**Software Requirements**

**Specification**

**for**

**Sign Language Recognition System**

**Prepared by**

**Sanskriti Bajpai [2000290120135]**

**Vidhi Pandey [2000290120188]**

**Vishakha Rana [2000290120192]**

**Under The Guidance of**

**Mr. RAJ KUMAR**

**(ASSISTANT PROFESSOR)**

**DEPARTMENT OF COMPUTER SCIENCE**

**KIET GROUP OF INSTITUTIONS, GHAZIABAD**

**APR-2023**

# Table of Contents

## Table of Contents ..................................................................................................................... 2 1. Introduction ........................................................................................................................ 3 1.1 Purpose ......................................................................................................................................3

1.2 Document Conventions .......................................................................................................... 4

1.3 Product Scope ........................................................................................................................ 5

1.4 References .............................................................................................................................. 5

1. **Overall Description ............................................................................................................ 7** 2.1 Product Perspective ................................................................................................................7
   1. Product Functions .................................................................................................................7
   2. User Classes and Characteristics ...........................................................................................7
   3. Design and Implementation Constraints ...............................................................................8
   4. Assumptions and Dependencies ............................................................................................8
2. **External Interface Requirements ..................................................................................... 9** 3.1 User Interfaces ....................................................................................................................... 9
   1. Hardware Interfaces ..............................................................................................................9
   2. Software Interfaces ...............................................................................................................9
3. **System Features ..............................................................................................................10**

4.1 System Feature ..................................................................................................................... 10

## 5. Other Non-functional Requirements ...............................................................................11

5.1 Performance Requirements .................................................................................................. 11

5.2 Security Requirements ......................................................................................................... 11

5.3 Software Quality Attributes ................................................................................................. 12 **Appendix A: Glossary ..........................................................................................................13**

**Appendix B: Analysis Models ............................................................................................... 15** a. Use Case Diagram.................................................................................................................. 15

1. Data Flow Diagram ................................................................................................................ 16
2. State Diagram ..........................................................................................................................17
3. Sequence Diagram ..................................................................................................................18
4. Class Diagram .........................................................................................................................19

# 1. Introduction

Verbal communication performed by humans is one of the most unique traits in the entire animal kingdom. Humans have used communication as a tool to share and expand our knowledge of the world. It is safe to say that humans have created settlements, societies, technologies, strategies and more, only through efficient communication. In today s world, communication between individuals is essential to the development and maintenance of society. But unfortunately, some individuals with hearing/speech disabilities are unable to perform this basic human interaction. This barrier in communication alienates them from society and hinders effective communication. Since it is not feasible to assume that every person who communicates with such disabled individuals knows sign language, we need a method that will eradicate this communication barrier. In this project, we are proposing one such method.

## 1.1 Purpose

The purpose of this document is to specify the features, requirements of the final product and the interface of Sign Language to Speech Convertor. It will explain the scenario of the desired project and necessary steps in order to succeed in the task. To do this throughout the document, overall description of the project, the definition of the problem that this project presents a solution and definitions and abbreviations that are relevant to the project will be provided. The preparation of this SRS will help consider all of the requirements before design begins, and reduce later redesign, recoding, and retesting. If there will be any change in the functional requirements or design constraints part, these changes will be stated by giving reference to this SRS in the following documents.

## 1.2 Document Conventions

" **Feature**: Features are individual measurable property or characteristic of a phenomenon being observed. These required for action recognition.

" **Label**: Labels are the final output. We can also consider the output classes to be the labels.

" **Model**: A machine learning model is a mathematical portrayal of a real-life problem. There are various algorithms that perform different tasks with different levels of accuracy.

" **Classification**: In classification, we will need to categorize data into a finite number of predefined classes.

" **LSTM**: Long Short-Term Memory is a kind of recurrent neural network (RNN) which can retain the information for a long period of time. It is used for processing, predicting, and classifying based on time-series data.

