW………………………………………………………6
* OpenCV………………………………………………………………………………………………..……..6
* Keras….……………………………………………………………………………………………..…….….7

**Chapter 2: INTRODUCTION TO “SIGN DETECTION & RECOGNITION”**

* INTRODUCTION…………………………………………………………………………………………..8
* PURPOSE & NEED..………………………………………………………………………………………8
* ADVANTAGES………………………………………………………………………………………………8
* SRS………………………………………………………………………………………………………………10
* Use
* Purpose
* Functional Requirements
* Non-Functional Requirements
* Software Tools
* Deployment
* Hardware Specification
* FEASIBILITY STUDY……………………………………………………………………………………..11

**Chapter 1: INTRODUCTION TO THE LANGUAGES USED**

* **WHY PYTHON?**

Python is a popular programming language for sign detection and recognition due to its simplicity, ease of use, and availability of libraries such as **OpenCV** and **Keras --** These libraries provide a wide range of tools for image processing, machine learning, and deep learning, which are essential for sign detection and recognition. Python’s syntax is also easy to read and write, making it an ideal choice for developers who want to create sign detection and recognition systems quickly and efficiently.

For instance, in a sign language recognition project, **OpenCV** and **Keras** modules of Python can be used to create a sign detector that detects numbers from 1 to 10 and can be extended to cover a vast multitude of other signs and hand gestures including the alphabets.

* **INTRODUCTION TO CNN & TENSORFLOW :**

**Convolutional Neural Networks (CNN):** CNNs are a type of neural network that are particularly well- suited for image classification tasks. They are designed to automatically learn and extract features from images, making them ideal for tasks such as sign detection and recognition

**TensorFlow:** TensorFlow is an open-source machine learning framework developed by Google. It is an open-source artificial intelligence package that builds models using data flow graphs. It enables developers to build large-scale neural networks with several layers. TensorFlow is mostly used for classification, perception, comprehension, discovery, prediction, and creation. It is used for a wide range of machine learning tasks, including image classification, natural language processing, and more. TensorFlow provides a high-level API for building and training machine learning models, making it easy to get started with machine learning .

* **OpenCV :**

One can process images and videos to identify objects, faces, or even handwriting of a human. When it integrated with various libraries, such as NumPy, python is capable of processing the OpenCV array structure for analysis. To Identify image pattern and its various features we use vector space and perform mathematical operations on these features. When OpenCV was designed the main focus was real-time applications for computational efficiency. All things are written in optimized C/C++ to take advantage of multi-core processing.

**OpenCV Functionality :**

* Image/video I/O, processing, display (core, imgproc, highgui)
* Object/feature detection (objdetect, features2d, nonfree)
* Geometry-based monocular or stereo computer vision (calib3d, stitching, videostab)
* Computational photography (photo, video, superres)
* Machine learning & clustering (ml, flann)
* **Keras:**

Keras is an open-source deep learning library written in Python. It is designed to be user-friendly, modular, and extensible, making it a popular choice for both beginners and experienced researchers in the field of artificial intelligence and machine learning. Keras simplifies the development and deployment of deep learning models by offering a high-level abstraction while maintaining flexibility and performance, making it a popular choice among deep learning practitioners and researchers.

**Chapter 2: INTRODUCTION TO “SIGN DEDTECTION AND RECOGNITION”**

* **Introduction:**

**Sign Language** is mainly used by deaf (hard hearing) and dumb people to exchange information between their own community and with other people. It is a language where people use their hand gestures to communicate as they can’t speak or hear.

**Sign Language Recognition** **(SLR)** deals with recognizing the hand gestures acquisition and continues till text or speech is generated for corresponding hand gestures. Sign language uses visual hand and body gestures to convey meaningful messages. Using, Deep Learning algorithms and Image Processing we can able to classify these hand gestures and able to produce corresponding text. An example of “A” alphabet in sign language notion to English “A” text or speech.

On the website “**SIGN DETECTION AND RECOGNITION**”, you can apply for a loan in just 5 easy steps and locate the best loans available to you with the lowest interest rates.

This is divided into 6 parts:

* Image Collection.
* Image Preprocessing.
* Pre-training using CNN model.
* Training and testing the model.
* Evaluation.
* Predicting the hand gestures.
* Implementation of these steps.
* **Purpose:**

The main objectives of this project are to contribute to the field of automatic sign language recognition and translation to text or speech. In our project, we focus on static sign language hand gestures.

