Software Requirements

Specification

For

*Sign Language Recognition System*

09/12/2022

Prepared by

|  |  |  |
| --- | --- | --- |
| **Specialization** | **SAP ID** | **Name** |
| B.Tech-CSE (AIML) | 500082942 | Raghav Akshat Upadhyay(B6-NH) |
| B.Tech-CSE(CCVT) | 500083018 | Rakshit Pratap Singh (B6-NH) |



Department of Informatics School of Computer Science

UNIVERSITY OF PETROLEUM & ENERGY STUDIES,

DEHRADUN- 248007. Uttarakhand

# Table of Contents

|  |  |  |
| --- | --- | --- |
| **Topic** | | **Page No** |
| Table of Content | | 4 |
| Revision History | | 4 |
| 1 | Introduction | 4 |
|  | 1.1 Purpose of the Project | 4 |
|  | 1.2 Target Beneficiary | 5 |
|  | 1.3 Project Scope | 5 |
|  | 1.4 References | 5 |
| 2 | Project Description | 6 |
|  | 2.1 Reference Algorithm | 7 |
|  | 2.2 Data/ Data structure | 7 |
|  | 2.3 SWOT Analysis | 7 |
|  | 2.4 Project Features | 8 |
|  | 2.5 User Classes and Characteristics | 8 |
|  | 2.6 Assumption and Dependencies | 8 |
|  | 2.7 Design Diagrams | 9-12 |
|  | 2.8 Assumption and Dependencies | 12 |
| 3 | System Requirements | 12 |
|  | 3.1 User Interface | 12 |
|  | 3.2 Software Interface | 13 |
|  | 3.3 Database Interface | 13 |
|  | 3.4 Protocols | 13 |
| 4 | Non-functional Requirements | 13 |
|  | 4.1 Performance requirements | 13 |
|  | 4.2 Security requirements | 13 |
|  | 4.3 Software Quality Attributes | 13 |
| 5 | Other Requirements | 13 |
| Appendix A: Glossary | | 14 |

**Revision History**

|  |  |  |  |
| --- | --- | --- | --- |
| **Date** | **Change** | **Reason for Changes** | **Mentor Signature** |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

## INTRODUCTION

The act of communicating or exchanging information, ideas, or feelings is referred to as communication. Both parties must be able to speak and understand the same language for two people to create communication. Deaf and dumb people, however, use various communication methods.. They use sign language to communicate with each other and with everyday people, yet everyday people do not value sign language as much as they should. Communication between a normal person and a deaf or dumb person is challenging since not everyone is familiar with or understands sign language.

For the deaf-mute and normal individuals to converse normally, a low-cost solution is required. Therefore, scientists are working to develop a means of communication for deaf-mute persons so that they can interact with hearing people. The Sign Language Recognition System is the innovation in this. The technology seeks to understand sign language and translate it either orally or in text form into the native tongue. But there are several issues with establishing sign language recognition, from picture acquisition to categorization. The ideal technique for acquiring images is still being investigated by scientists. The challenges of picture pre-processing are presented by image collection using a camera. Using an active sensor device, meanwhile, can be expensive. Researchers' use of classification systems has certain downsides as well. Researchers are unable to choose the optimum recognition method because there are so many options. By concentrating on one method, it tends to prevent testing of other methods that might be more appropriate for sign language recognition. We aim to use Convolutional Neural Networks (CNN) to perform our classification and to design a system that can classify these signs in real time.

1.1 Purpose of the Project

The goal of this project is to build a Convolutional Neural Network based recognition system which will be able to classify which letter of the American Sign Language (ASL) alphabet is being signed, given an image of a signing hand. This project is a step towards building a possible sign language translator, which can take communications in sign language and translate them into written and oral language. Such a system would greatly lower the barrier for many deaf and mute individuals to be able to better communicate with others in day-to-day interactions. This goal is further motivated by the isolation that is felt within the deaf community. Loneliness and depression exist in higher rates among the deaf population, especially when they are immersed in a hearing world [1]. Large barriers that profoundly affect life quality stem from the communication disconnect between the deaf and the hearing. Some examples are information deprivation, limitation of social connections, and difficulty integrating in society [2].