" **Training-set**: This is the data set over which LSTM model is trained. The predictions are completely dependent on the training-data set.

" **Testing-set**: The test dataset is a subset of the training dataset that is utilized to give an objective evaluation of a final model.

" **Categorical Accuracy**: Categorical Accuracy calculates the percentage of predicted values (yPred) that match with actual values (yTrue) for one-hot label

" **MediaPipe Holistic**: MediaPipe is a Framework for building machine learning pipelines for processing time-series data like video, audio, etc. For example -> We feed a stream of images(Hands here) as input which comes out with hand landmarks rendered on the images.

" **TensorFlow:** TensorFlow is an open-source end-to-end platform for creating Machine Learning applications. It is a symbolic math library that uses dataflow and differentiable programming to perform various tasks focused on training and inference of deep neural networks.

" **OpenCV:** OpenCV(Open Source Computer Vision) is an open source library of programming functions used for real-time computer-vision. It is mainly used for image processing, video capture and analysis for features like face and object recognition.

## 1.3 Product Scope

This system is primarily intended for making an Interpreter. This will have applications in Business who want to employ deaf and mute employees can use it to convey employee messages to the end consumer. It will be used majorly by the deaf and mute to communicate.

The applications can further be extended to security purposes, by developing a sign language of your own. And even observing and analyzing any suspicious actions.

Some other applications and scopes of this project are:

" It can be used to provide live captions for the online meetings.

" It can be used to detect mistakes in sign languages.

" It can be used for learning and practicing sign languages.

" Use hand gestures to control and automate other devices.

## 1.4 References

1. Akshay Divkar, Rushikesh Bailkar, Dr. Chhaya S. Pawar, <Gesture Based Real-time Indian Sign Language Interpreter=, International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 7 Issue 3, pp. 387-394, May-June 2021. Available at DOI : https://doi.org/10.32628/CSEIT217374
2. Hema B N., Sania Anjum, Umme Hani, Vanaja P., Akshatha M., =Sign Language and Gesture Recognition for Deaf and Dumb People=, International Research Journal of Engineering and Technology (IRJET) , e-ISSN: 2395-0056 Volume: 06 Issue: 03 | Mar 2019 www.irjet.net p-ISSN: 2395-0072
3. Ss, Shivashankara & S, Dr.Srinath. (2018). American Sign Language Recognition System: An Optimal Approach. International Journal of Image, Graphics and Signal Processing. 10. 10.5815/ijigsp.2018.08.03.
4. Shreyas Viswanathan, Saurabh Pandey, Kartik Sharma, Dr P Vijayakumar, <SIGN LANGUAGE TO SPEECH AND SPEECH CONVERSION USING CNN=, International Research Journal of Modernization in Engineering Technology and Science, e-ISSN: 2582-5208, Volume:03/Issue:05/May-2021, www.irjmets.com
5. Shruty M. Tomar, Dr.Narendra M. Patel, Dr. Darshak G. T., <A Survey on Sign Language Recognition Systems=, International Journal of Creative research Thoughts

(IJCRT) 2021, Volume 9, Issue 3 March 2021 | ISSN: 2320-2882, https://ijcrt.org/IJCRT2103503.pdf

1. Mahesh Kumar N B, <Conversion of sign language into speech=, International Journal of Creative research Thoughts (IJCRT) 2021, ISSN 0973-4562 Volume 13, Number 9 (2018) pp. 7154-7161
2. He Siming, <Research of a Sign Language Translation System Based on Deep Learning=, International Conference on Artificial Intelligence and Advanced

Manufacturing (AIAM), 2019, Publisher: IEEE, DOI:

10.1109/AIAM48774.2019.00083

1. Kothadiya, D.; Bhatt, C.; Sapariya, K.; Patel, K.; Gil-González, A.-B.; Corchado, J.M., <Deepsign: Sign Language Detection and Recognition Using Deep Learning.= Electronics 2022,11,1780. https://doi.org/10.3390/electronics11111780
2. Sakshi Mankar, Kanishka Mohapatra, Ashwin Avate, Mansi Talavadekar, Prof. Surendra Sutar, <Realtime Hand Gesture Recognition using LSTM model and

