A Mini Project Report

On

Sign Language Translator

Submitted in partial fulfilment of the requirement of

University of Mumbai

For the Degree of

Bachelor of Engineering (SE)

in

COMPUTER ENGINEERING

Submitted by

Sahil Raj (1020129)

Anushka Gaat (1020140)

Saatvik Inampudi (1020148)

Vithi Kulkarni (1020164)

Supervised by

Ms. Snehal Nikalje

Department of Computer Engineering

Fr. Conceicao Rodrigues Institute of Technology
Sector 9A, Vashi, Navi Mumbai - 400703

UNIVERSITY OF MUMBAI

2021-2022

APPROVAL SHEET

This is to certify that the Mini project entitled

Sign Language Translator

Submitted by

Sahil Raj	**1020129**
Anushka Gaat	**1020140**
Saatvik Inampudi	**1020148**
Vithi Kulkarni	**1020164**

Supervisors: **Ms. Snehal Nikalje**

Project Coordinator: **Mrs.Dakshayani R**

Examiners: **1.**

2.

Head of Department: Dr.Lata Ragha

Date:

Place:

Declaration

We declare that this written submission for S.E. Mini Project entitled ”**Sign Language Translator**” represent our ideas in our own words and where others’ ideas or words have been included. We have adequately cited and referenced the original sources. We also declared that we have adhere to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any ideas / data / fact / source in our submission. We understand that any violation of the above will cause for disciplinary action by institute and also evoke penal action from the sources which have thus not been properly cited or from whom paper permission have not been taken when needed.

Project Group Members:

1. Sahil Raj, 1020129

2. Anushka Gaat, 1020140

3. Saatvik Inampudi, 1020148

4. Vithi Kulkarni, 1020164

Abstract

Communication is giving, receiving or exchanging ideas, information, signals or messages through appropriate media, to give information or to express emotions. There are some people in our society who face many challenges while communicating with each other and also to the normal people. Few techniques that are used for image recognition are gloves. A glove is constructed by embedding sensors at the desired positions. Gloves containing flex sensors for detecting the bending of the fingers are made of a material which changes its resistance upon bending. Since sensor gloves are expensive and complicated to use it is always not the best method for image recognition. Another way of image recognition is taking image input as HSV and converting it to grayscale then using three layers of CNN, the images are classified. Instead of converting HSV input to RGB and then to grayscale, image can be taken directly in RGB values, filtered and feed the extracted features to the neural network for classification.

Keywords: Feature extraction, Convolutional neural networks, ANN Layers

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Sign language recognition is a problem that has been addressed in research for years. However, we are still far from finding a complete solution available in our society.

Indian sign language was created in 1999.

The term machine learning was coined in 1959 by Arthur Samuel, an American IBMer and pioneer in the field of computer gaming and artificial intelligence.

1.2 Motivation

One of the most valuable gifts to a human being is a capability to envision, listen, speak and respond according to the situations. But there are some unfortunate ones who are deprived of this. Communication between hearing impaired and normal peers have always been a challenging task. So this creates a barrier in the communication between these two communities.

To solve this issue we thought about making an user-friendly interface to bridge the communication gap and make them feel less separated from the society.

We took this project as an opportunity to learn Machine learning in depth and expand our creativity and imagination. Just the thought that given the data you can extract something useful from it is already very motivating.

1.3 Aim and Objective

Our aim is to build a reliable program which acquires image data of sign language using webcam and converts it to text and audio output. The task is accomplished using Image processing and Machine learning. Our proposed project can find its applicability in the day to day life for communication.

Our objective of developing such a user interface is to make the life easier for those in need so that they don’t feel left out and it is better than other user interfaces because of its compatibility and user friendly environment.

1.4 Report Outline

Chapter 2 is about the literature survey which consists of several research papers. Primarily, this section discusses of detailed information about previous methodologies on this topic followed by a summary table. It also consists of the existing systems and its drawbacks.

Chapter 3 includes the problem statement and scope of the project.