* 1. Target Beneficiaries

The target beneficiary of the proposed project is especially abled people who are unable to or have trouble in communication via speech and the people around the who wish to communicate with them. Industry wise the project will benefit the Healthcare Industry and the Telecommunication Industry.

According to World Health Organization, they represent more than 5% of the world's population - 466 million people (432 million adults and 34 million children) and this number could rise upto 1.5 billion by the year 2050.This require an easy and cheap system for them to be able to communicate effectively communicate.

Project Scope

The scope of this project is to use to build a Sign Language Recognition System. This system is planned to be used by speechless people in order to ease their life, also anyone that want to reach and serve speechless people as well, cooperation’s and foundations which aims to help speech-disordered people.

The project is aimed to use the machine learning algorithm convolutional neural networks where the application will work as the person(user) whose gestures to be recognized stands in front of the camera, the user should use their one hand in a specific way for the system to capture the image. After capturing the image, the system will provide an output matching the image with a predefined gesture according to its training. In this project it is aimed to recognize American Sign Language and translate it into a text in English only. Other languages will neither be considered as input language (the sign language) nor as output language.

* 1. References

[1] Farnaz D. Notash and Elahe Elhamki. “Comparing loneliness, depression and stress in students with hearingimpaired and normal students studying in secondary schools of Tabriz”. In: International Journal of Humanities and Cultural Studies February 2016 Special Issue (2016). issn: 2356-5926

[2] “The Cognitive, Psychological and Cultural Impact of Communication Barrier on Deaf Adults”. In: Journal of Communication Disorders, Deaf Studies Hearing Aids 4 (2 2016). doi: 10.4172/2375-4427.1000164

[3] Huang, J., Zhou, W., Zhang, Q., Li, H., & Li, W. (2018, January 30). Video-based Sign language recognition without temporal segmentation. arXiv.org. Retrieved October 15, 2022, from https://arxiv.org/abs/1801.10111

[4] Real-time American Sign Language Recognition with -stanford university. (n.d.). Retrieved November 7, 2022, from http://cs231n.stanford.edu/reports/2016/pdfs/214\_Report.pdf

## PROJECT DESCRIPTION

The overall design algorithm of the project consists of five components:

The figure describes the overall steps involved in this project.