Conversion into Speech=, March 2022, International Journal of Innovative Research in Technology, Volume 8 Issue 10, ISSN: 2349-6002

# 2. Overall Description

## 2.1 Product Perspective

There’s is a huge communication barrier between Sign language users and the verbal language users. The sign language converter addresses this problem by converting the hand gestures to the English language words through the image processing algorithm. Our project is different from the existing systems because it focuses on the word recognition through gestures while the existing systems focuses on letter recognitions through hand signs which is very slow and makes having an actual conversation quite difficult.

## 2.2 Product Functions

¥ Capturing the gestures made by the sign language user through an image sensor.

¥ Tracking the Gestures through OpenCV by identifying feature points.

¥ Pre-processing the captured data .

¥ Feeding the data to the model.

¥ LSTM Model will process the data provided. ¥ Predicting the word based on processed data.

¥ Selecting the word of highest possibility upto three words.

¥ Displaying the word on the UI or Output area.

## 2.3 User Classes and Characteristics

The project will be useful to the people who have trouble understanding sign language or the people who encounter the usage of sign language in their day-to-day communications.

" People with hearing disability.

" People with mute disability.

" People who don t know sign language

" People who communicate with sign language users.

## 2.4 Design and Implementation Constraints

" Hardware limitation on mobile devices as mobile devices have very limited hardware power.

" Full-fledged translation is not possible because the English language has more than 1,000,000 words.

" For the initial phase, the word that can be translated are limited to 54 words.

" Only one way communication is possible through this project.

" Fast paced conversations are possible as the data captured requires some time to process and predict the words and the hardware is not capable to process that fast.

## 2.5 Assumptions and Dependencies

" It is assumed that the user will have an embedded or external Camera\Image sensor available and installed on the host system or device.

" OpenCV is a dependency of the project.

" MediaPipe is a dependency of the project.

" Python and its numpy library is a dependency of the project.

" It is assumed that user is running the project on capable hardware as described in minimum hardware requirement .

" Tkinter/Kivy/PyQT

" Matplotlib, Scikit Learn.

# External Interface Requirements

## User Interfaces

There will be an output screen where the video stream used for processing will be displayed and on the bottom side of the video display window the predicted words will be displayed. There will be three words displayed which will be arranged in order of high to low possibility in a left to right manner. Word with highest possibility will be highlighted using a coloured outline.

### 3.2 Hardware Interfaces

If there is no embedded camera in the system, then there will be the need of an external camera sensor along with the driver needed to enable the functionality on that specific operating system and the hardware platform.

### 3.3 Software Interfaces

#### OpenCV

OpenCV is used to track the gestures from the input stream and then it is fed to the MediaPipe interface.

#### MediaPipe

It extracts the feature points tracked by OpenCV and the feeds it to the LSTM model for prediction.

#### TensorFlow

TensorFlow is an open source software library for high performance numerical computation. Its flexible architecture allows easy deployment of computation across a variety of platforms

(CPUs, GPUs, TPUs), and from desktops to clusters of servers to mobile and edge devices.

# System Features

## System Feature

The Primary features of this system is to translate the sign language into speech.

¥ Initially, widely used gestures have been tracked to train the system.

¥ The system works at word-level translations.

¥ The captured images need to be pre-processed. The system modified the images captured and trained the LSTM model to classify the signals into labels.

# Other Non-functional Requirements

## Performance Requirements

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

### 1. Response Time

1. **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, LSTM model requires 400-700K parameters.
2. **Scalability** — Highly scalable using ML-based cloud services like TensorFlow, AWSML, Google cloud-ML.

### 4. Platform -

¥ No OS bound.

