ST. VINCENT PALLOTTI COLLEGE OF ENGINEERING & TECHNOLOGY, NAGPUR

Department of Computer Engineering

6th SEM Mini Project

REAL-TIME SIGN LANGUAGE CONVERSION TO TEXT

Group No. : 3A10

Members:

Naina Makhijani   15

Shreeyash Gaiki   57

Yamini Mirashe   26

Saloni Sable   19

Project Guide:

Dr. Kapil Gupta

CONTENT

- **Project Overview**
- **Project Aim**
- **Objectives**
- **Dataset**
- **Algorithm**
- **ML Model**
- **Workflow**
- **Schedule Diagram**
- **Requirements**
- **References**

PROJECT OVERVIEW

Project Title:

Real-time Sign Language Conversion to Text

Problem Statement:

To develop a system that can recognize sign language gestures in real-time and convert them into text.

PROJECT AIM

Aim:

Our project aims to create a computer application and train a model which when shown a real time video of hand gestures of American Sign Language shows the output for that particular sign in text format on the screen.

The primary goal is to bridge the communication gap between individuals who use sign language and those who may not be familiar with it, facilitating seamless and inclusive communication.

OBJECTIVES

To implement a machine learning model which will identify 26 English alphabet (A-Z or a-z), Numbers (0 to 9) and Black space.