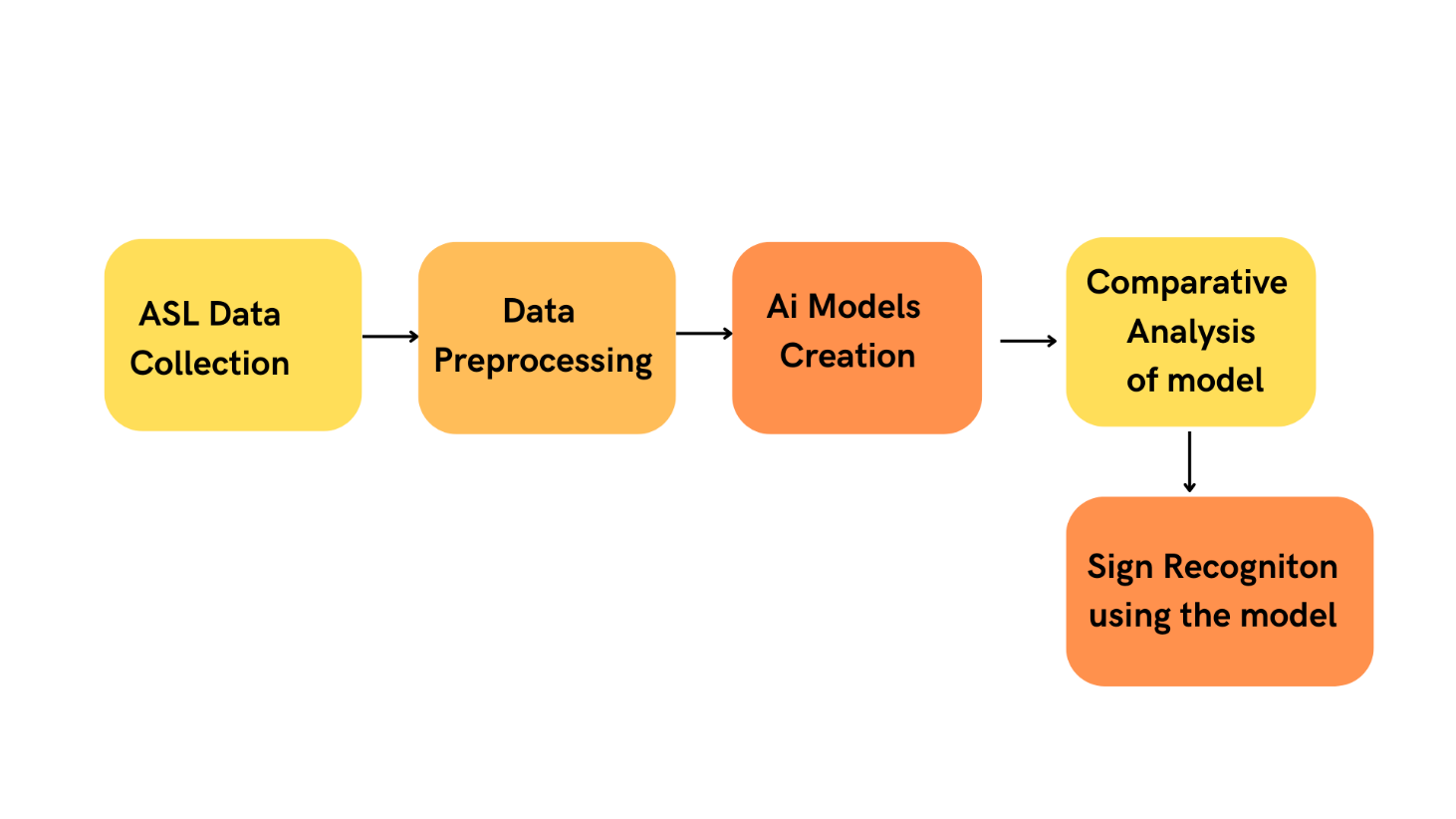


Fig.1 Project model workflow

* 1. Reference Algorithm

For this project, we have used one reference algorithms:

* Convolutional Neural Network: is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a CNN is much lower as compared to other classification algorithms. Initially we plan to create a simple neural network with one convolution layer then moving forward we will apply parameter tuning on the model using multiple layers, Batch Normalization, and Dropout layer.
  + - Data Collection – This project works upon American Sign Language MNIST dataset which is publicly available on the Kaggle platform. <https://www.kaggle.com/datasets/datamunge/sign-language-mnist>. The dataset consists of pixel values of hand sign images of various alphabets of the English Language.
    - Data Pre-Processing – After the initial phase of data collection, the next task in line was pre-processing the data we received was already in csv format and had pixel values in it. For some models we had to apply augmentation in the data so as to get a better accuracy.
    - Model Creation – After pre-processing, we will use the Tensorflow and keras library of python to create our neural network. We will try to make a number of models using different number of convolution layer, adding or removing batch-normalization, augmentation and dropout layer.
    - Real Time Detection– If we try to create a Real Time System we will use OpenCV library of python to get input images from our camera.
    - Testing and Analysis – This step contains the conclusion, results and observations of the model after proper testing and analysis.
  1. Characteristic of Data

The original MNIST image dataset of handwritten digits is a popular benchmark for image-based machine learning methods but researchers have renewed efforts to update it and develop drop-in replacements that are more challenging for computer vision and original for real-world applications.

The American Sign Language letter database of hand gestures represent a multi-class problem with 24 classes of letters (excluding J and Z which require motion). The dataset format is patterned to match closely with the classic MNIST. Each training and test case represents a label (0-25) as a one-to-one map for each alphabetic letter A-Z (and no cases for 9=J or 25=Z because of gesture motions). The training data (27,455 cases) and test data (7172 cases) are approximately half the size of the standard MNIST but otherwise similar with a header row of label, pixel1, pixel2 to pixel784 which represent a single 28x28 pixel image with grayscale values between 0-255. The original hand gesture image data represented multiple users repeating the gesture against different backgrounds. The Sign Language MNIST data came from greatly extending the small number (1704) of the color images included as not cropped around the hand region of interest. We can check a random sample of the training data in Fig 2