¥ CPU: Core i5 10gen or Higher

¥ GPU: GeForce GTX 980 or higher

¥ RAM: 8GB or Higher

### 5.2 Security Requirements

Currently our system does all the processing and temporary data-storage on the local device.

" **Confidentiality:** Our System preserves the access control and disclosure restrictions on information. Guarantee that no one will be break the rules of personal privacy and proprietary information;

" **Integrity:** Our system avoids the improper (unauthorized) information modification or destruction.

" **Availability:** All of the private translated speech/conversation stays right within the local application, avoiding any foreign interventions.

### 5.3 Software Quality Attributes

1. **Usability**

Our application is simple to use, and it is user friendly.;

1. **Availability**

Our system is available — holds integrity, dependability, and confidentiality.

#### 3. Functionality

Currently our system is functionally under progress. Our goal being able to translate the word-level sign is still under progress.

# Appendix A: Glossary

|  |  |
| --- | --- |
| **Accuracy** | **:** Accuracy is one metric for evaluating classification models. Informally, accuracy is the fraction of predictions our model got right. |
| **Artificial**  **Intelligence** | **:** Artificial intelligence (AI) refers to the simulation of human intelligence in machines that are programmed to think like humans and mimic their actions. |
| **Cloud-ML** | **:** Cloud ML helps developers to easily build high quality custom machine learning models with limited machine learning expertise needed. |
| **Framework** | **:** ML frameworks are interfaces that allow data scientists and developers to build and deploy machine learning models faster and easier. |
| **Gesture** | **:** A gesture is a movement that you make with a part of your body, especially your hands, to express emotion or information. |
| **Machine**  **Learning** | **:** Machine learning (ML) is a type of artificial intelligence (AI) that allows software applications to become more accurate at predicting outcomes without being explicitly programmed to do so |
| **Model** | **:** A Model is a file that has been trained to recognize certain types of patterns. |
| **NumPy** | **:** NumPy is a Python library used for working with arrays. |
| **OpenCV** | **:** OpenCV (Open-Source Computer Vision Library) is an open-source computer vision and machine learning software library. |
| **Optimal**  **Approach** | **:** An optimal approach is a decision that leads to at least as good a known or expected outcome as all other available decision options. |
| **Pandas** | **:** Pandas is a Python library used for working with data sets. It has functions for analyzing, cleaning, exploring, and manipulating data. |
| **Deep**  **Learning** | **:** a type of machine learning based on artificial neural networks in which multiple layers of processing are used to extract progressively higherlevel features from data. |
| **LSTM** | **:** Long Short-Term Memory is a kind of recurrent neural network (RNN) |