* This work focused on recognizing the hand gestures which includes **26 English alphabets** (A-Z) and **10 digits** (0-9) using Deep **Neural Networks** (DNN).
* We created a convolution neural networks classifier that can classify the hand gestures into English alphabets and digits.
* We used the horizontal voting ensemble technique to achieve the maximum accuracy of the model.

1.Hand gestures

2.Sign language recognition.

3.Convolution neural networks.

4.Computer vision.

* **Need:**

The future of sign language recognition systems is intertwined with advancements in technology and a growing awareness of the importance of accessibility and inclusion. As these systems continue to evolve, they have the potential to transform the way deaf and hard-of-hearing individuals communicate and interact with the world.

* A Model can be trained to recognize the gestures from a live video feed in real time
* As the Model converts the image to text, a text to speech model can be used as if the speaker speaks.
* Enhance the recognition capabilities for various lighting conditions.
* Implementing & identifying more number of gestures.
* Provide editing mechanism by using gestures.
* **Advantages:**

Advantages of building a Sign Language Recognition system includes:

* **Gesture Detection**: It can detect and interpret sign language gestures made by users.
* **Video Input**: Typically, it takes video or image input from cameras or sensors to analyse gestures.
* **Real-time Interaction**: It allows real-time communication between sign language users and non-signers.

* **Inclusivity:** It facilitates communication for deaf and hard-of-hearing individuals, promoting inclusivity.
* **Bridging the Communication Gap**: Bridges the gap between sign language users and those who don't understand sign language.
* **Educational Tool**: It can be used as an educational tool to teach sign language to non-signers.
* **Multimodal Interaction**: Can be integrated with other technologies, such as speech recognition, for more versatile communication.
* **SRS- SOFTWARE REQUIREMENT SPECIFICATION**

A software requirements specification (SRS) is a document that describes what the software will do and how it will be expected to perform. It also describes the functionality the product needs to fulfil all stakeholders (business, users) needs.

1. **Use:**

An SRS establishes the [Sign language recognition is an area of research that aims to develop algorithms that can understand and interpret sign language, enabling people who use sign language as their primary mode of communication to communicate more easily with non-signers](https://paperswithcode.com/task/sign-language-recognition) . An SRS provides a reference for validation of the final product/software.

1. **Purpose:**

[The purpose of sign detection and language is to recognize hand gestures and movements used in sign language and translate them into written or spoken language](https://link.springer.com/article/10.1007/s10639-022-11391-z).

1. **Functional Requirements:**

A real-time sign language detector based on a Convolutional Neural Network (CNN) can be developed to recognize selected sign language signs with an accuracy of 70-80% without a controlled background with small light.

The system can be integrated into a small portable device so that deaf-mute people can easily take the device anywhere

The functional requirements of a sign language detection system can be broadly classified into the following categories: -

* Real-time detection: The system should be able to detect signs in real-time.
* Accuracy: The system should be able to detect signs with high accuracy.
* Robustness: The system should be able to detect signs in different lighting conditions, with different backgrounds, and with different hand orientations.
* Portability: The system should be portable and easy to use.
* Cost-effectiveness: The system should be cost-effective.

1. **Non- Functional Requirements:**

To assess the performance of a system the following are the parameters:

1. **Response Time-** The response time must be minimum.

2. **Workload -** Workload is sure little heavy but compared to CNN based model its more efficient. Compared to CNN model which would require 40-50 million parameters.

3. **Scalability -** Highly scalable using ML-based cloud services like TensorFlow, AWS-ML, Google cloud- ML.

4. **Platform –**

* No OS bound.
* CPU: Core i5 10gen or Higher

1. **Software Tools**

* You can use any software to run this project like google chrome, Microsoft edge, Mozilla Firefox, etc. We are using google chrome.
* We are working with Python for the front-end.
* For back-end support we are using Python.

**6. Deployment**

Operating system server: Windows 11.