Chapter 4 talks about the overall design of the project, which includes the flowchart and image of training-testing split, Model working and functioning of CNN layers of the project.

Chapter 5 consists of the conclusion and the future scope of the project.

Chapter 2

Study of the System

2.1 Literature Survey

2.1.1 A Naives Bayes Classifier with Distance Weighting for Hand-Gesture Recognition,2009

Idea mentioned in the paper

Gestures are classified using a weighted K-Nearest Neighbours Algorithm, which is combined with a naïve Bayes approach to estimate the probability of each gesture type.

Technology/tool used

CNN (Convolutional Neural Network)

KNN algorithm: It can be used for Regression as well as for Classification but mostly it is used for the Classification problems.

Naïve Bayes Classifier: It is one of the simple and most effective Classification algorithms, which helps in building the fast machine learning models that can make quick predictions. It is a probabilistic classifier, which means it predicts based on the probability of an object.

[1]

2.1.2 Real-time Conversion of Sign Language to Text and Speech, 2020.

Method followed in the paper

The steps involved in the pre-processing of the frame are Contour masking, Skeletonization and Canny edges

The steps involved in post processing of frame are Upscaling and Resizing

Technology/tool used

JAVA based openCV is used. Support Vector Machine (SVM) is used for classification. In SVM, data items are plotted in n-dimensional space where n is the number of features.

[2]

2.1.3 Application of Machine Learning Techniques for Real-Time Sign Language Detection using Wearable Sensors,2020

Idea mentioned in the paper

A glove is constructed by embedding sensors at the desired positions. With the constructed glove a significant amount of data is needed to be collected which are fed to train a Machine Learning (ML) model to classify the captured data.

Technology/tool used

Flex sensors for detecting the bending of the fingers. Flex sensors are made of a material which changes its resistance upon bending. The constructed glove is interfaced with an Arduino Mega microcontroller board via circuitry concise in a veroboard.

A suitable Machine Learning algorithm has been chosen from four after a comparative study for the final deployment of the system. The four algorithms are Artificial Neural Network (ANN), Support Vector Machine (SVM), and K-Nearest Neighbor (KNN).

[3]

2.1.4 Conversion of Sign Language into Text Using Machine Learning Technique, 2021

Idea mentioned in the paper

The input is captured through the camera using OpenCV and passed through the CNN classifier. The captured input is then passed through several layers to extract features from the image and reduce the number of parameters when images are too large. Then softmax layer is used to show the output.

Technology/tool used

Python libraries like OpenCV, Matplot lib, Numpy are used.

OpenCv (Open Computer Vision) is used for Image Processing. This will be required to perform various actions and processing of the captured images.

Matplotlib is a plotting library for the Python programming language.

Numpy for adding support to large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays.

[4]

2.1.5 Sign Language Interpreter using Image Processing and Machine Learning, 2019

Idea mentioned in the paper

The proposed system consists of two main stages, Segmentation of hand and Recognition of hand sign. In this system, segmenting the hand using YCbCr colour space is firstly done and then processing the image through HOG and then it is provided to the model.

Technology/tool used

Histograms of oriented gradients (HOG) is used as a feature descriptor. The image is divided into small cells and a histogram of gradient is calculated for each cell. It creates a bin and combines the histogram of different samples based on magnitude and angle. Support Vector Machine (SVM) is used for classification. In SVM data items are plotted in n-dimensional space where n is the number of features. Each feature is associated with a coordinate value. Then it finds a hyperplane that differentiates the classes.
[5]

2.1.6 Recognition of ASL using Image Processing and Machine Learning, 2019

Methods followed in the paper

Image Acquisition, Feature extraction, Orientation detection, Gesture recognition and Text to speech

Technology/tool used

General architecture was a common CNN architecture, consisting of multiple convolutional and dense layers. The architecture included 3 groups of 2 convolutional layers followed by a max-pool layer and a dropout layer, and two groups of fully connected layer followed by a dropout layer and one final output layer.
[6]

2.1.7 Hand gesture recognition in images and video,2010

Idea mentioned in the paper

Image processing technique, cancellation of background, noise effect, it then extracts relevant features for classification and finally classifies the gesture features using a multiclass Support Vector Machine classifier.