To develop a user interface consisting of a camera screen to capture real-time video of hand gestures and a text-box to display the converted sentence in English.
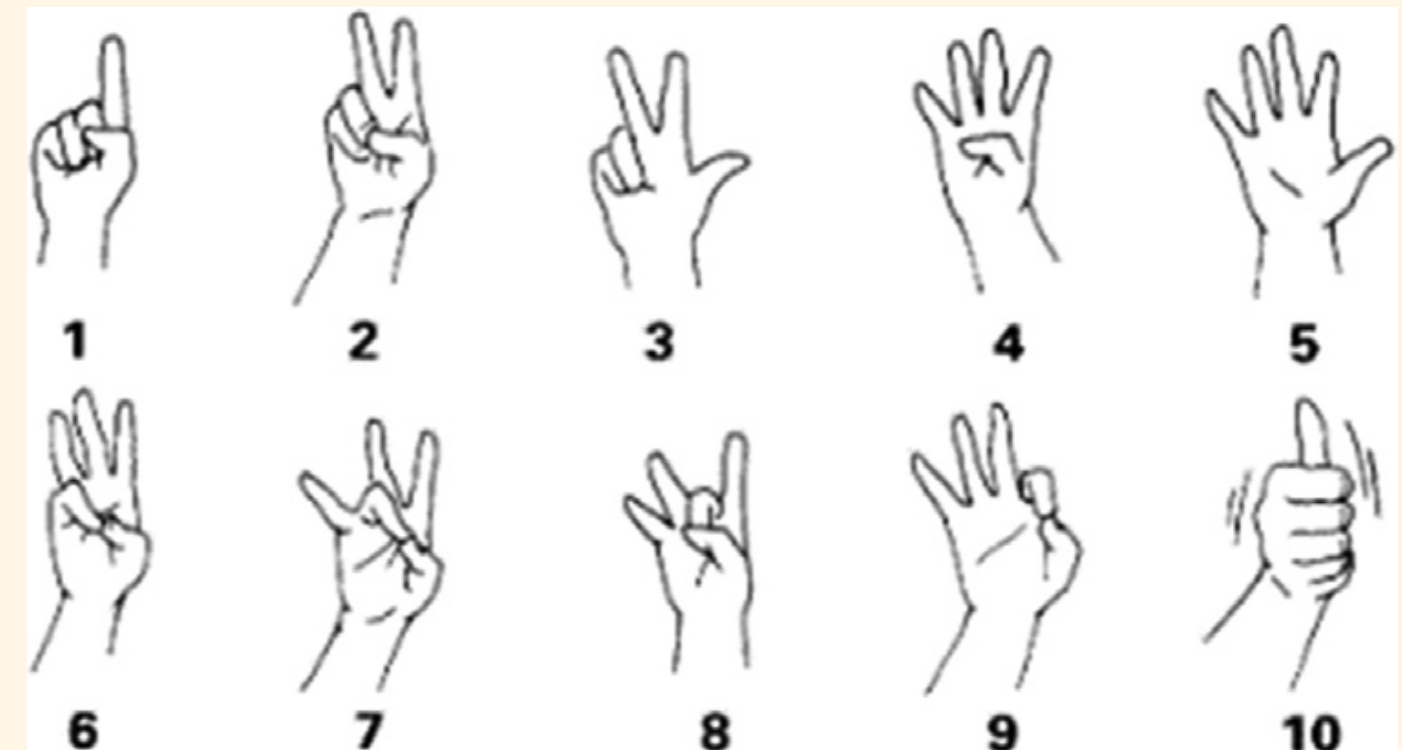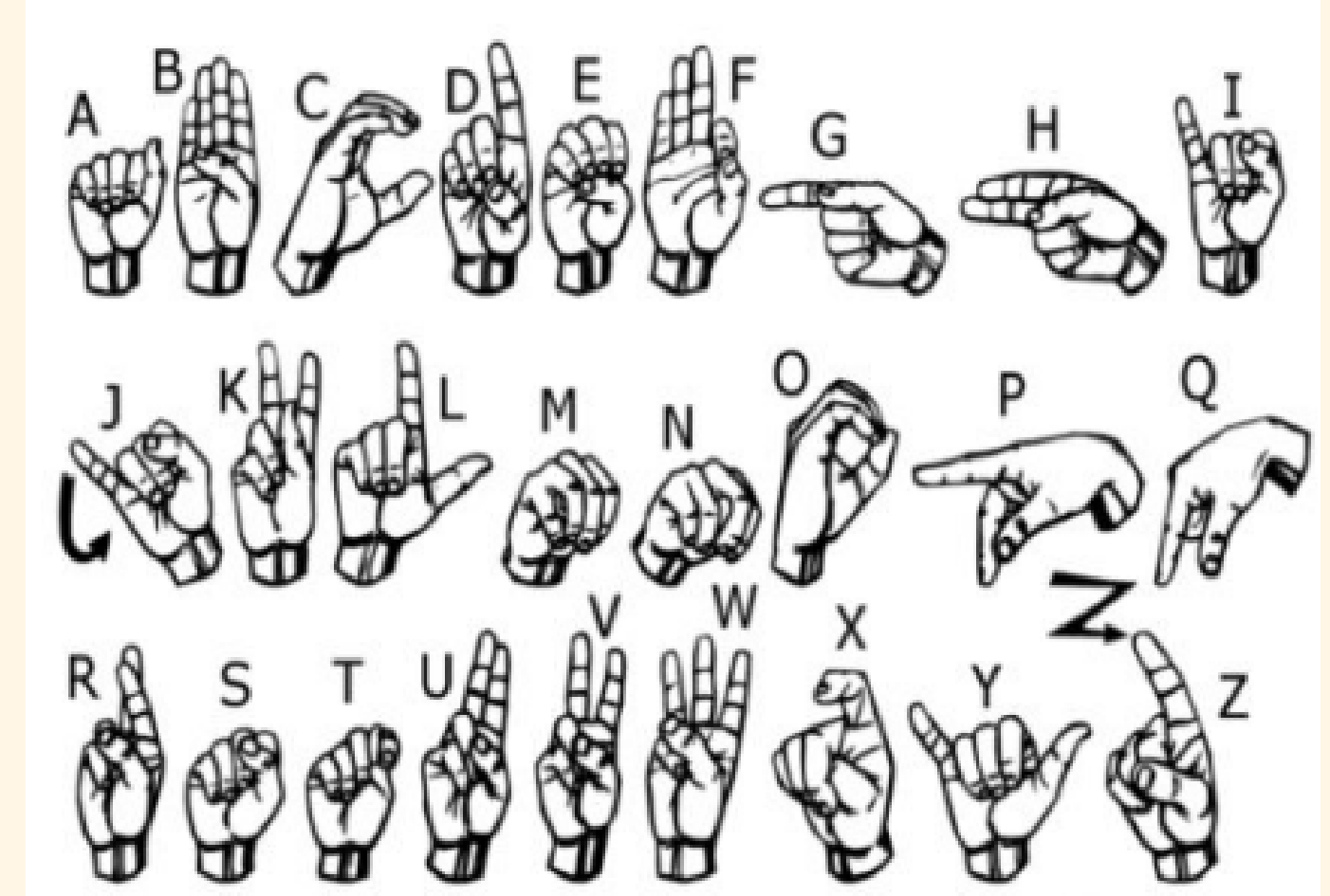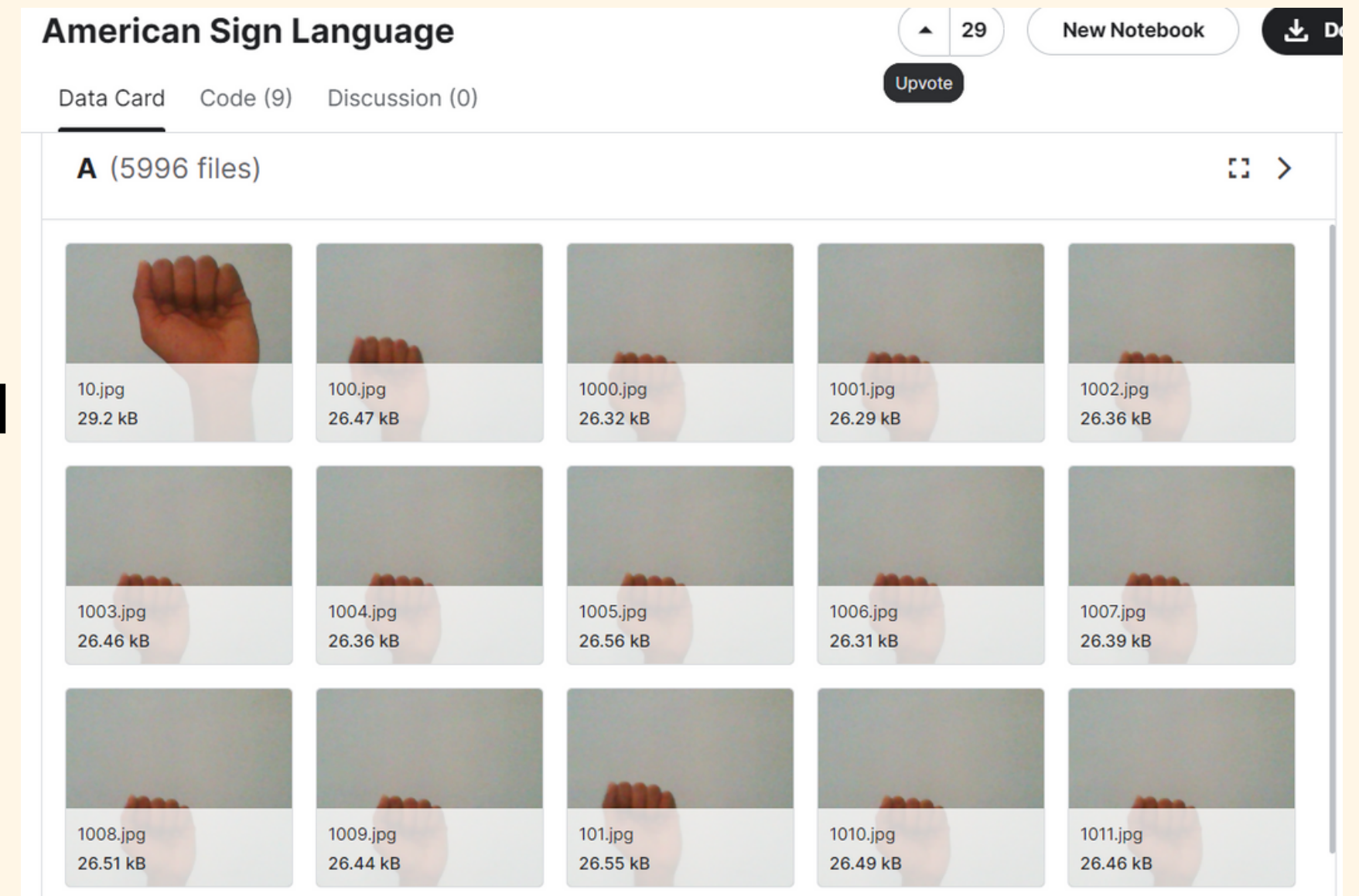
DATASET