* **FEASIBILITY STUDY**
* Economical Feasibility
* Technical Feasibility
* Operational Feasibility
* Legal Feasibility
* Scheduling Feasibility
* **Economical Feasibility:**
* Economical feasibility refers to the assessment of whether a project or endeavour is financially viable and justifiable.
* The feasibility of using Sign Language depends on the context and the resources available. In some cases, it may be more economical to use SL than other forms of communication, such as hiring an interpreter or using text-based communication. For example, in India, where there is a shortage of qualified interpreters, researchers are developing apps that can interpret Indian Sign Language.
* Overall, the system costs under **$35** (approx. Rs.3,000/-) which is lower when compared to similar systems investigated, that cost at least **$100 (**approx. Rs.8,500/-). The final system meets the initial goals of sign language to speech translation, portability and cost-effectiveness.
* **Technical Feasibility:**

The proposed project or system can be developed, implemented, and operated using the available technology: -

1. **Convolutional Neural Networks (CNN):** CNNs are a type of neural network that are particularly well- suited for image classification tasks. They are designed to automatically learn and extract features from images, making them ideal for tasks such as sign detection and recognition
2. **TensorFlow**: It is mostly used for classification, perception, comprehension, discovery, prediction, and creation. It is used for a wide range of machine learning tasks, including image classification, natural language processing, and more., resources, and expertise within a given timeframe and budget.
3. **OpenCV**: One can process images and videos to identify objects, faces, or even handwriting of a human. When it integrated with various libraries, such as NumPy, python is capable of processing the OpenCV array structure for analysis.
4. **Keras**: Keras is an open-source deep learning library written in Python. It is designed to be user-friendly, modular, and extensible, making it a popular choice for both beginners and experienced researchers in the field of artificial intelligence and machine learning. The technical feasibility of Keras as a library primarily revolves around its functionality, performance, ease of use, compatibility with different hardware and software configurations, and integration with other tools and frameworks.

In Technical Feasibility current resources both hardware software along with required technology are analysed/assessed to develop project. The project can run on any operating system, and can also be operated on any low-end systems.

* **Operational feasibility:**
* This evaluation considers the practical aspects and impact of the project on day-to-day operations, processes, and people within an organization or system.
* The project tells how much easy product will be to operate and maintenance as we use python language because it includes many **PYTHON LIBRARIES**. The project tells us how the sign language can be understood by the normal people and vice versa for the deaf and blind people. The system is easy to handle by the customers.
* The system measure of how well a proposed system solves the problem, as sign language is not understood by the normal people and vice versa the deaf people, so for the same we need and **INTERPRETER** who can convey the message between the both. The average amount charge by interpreter is $5 to $18 and in rupees its amount is 1500 to 4000/. Our system in that scenario is free of cost where we don’t need help of interpreter as the application what have been prepare solve the problem itself.
* **Legal Feasibility:**

Legal feasibility refers to the assessment of whether a proposed project or business venture complies with the laws, regulations, and legal requirements imposed by the relevant authorities. The legal recognition of sign languages varies widely across the world. In some countries, signed languages are recognized as official languages, while in others, they have a protected status in certain areas such as education. The **World Federation of the Deaf** considers the lack of meaningful sign language legislation on the national level a grave violation of deaf peoples’ fundamental rights. The **WFD’s** 2020-2030 Strategic Direction calls for further promotion of the legal recognition of national sign languages.

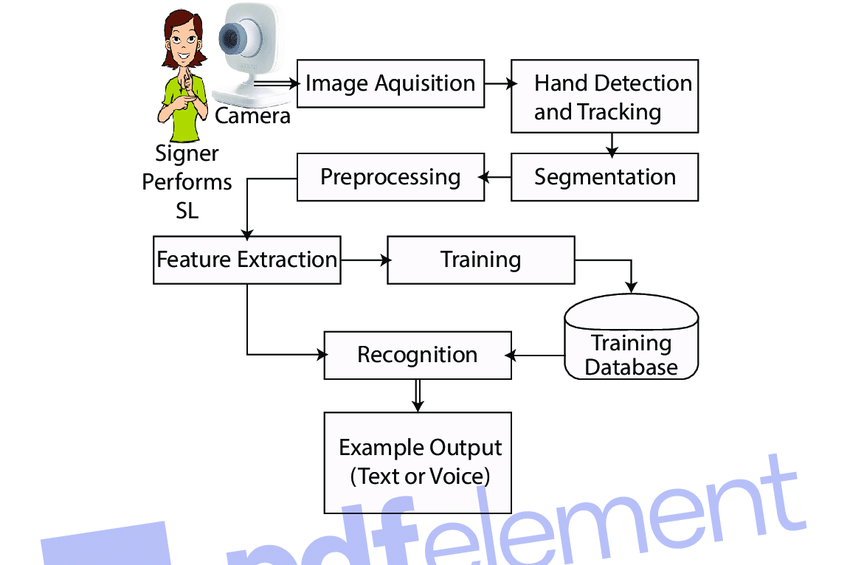
* **Scheduling Feasibility:**

Scheduling feasibility refers to the evaluation of whether a proposed project or initiative can be completed within a reasonable and realistic timeframe. It involves assessing the timeline, milestones, and overall schedule of the project to ensure it can be accomplished within the available time constraints and deadlines.