Technology/tool used

CNN (Convolutional Neural Network)
Their algorithm first applies image processing techniques on the images in order to cancel background and noise effects on the image, it then extracts relevant features for classification and finally classifies the gesture features using a multiclass Support Vector Machine classifier. The algorithm is robust and operates well on several different backgrounds, lighting and noise conditions.
[7]

2.1.8 Summary table

Table 2.1: Table of Comparison

Sr no.	Paper name	Author name	Methodology	Advantages	Disadvantages
1.	A Naïve Bayes Classifier with Distance Weighting for Hand-Gesture Recognition, 2009.[2]	Thomas Müller, Mary Ellen Foster and Pujan Ziaie	Gestures are classified using a weighted K-Nearest Neighbours Algorithm which is combined with a nave Bayes approach to estimate the probability of each gesture type.	Gestures are extracted from each frame of the video, with a static background. This method is not dependent on skin color	The k-nearest neighbours (KNN) algorithm significantly slows as the size of that data in use grows.
2.	Real-time Conversion of Sign Language to Text and Speech, 2020.[3]	Kohsheen Tiku, Jayshree Maloo, Aishwarya Ramesh, Indra R	JAVA based openCV is used. The steps involved in the pre-processing of the frame are: 1.Contour masking 2. Skeletonization 3.Canny edges The steps involved in post processing of frame are: 1.Upscaling 2.Resizing	They have added functionality/options like adding an alphabet, erasing the detected alphabet, clearing the detected sentence and converting entire text to speech.	Proposed system only recognises alphabets from customised SVM models.

			Classification is done using SVM prediction		
3.	Application of Machine Learning Techniques for Real-Time Sign Language Detection using Wearable Sensors, 2020.[4]	Ashikur Rahman, Nazmus Saquib	A glove is constructed by embedding sensors at the desired positions. With the constructed glove a significant amount of data is needed to be collected which are fed to train a Machine Learning (ML) model to classify the captured data.	This work provides a detailed explanation of an FSM based method of dynamic character recognition.	Sensor gloves are expensive and complicated to use.
4.	Conversion of Sign Language into Text Using Machine Learning Technique, 2021.[5]	Sayali Gore, Namrata Salvi, Swati Singh	The input is captured through the camera using OpenCV and passed through the CNN classifier. The captured input is then passed through several layers to extract features from the image and reduce the number of parameters when images are too large. Then softmax layer is used to show the output.	The number of convolution layer and pooling layer can be increased for more accuracy. This algorithm is particularly used to improve the recognition accuracy under challenging conditions such as a change in scale, rotation and translation.	Since CNN is used, lots of training data is required.
5.	Sign Language Interpreter using Image Processing and Machine Learning, 2019. [6]	Omkar Vedak, Prasad Zavre, Abhijeet Todkar, Manoj Patil	The proposed system consists of two main stages: 1.Segmentation of hand 2. Recognition of hand sign. In this system, segmenting the hand using YCbCr colour space is firstly done and then processing the image through HOG and then it is provided to the model.	The system produces 88% accuracy. HOG is better than other descriptors because it can adapt to changing illuminations and rotation of objects. It does not consider an image as a whole, but divides it into smaller cells and then for the pixels within the cells edge or gradient direction histogram is calculated.	The final descriptor vector grows larger, thus taking more time to extract and to train using a given classifier.
6.	Recognition of ASL using Image Processing and Machine Learning, 2019 [7]	Sharmila Gaikwad, Akanksha Shetty, Akshaya Satam, Mihir	a] Image Acquisition b] Feature extraction c] Orientation detection c]gesture recognition d] text to speech	The system is not only can apply in family environment, but also can apply in public.	we may move on to recognizing words, from as large a dictionary as possible, from

		Rathod, Pooja Shah			Indian Sign Language.
7.	Hand gesture recognition in images and video,2010.[8]	Ilan Steinberg, Tomer M. London, Dotan Di Castro	Image processing technique ,cancellation of background, noise effect, it then extracts relevant features for classification and finally classifies the gesture features using a multiclass Support Vector Machine classifier.	97.8% accuracy-real time and offline classification.	SVM algorithm is not suitable for large data sets.