We will be using following datasets:
- **American Sign Language Dataset**
- **American Sign Language**

While training each image will be labelled to its character.

The dataset consist of 6000 different images of each letter and 70 different images of each number, total 169k images of training data.
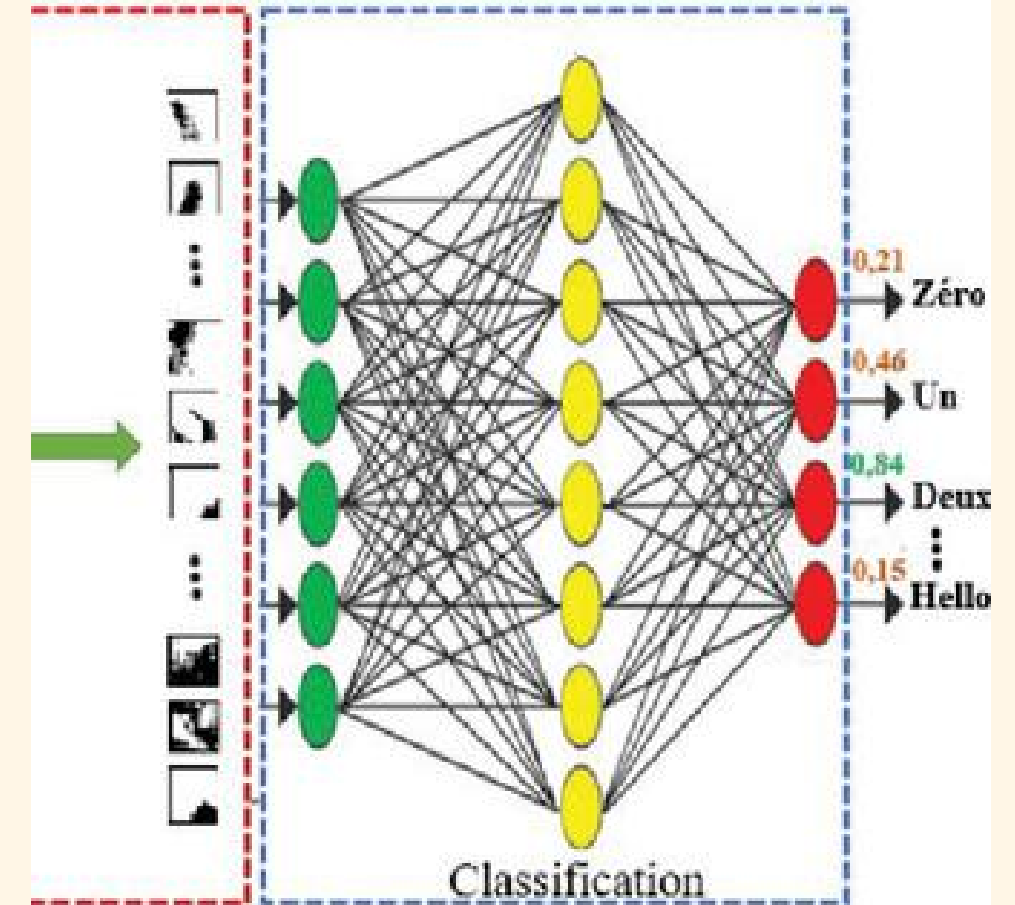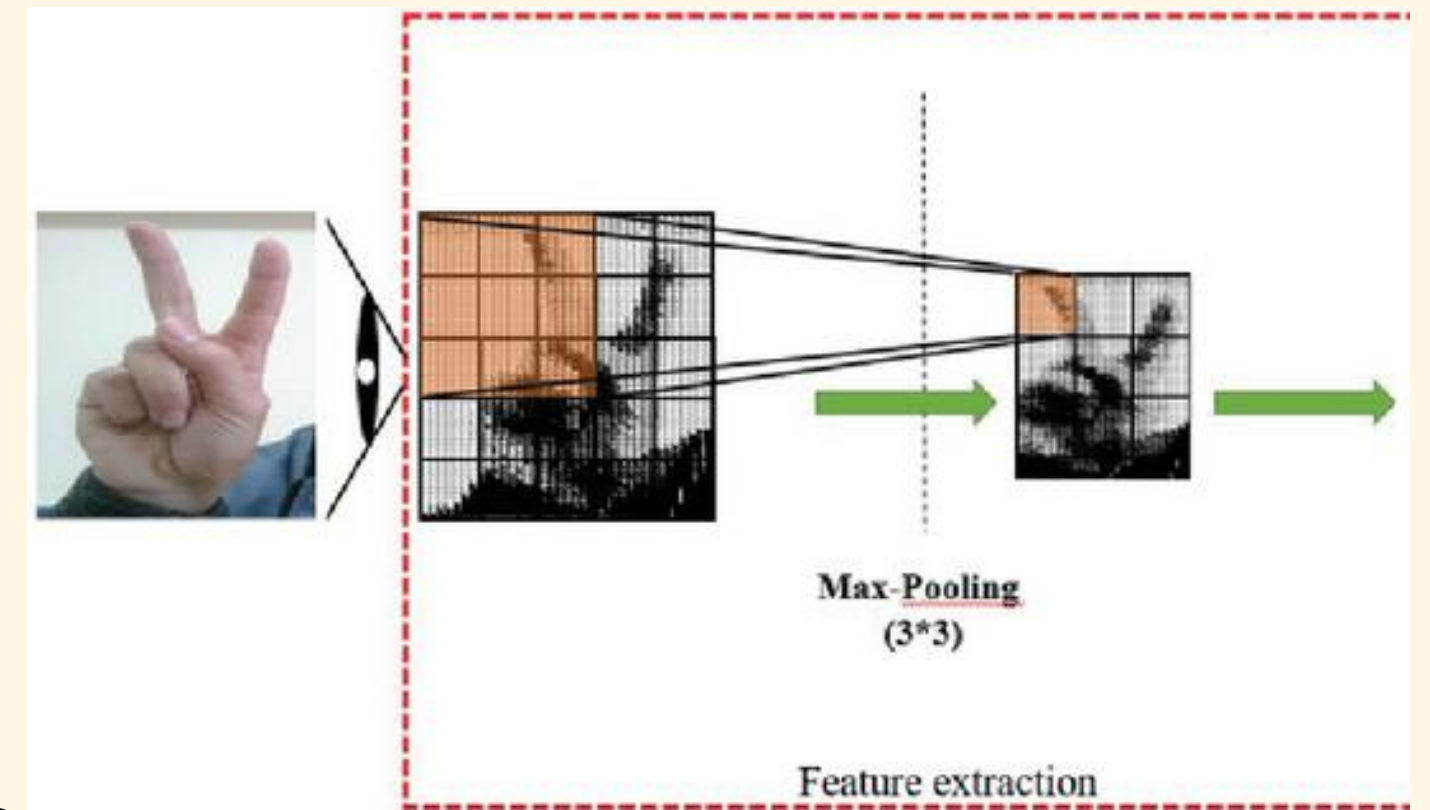
ALGORITHM