**This project for Sign Language Recognition: -**

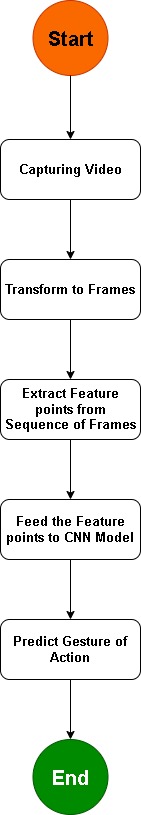
* Will require **approx. 10 months** for the completion of whole project.
* Firstly, we have to gather the data for the hand gesture which will be require lot of time for the collection of data nearly about a month or more.
* After the data collection, the implementation of the data on the real time project will be done.
* And the rest time will be used for the coding purpose, for building the project using CNN, OpenCV and Tensorflow.

**DATA FLOW DIAGRAM**



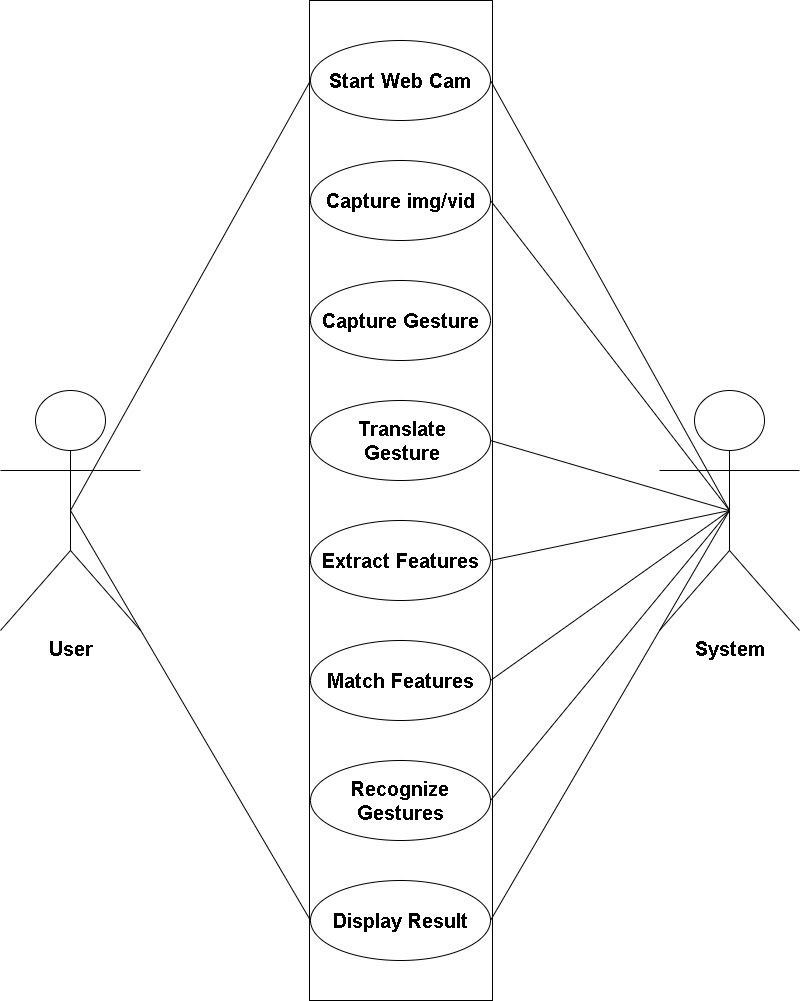
**Fig. 1: Data flow diagram**

**Flowchart Diagram**

****

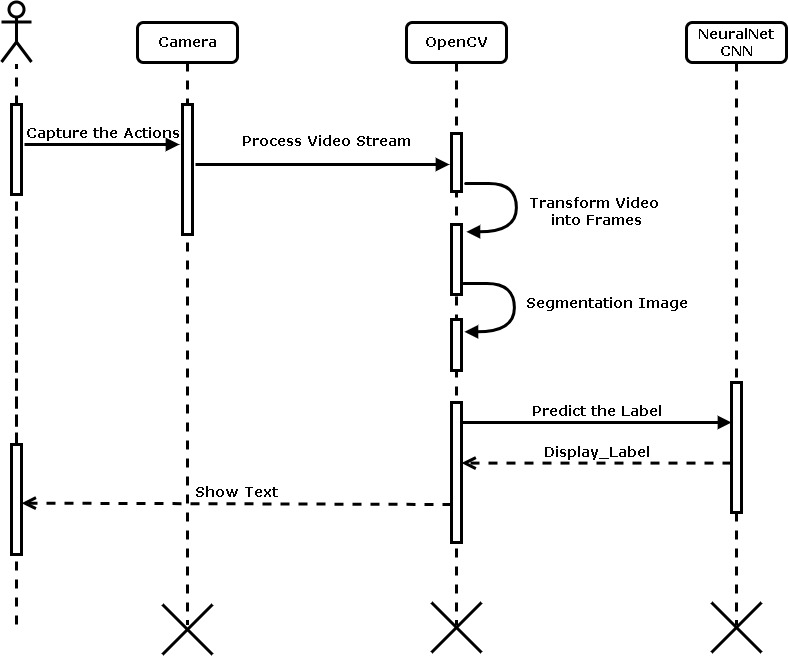
**Fig. 3: Flowchart Diagram**

**Use-Case Diagram**



**Fig. 4: Use Case Diagram**

**UML Diagram**

****

**Fig. 5: UML Diagram**

**Methodology**

**Hand Detection**

**and Tracking**

**Hand Region**

**Segmentation**

**Hand Posture**

**Recognition**

**Classified Gesture**

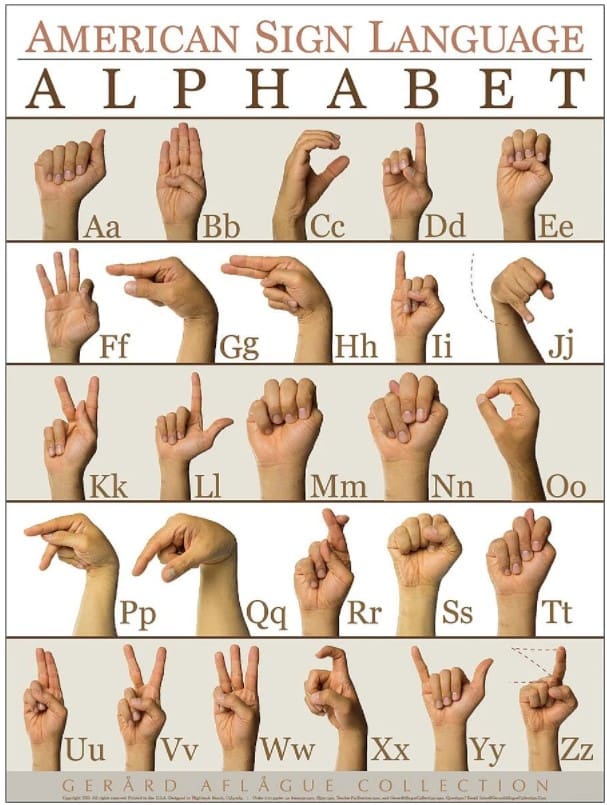
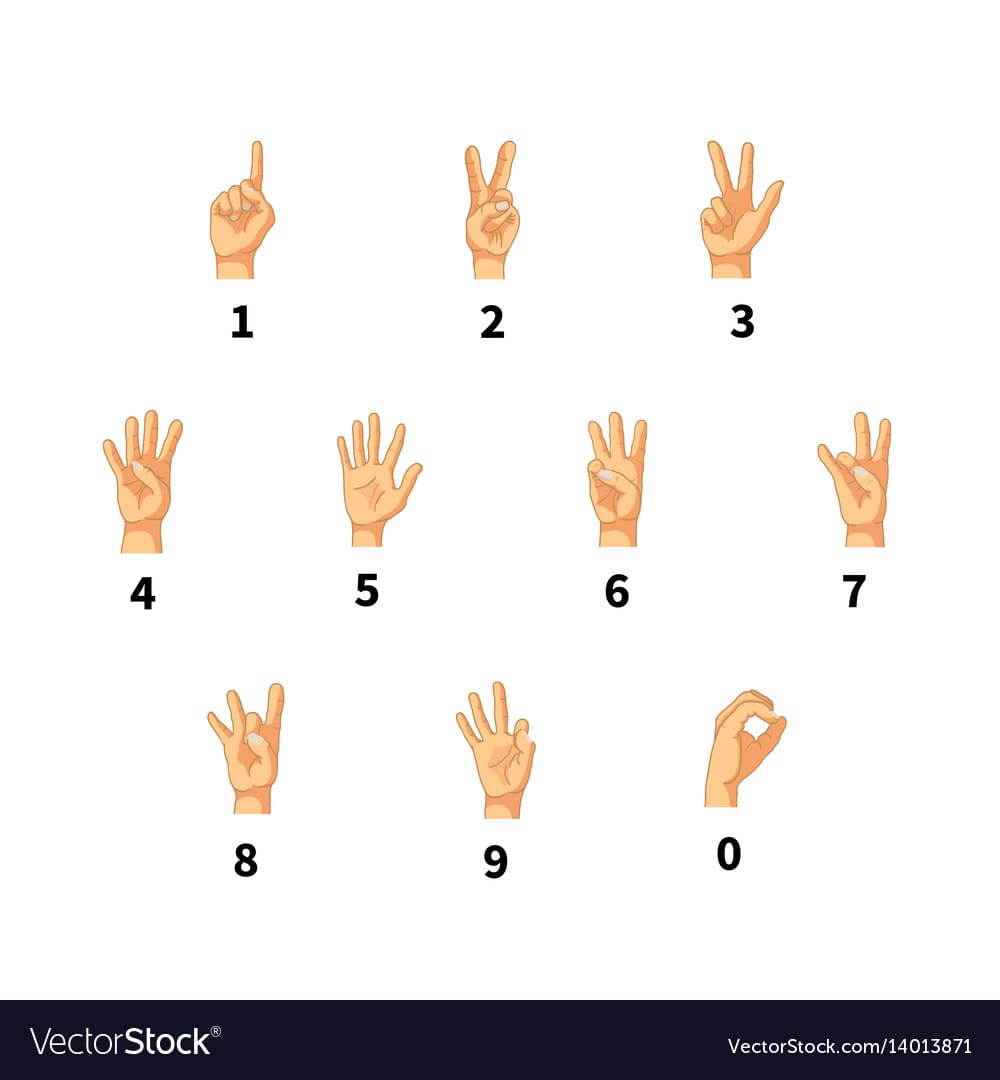