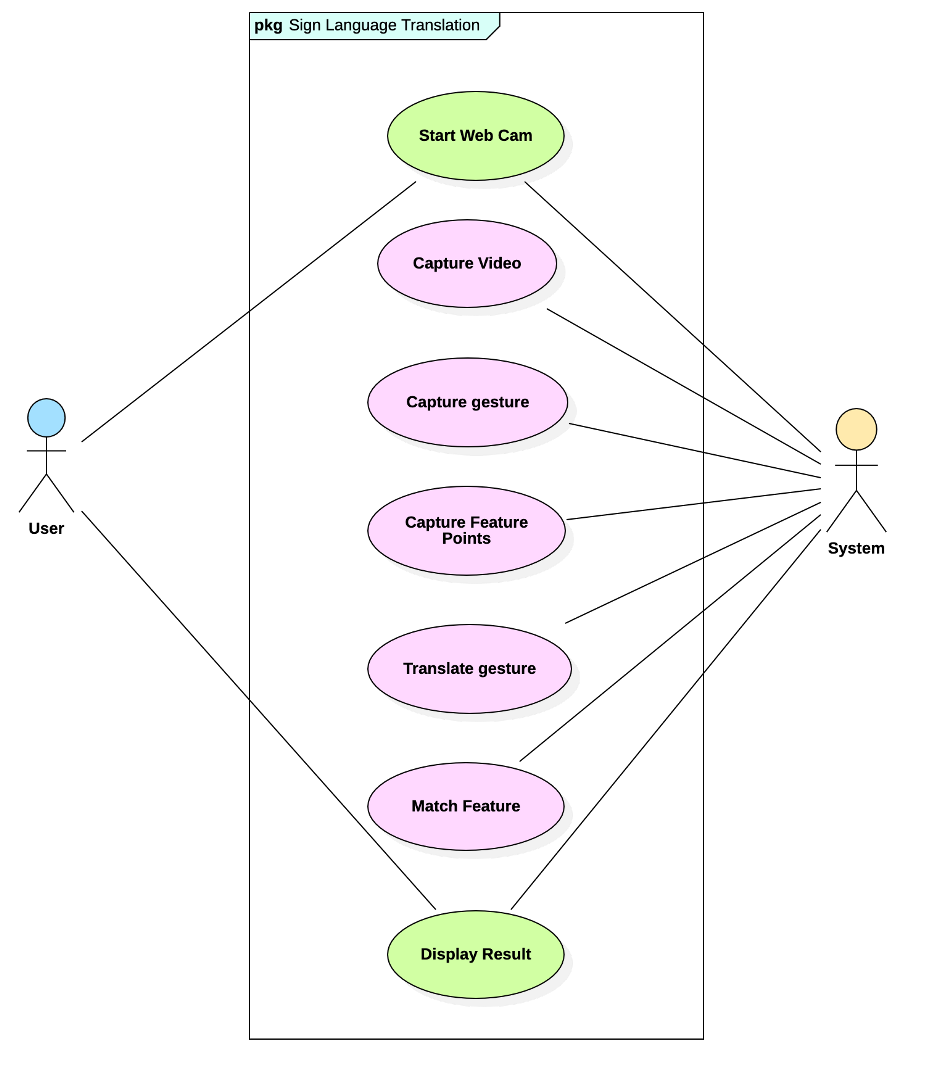
which can retain the information for a long period of time

|  |  |
| --- | --- |
| **Signal**  **Processing** | **:** Signal processing is a broad engineering discipline that is concerned with extracting, manipulating, and storing information embedded in complex signals and images. |
| **Computer**  **Vision** | **:** Computer vision is an interdisciplinary scientific field that deals with how computers can gain high-level understanding from digital images or videos. |
| **System** | **:** A system is a collection of elements or components that are organized for a common purpose. In this case the <Sign Language to Speech Convertor= is a system. |
| **TensorFlow** | **:** TensorFlow is an open-source end-to-end platform for creating |