- **Continuous stream of images is captured and processed for image normalization and other filtering to feed into ML model.**

- **Processed frames are given to trained model , which classifies the input image with the gesture letter.**

- **Model maps each frame to a letter and if same frame is detected for more than few frames, then it is taken into consideration for word forming.**

- **Space between the words are considered using blank symbol.**

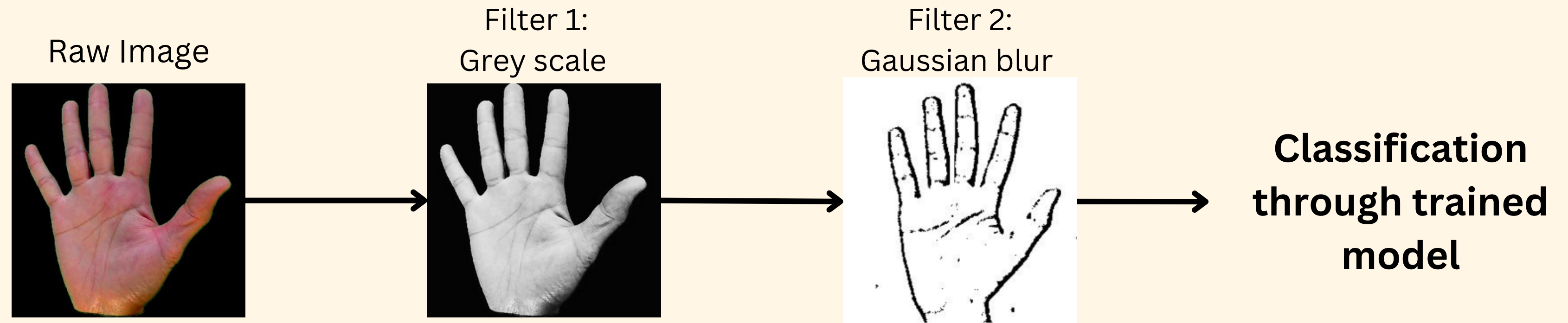- **The sentence formed is displayed on the output screen.**

ML MODEL

Convolutional Neural Networks

- CNNs consist of multiple convolutional layers each layer containing numerous "filters" which perform feature extraction.