2.2 Existing System

2.2.1 Indian sign language interpreter using image processing and Machine Learning

In this methodology, the input is taken in RGB format. For feature extraction, HSV method is used and 3-layer CNN network is used for classification. CNN is lightweight as it requires lower computational power.

The images are acquired using a camera and then the features are extracted using HSV filtering. The skin tone values are selected to isolate the hands form the background. Then it is converted into the grayscale image to obtain the features of the hand. These images are then separated as training and test sets and then fed to the model.

For simplicity and to make the code lightweight for deploying in Raspberry Pi this model is chosen. Then the trained model uses classification of the various signs from the given dataset for recognition and the output is produced on the screen with the recognized letter.

Hardware used: Raspberry pi, Webcam
Software used: Python Language
[8]

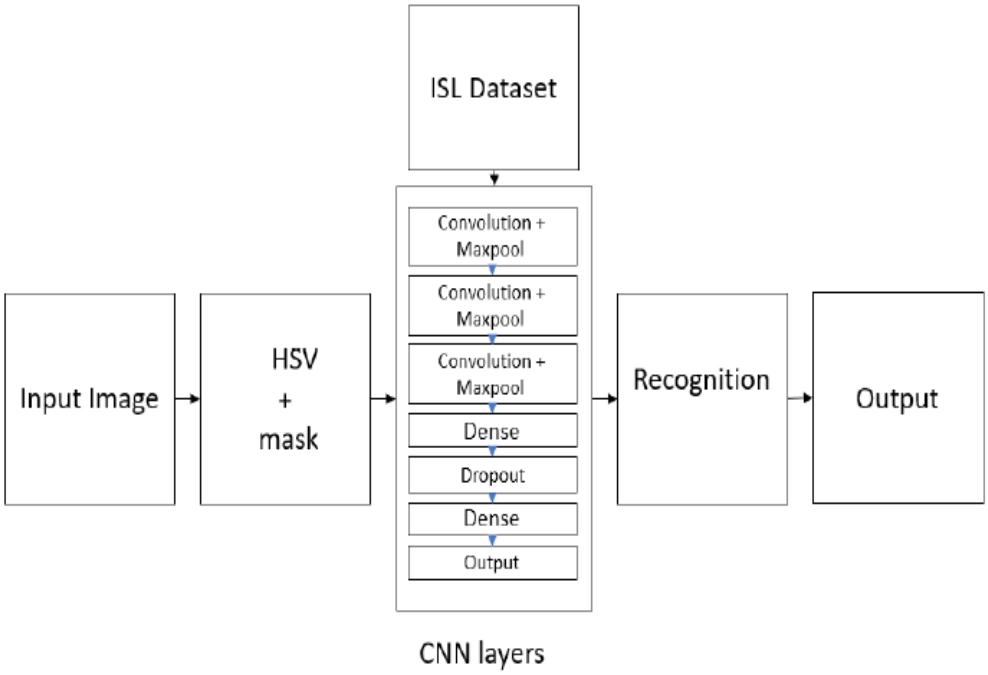

Fig 2.2: Design of Existing system

Name	Author	Methodology	Drawbacks	Rectifications
Sign Language Recognition Using Modified Convolutional Neural Network Model, 2018.[1]	Shubhendu Apoorv, Sudharshan Kumar Bhowmick and R Sakthi Prabha	HSV method for feature extraction is used and 3-layer CNN network is used for classification.	1. tanh activation function is used which makes activation dense and costly. 2. Model deployed in Raspberry-Pi	1. ReLu activation will be used since it is sparse and efficient. 2. Raspberry-pi does not replace the computer, and the processor is not as fast.