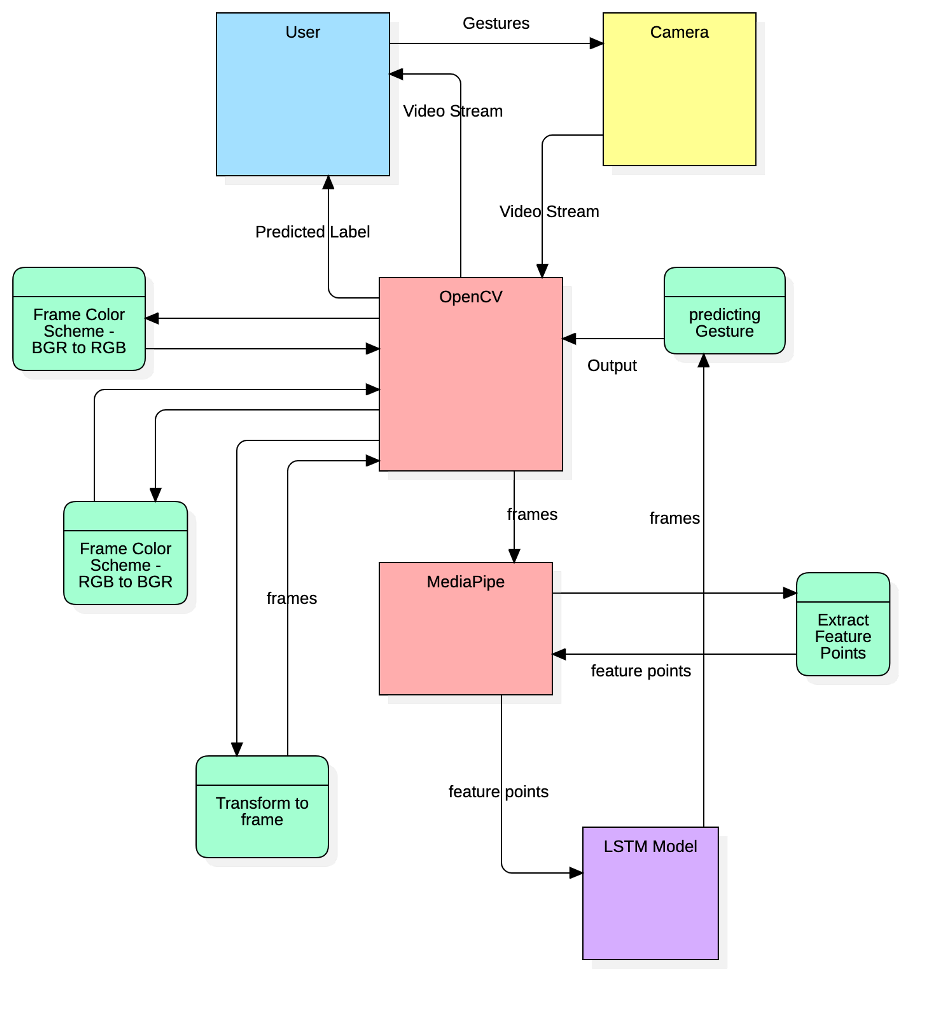
Machine Learning applications.