**Display as Text or**

**Voice**

**Image Acquisition**

**Form Camera**

* **Image Acquisition Form Camera-** An Image Acquisition Form Camera typically refers to a software or hardware system designed to capture images from a camera device and present them in a user-friendly interface for further processing or analysis.
* **Hand Region Segmentation-** Hand region segmentation refers to the process of isolating or extracting the region of an image or video frame that contains a person's hand. This segmentation is a crucial step in many computer vision applications, especially those related to gesture recognition, sign language interpretation, human-computer interaction, and augmented reality.
* **Hand Detection and Tracking-** Hand detection and tracking are fundamental techniques in computer vision and robotics that involve identifying and following the movement of human hands in images or videos. These techniques are crucial for applications such as gesture recognition, human-computer interaction, augmented reality, and robotics control.
* **Hand Posture Recognition-** Hand posture recognition, also known as hand gesture recognition, is the process of identifying and interpreting the configuration or shape of a hand in an image or video frame. This recognition is used in various applications, including human-computer interaction, virtual reality, robotics, sign language interpretation, and gesture-based control systems.
* **Classified Gesture-** A classified gesture refers to a specific hand movement or configuration that has been categorized and labeled according to its meaning or intended action. In the context of gesture recognition systems, classified gestures are predefined gestures that are associated with specific commands, functions, or actions within an application or system.
* **Display as Text or Voice-** Displaying as text or voice refers to presenting information or content in either textual form (text) or through spoken words (voice). This approach is commonly used in user interfaces, communication systems, accessibility features, and interactive applications to provide information to users in a format that suits their preferences or needs.

****

**Fig. 6: Dataset Samples**

**Dataset Samples**