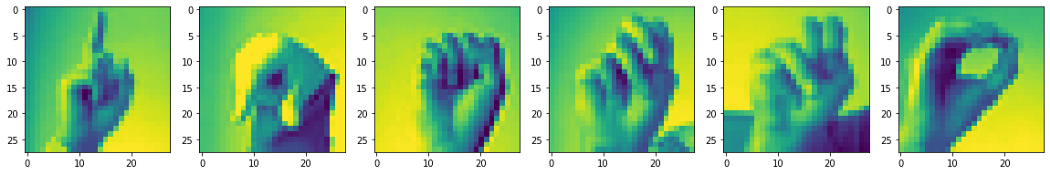


Fig 2. Random sample from dataset

* 1. SWOT Analysis
* Strength: To perform sign language detection over the different hand signs of American Sign Language and performing pre-processing using ImageDatagenerator class of Tensorflow. Strength of the project is derived by the fact that comparatively less work has been done in the fine tuning of American Sign Language Recognition using CNN and that we are classifying individual alphabets which allows the user to say anything by a combination of words
* Weakness: During the preprocessing phase, our general low-end machines will consume huge amount of memory and time.
* Threat: The project is planned and made for the assistance of the deaf/dumb people and could not be directly used by a layman
* Opportunity: This gives us an opportunity to help the especially abled people and give back to society.

2.4 Project Features

* In this project, pre-processing of data is performed by using Image-Data-Generator which is one of the best classes for processing of data in Tensor-flow.
* Our project uses Asl dataset to classify images into one of the given Hand Sign in English language

Use Case Diagram:

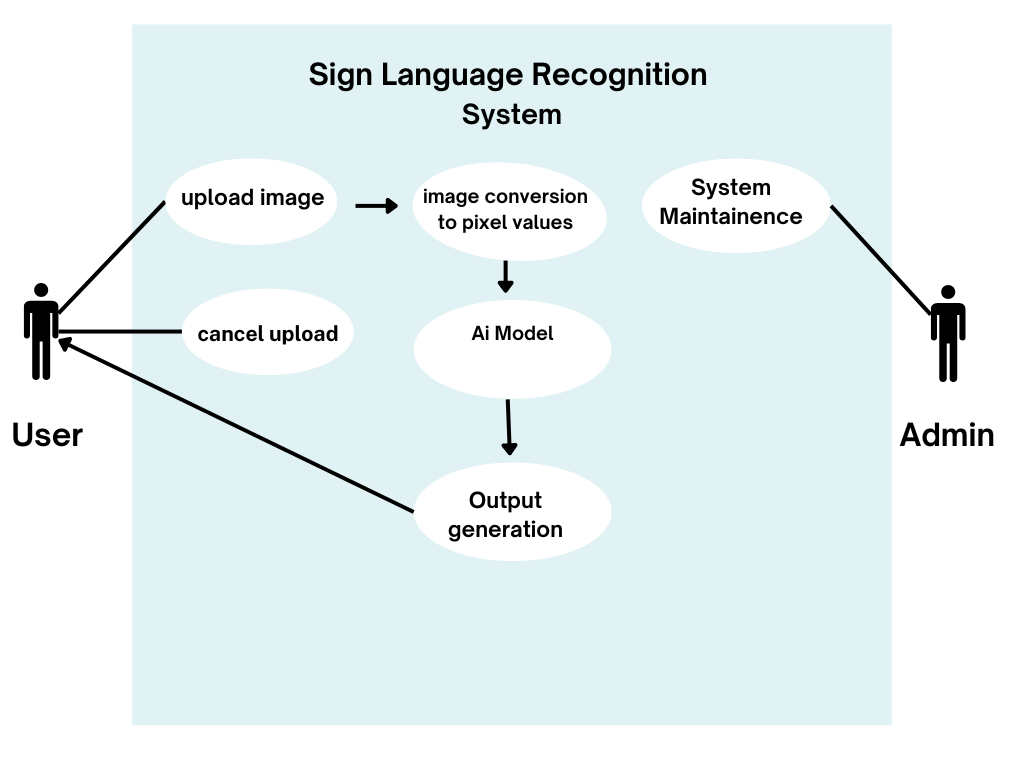


Fig 3. Use Case Diagram

2.5 User classes and characteristic

User classes of the project are deaf and dumb people also people who interact with them. The project focusses on accurately detecting the alphabets of American Sign Language using Convolutional Neural Networks.