# Appendix B: Analysis Models

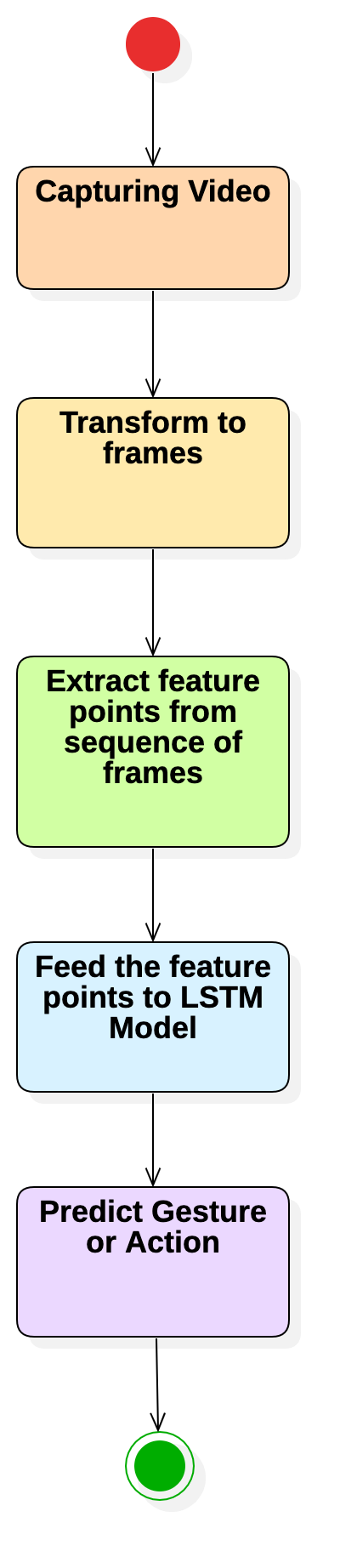
## a. Use Case Diagram



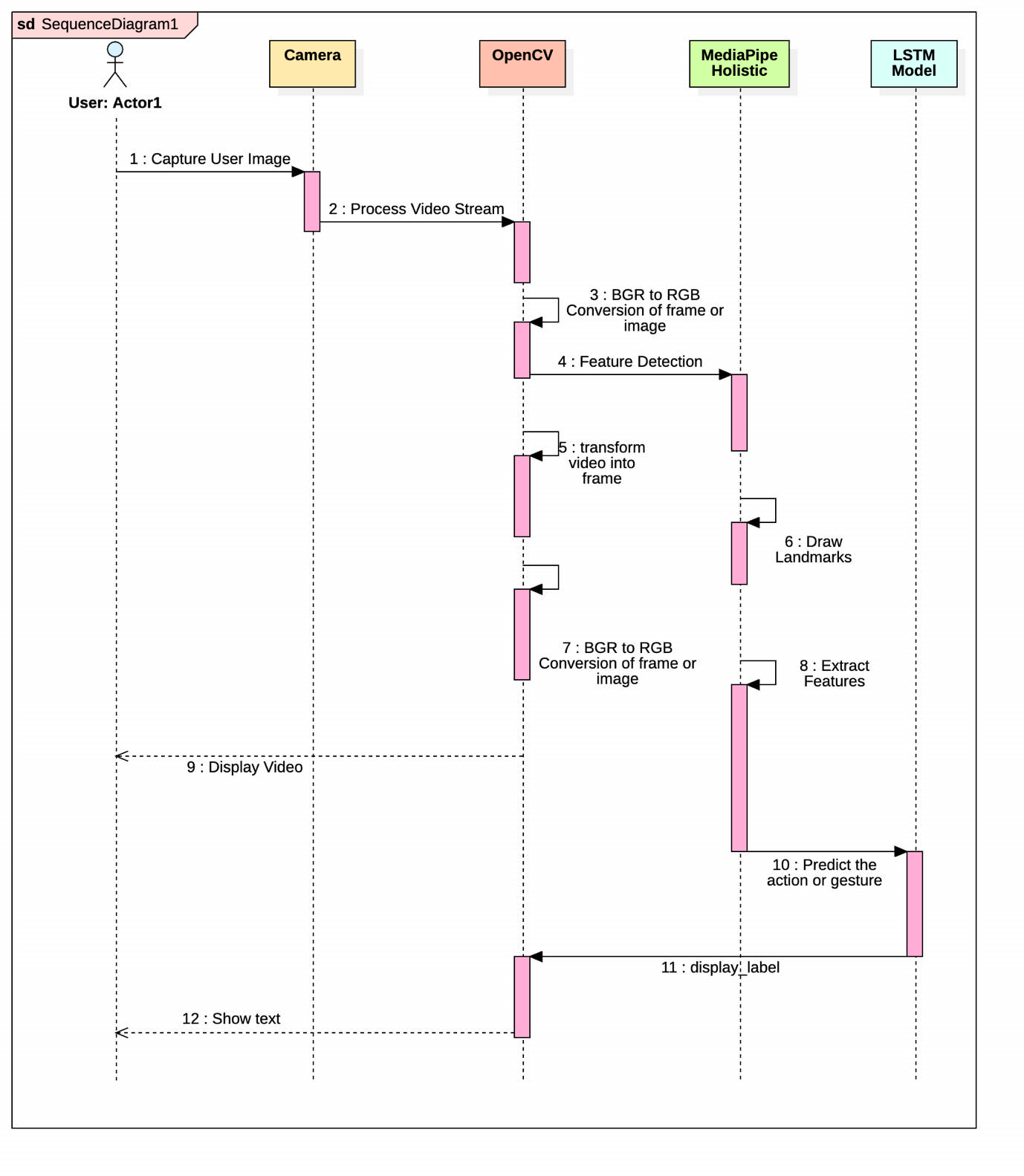
## b. Data Flow Diagram



## c. State Diagram



## d. Sequence Diagram



## e. Class Diagram