- Initially these "filters" are random and by training, the feature extraction gets better.

- CNN is primarily used for image classification.

- During training, the CNN learns to optimize its internal parameters to accurately map input frames to corresponding gesture.

- Adjustments are made to improve the model's accuracy and generalization.

Max-Pooling (3*3)

Feature extraction

0,21 Zéro
0,46 Un
0,84 Deux
0,15 Hello

Classification

ML MODEL

Convolutional Neural Networks

Raw Image

Filter 1:
Grey scale

Filter 2:
Gaussian blur

Classification through trained model

- **Different filter layers will be applied to each image frame to ease the model processing, also increasing its accuracy in different conditions.**

WORKFLOW

- **Project Planning and Research**
- **Data Collection and Preprocessing**
- **Model selection and Architecture**
- **Model training**
- **Integration with Text Generation**
- **Real-Time Implementation**
- **User Interface Development**
- **Testing and Evaluation**

	ACTIVITIES	ASSIGNEE	EH	START	DUE	%
	Project Planning and Research:		-	29/Jan	03/Feb	73%
1	Objective Definition	Na, Sh, Ya	-	29/Jan	03/Feb	100%
2	User Requirements	Saloni	-	29/Jan	03/Feb	100%
3	Research	Na, Sh, Ya	-	29/Jan	03/Feb	20%
	Data Collection and Preprocessi...		-	05/Feb	11/Feb	33%
5	Dataset Selection	All	-	05/Feb	07/Feb	100%
6	Annotation	All	-	07/Feb	11/Feb	0%
7	Preprocessing	All	-	07/Feb	11/Feb	0%
	Model Selection and Architectu...		-	12/Feb	17/Feb	0%
9	Choose Model	All	-	12/Feb	17/Feb	0%
10	Architecture Design	All	-	12/Feb	17/Feb	0%
	Training the ASL Recognition M...		-	18/Feb	02/Mar	0%
12	Data Splitting	All	-	18/Feb	24/Feb	0%
13	Model Training	All	-	18/Feb	28/Feb	0%
14	Fine-Tuning	All	-	28/Feb	02/Mar	0%
	Integration with Text Generati...		-	03/Mar	09/Mar	0%
16	Choose Text Generation Mo...	All	-	03/Mar	09/Mar	0%
17	Integration	All	-	03/Mar	09/Mar	0%
	Real-Time Implementation:		-	10/Mar	17/Mar	0%
19	Real-Time Processing	All	-	10/Mar	17/Mar	0%
20	Stream Processing	All	-	10/Mar	17/Mar	0%
	User Interface Development:		-	18/Mar	24/Mar	0%
22	Design Interface	All	-	18/Mar	24/Mar	0%
23	Integration	All	-	18/Mar	24/Mar	0%
	Testing and Evaluation:		-	25/Mar	03/Apr	0%
25	Functional Testing	All	-	25/Mar	29/Mar	0%
26	User Testing	All	-	27/Mar	03/Apr	0%