2.6 Design and Implementation Constraints

DESIGN:

The user interface of our project basically takes input by the user and processes it to give output.

CONSTRAINTS:

1. The constraints of this can be that if the user tries to process a lot of data the system might take a lot of time and memory space to render the raw data.

2. If the user gestures an incorrect gesture even then the system will classify it one of the alphabets as we do not have a label for none

2.7 Design Diagrams

Data Flow Diagram:

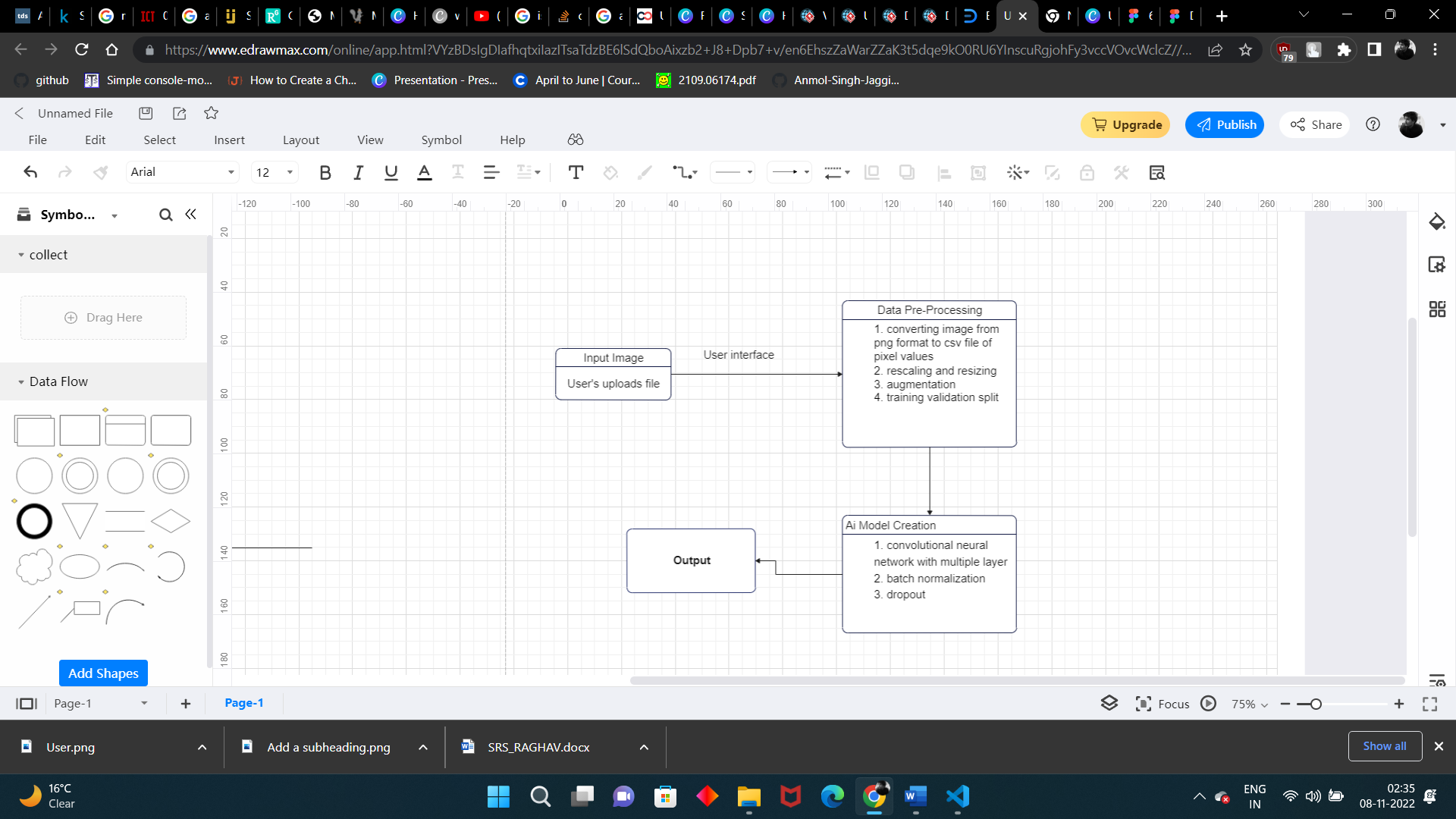


Fig 4: Data Flow Diagram

Class Diagram:

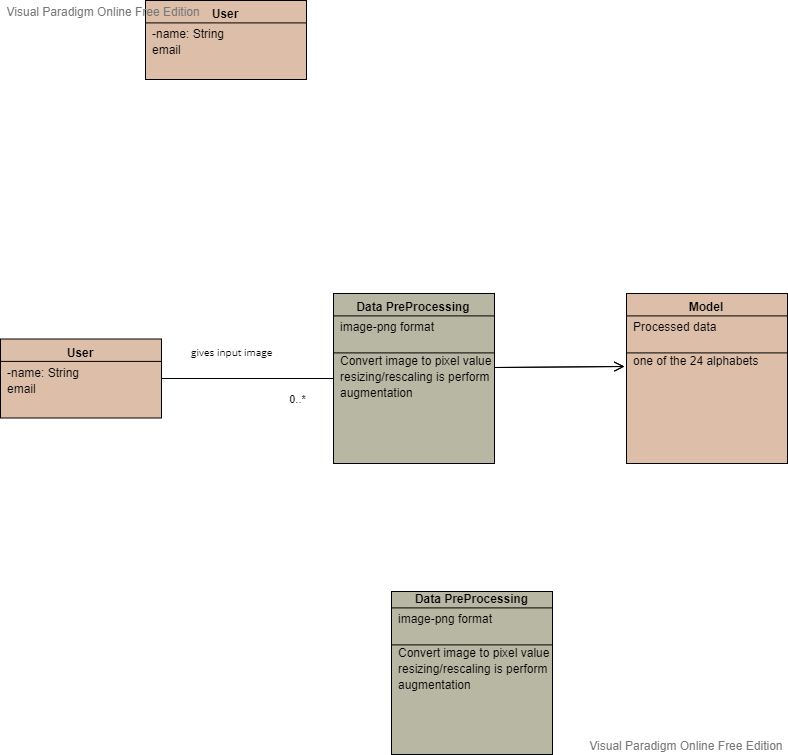
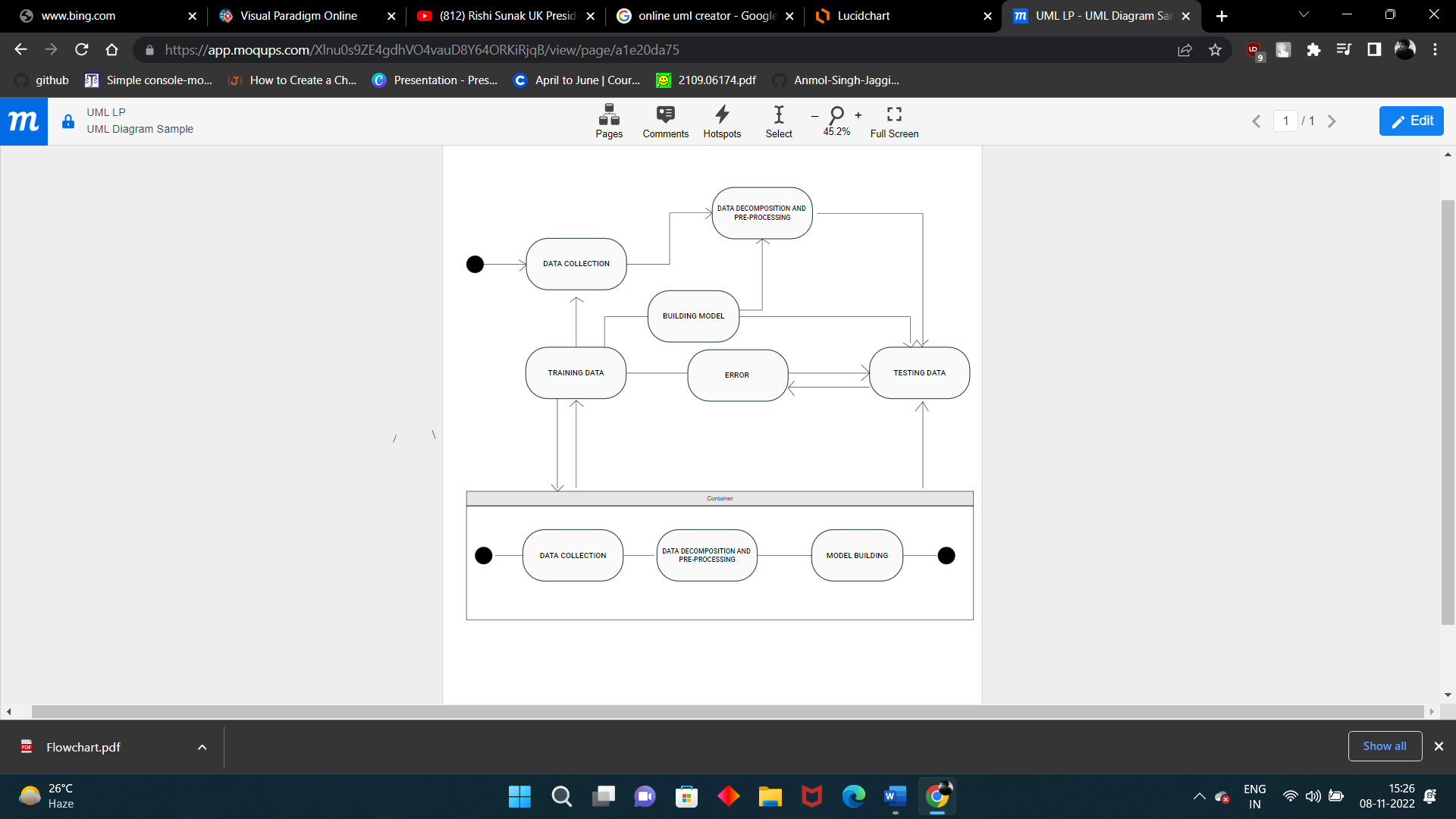