SCHEDULE DIAGRAM
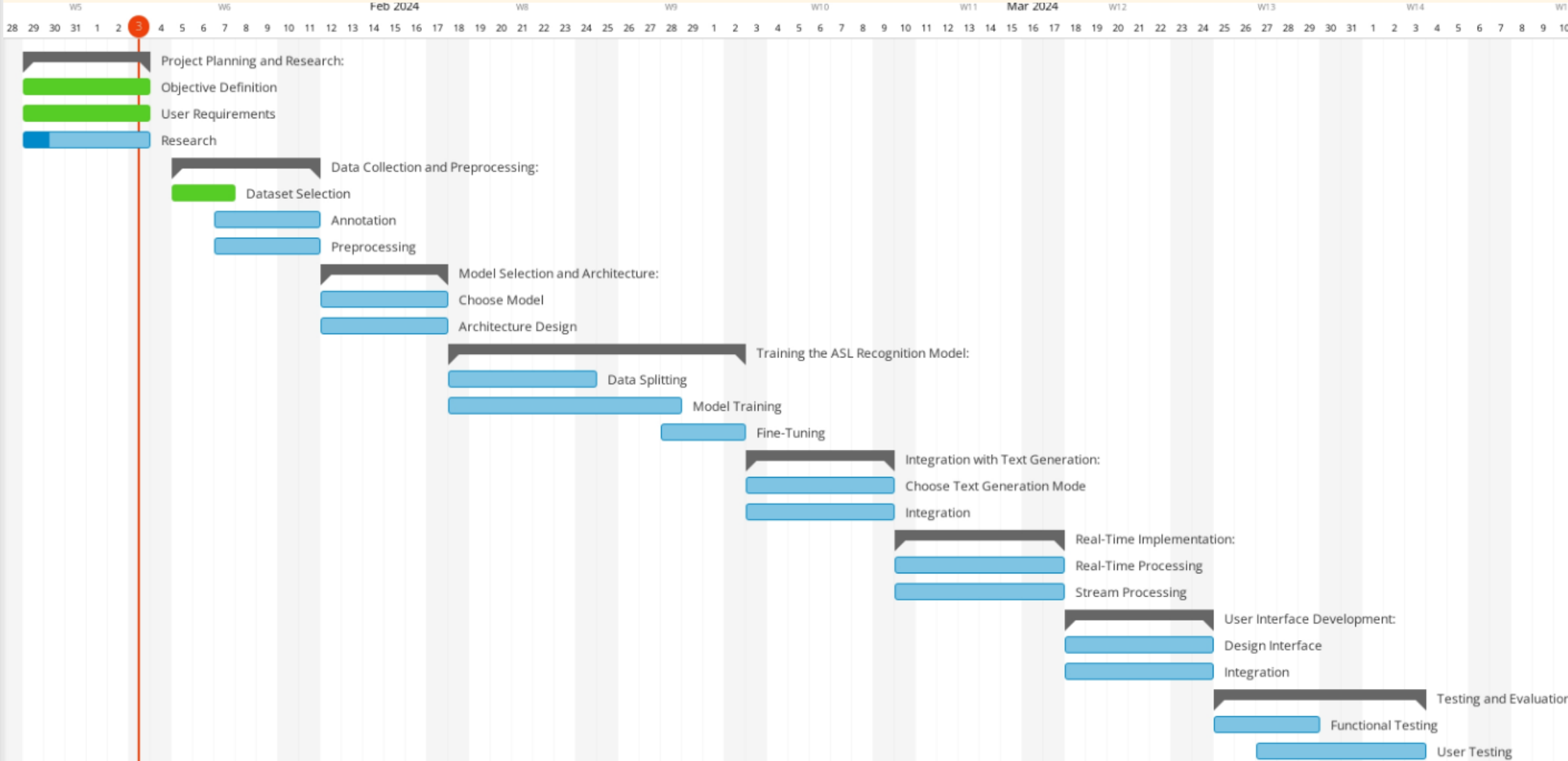
REQUIREMENTS

Software:

- **Python 3**
- **Tensorflow**
- **OpenCV**
- **NumPY**
- **Keras**
- **pyttsx3**
- **Dataset**

Tools:

- **IDE**
- **GitHub**
- **Documentation**
- **Jupyter Notebook**

REFERENCES

- **What is CNN?(https://youtu.be/QzY57FaENXg si=VqNrYcp8trtWBPaZ)**
- **Simple explanation of CNN (https://youtu.be/zfiSAzpy9NM?si=afI-3hDU8HaAX-4Y)**
- **Sign language using CNN (https://youtu.be/NQPV2344cGE?si=fSUcEZ94BXPRQiRn)**
- **A brief guide to CNN (https://medium.com/nybles/a-brief-guide-to-convolutional-neural-network-cnn-642f47e88ed4)**
- **Tools used in ML (https://neptune.ai/blog/top-tools-to-run-a-computer-vision-project)**
- **Kaggle dataset (https://www.kaggle.com/datasets/kapillondhe/american-sign-language), (https://www.kaggle.com/datasets/ayuraj/asl-dataset)**

THANK YOU