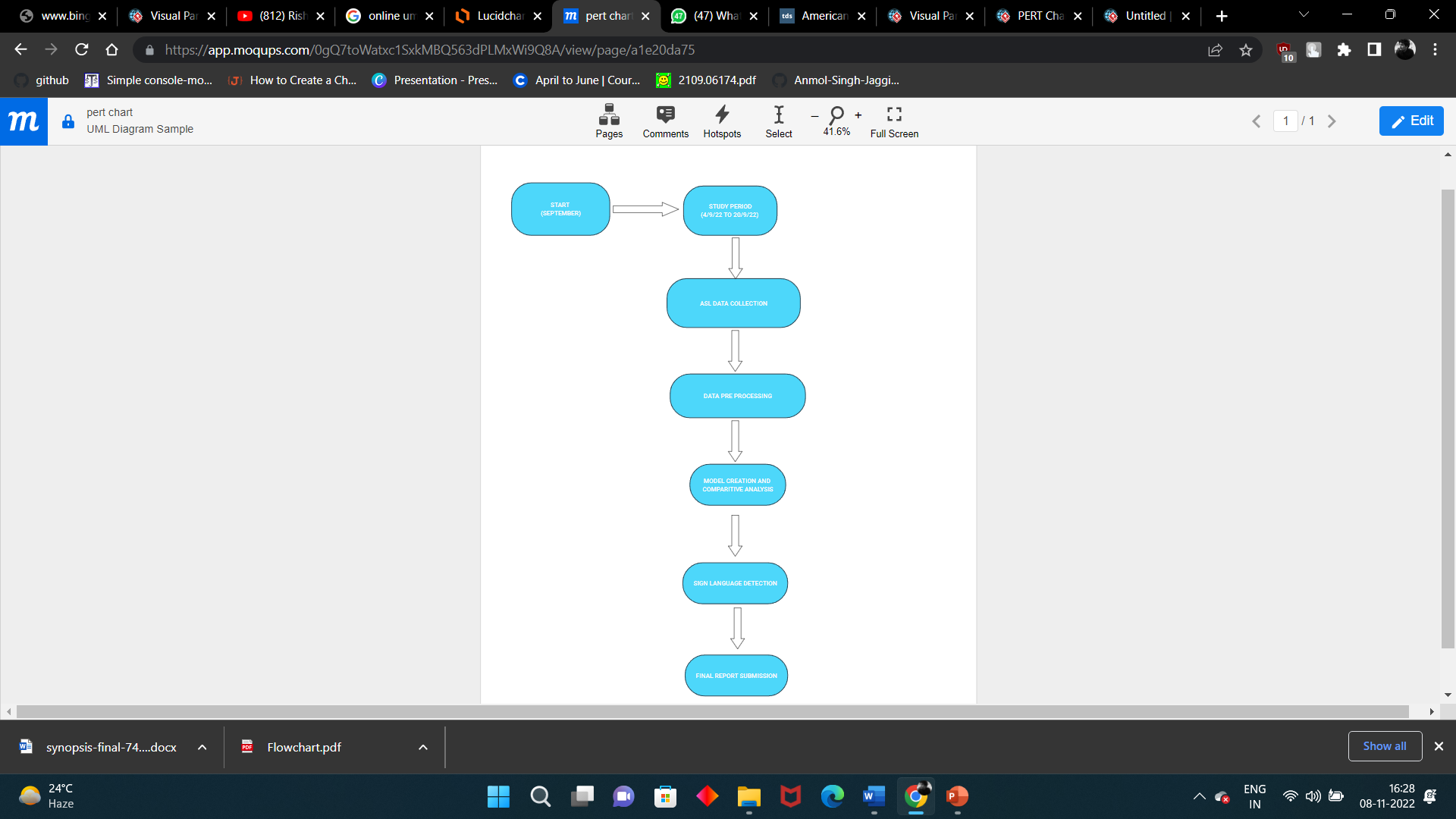


Fig: Class Diagram

UML Diagram:



PERT Chart:



2.8 Assumption and Dependencies

* No assumption and dependencies

## SYSTEM REQUIREMENTS

* 1. User Interface
     + ASL dataset is used to train the model.
     + Libraries used-

1. OpenCV
2. Numpy
3. Matplotlib
4. OS
5. Tensorflow
6. Tensorflow-keras
7. Seaborn
8. pandas
9. PIL

3.2 Software Interface

* + - The software use jupyter-notebook and VS code to implement the code.
  1. Database Interface No use of database.
  2. Protocols

No protocols have been used in our proposed model.

## NON- FUNCTIONAL REQUIREMENTS

* 1. Performance Requirements
     + The project should be able to detect and segment each and every kind of image, the dataset is containing.
     + The project should have high accuracy.
     + The proposed technique should be better than the existing models 4.2 Security requirements.
     + There is no specific security requirement for our project.

4.3 Software Quality Attributes

* + - The output of the project should not be misleading, as it can cause a drawback in the

Communication process of people.

## OTHER REQUIREMENTS

* There are no other requirements for our project

## APPENDIX A: GLOSSARY

* + Pre-processing of ASL data- Preprocessing data is a common first step in any project to prepare raw data in a format that the AI model can accept. For example, we’ll be converting the image data to pixel format. You can also preprocess data to enhance desired features or reduce artifacts that can bias the AI model.
  + Data Augmentation- Data Augmentation is an essential step in training the neural network. For example, in the training dataset, we have hand signs of the right hands but in the real world, we could get images from both right hands as well as left hands. Data Augmentation allows us to create unforeseen data through Rotation, Flipping, Zooming, Cropping, etc.
  + Test and analysis- The following step will handle the accuracy of the AI model and give the overall analysis of the project and its working.

## APPENDIX B: ANALYSIS MODEL

The waterfall model has helped us derive a systematic approach towards our motto.

## APPENDIX C: ISSUES

## Not-Applicable