Table 2.2: Existing system comparison

Chapter 3

Proposed System

3.1 Problem Statement

Converting sign language to text using Machine Learning and Convolutional Neural Networks to help reduce the gap of communication between the deaf-dumb and others in the society.

3.2 Scope

We want to build a reliable program which acquires image data of sign language numbers(1-9) using webcam and converts it to text output. We are focusing on converting numbers in sign language to text. The task is accomplished using Image processing and Machine learning. Our proposed project can find its applicability in the day to day life for communication. The main goal of developing such a user interface is to make the life easier for those in need so that they don't feel left out and it is better than other user interfaces because of its compatibility and user friendly environment.

3.3 Hardware requirements:

Webcam, Laptop/PC

3.4 Software requirements:

TensorFlow:
It is an open source software library for numerical computation. TensorFlow is widely used in Machine Learning.

Keras:
It is a high-level neural networks library written in python that works as a wrapper to TensorFlow. It is used in cases where we want to quickly build and test the neural network with minimal lines of code.

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.

Chapter 4

Design of the System

4.1 Approach

4.1.1 Flowchart:

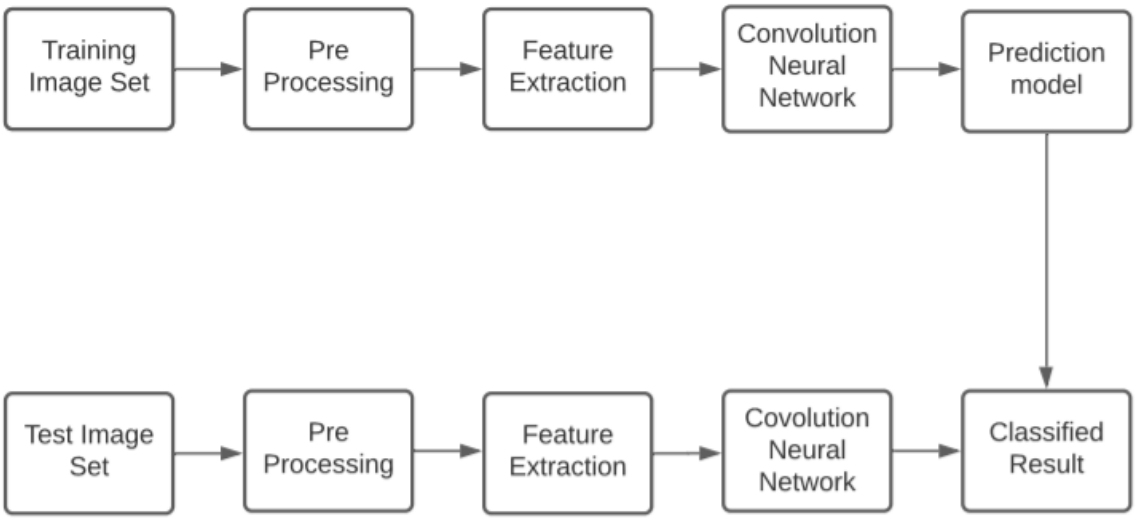

Fig 4.1 Flowchart

4.2 Approach

4.2.1 Block Diagram:

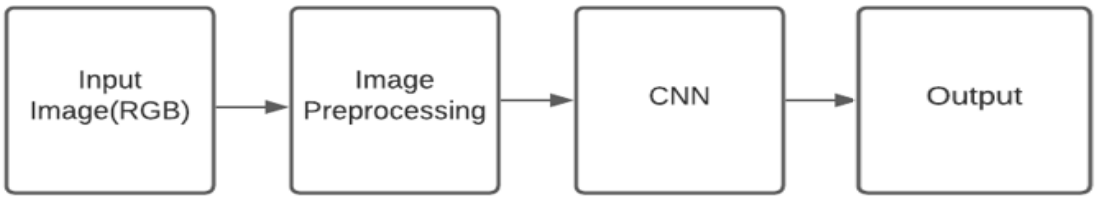

Fig 4.2 System architecture of Sign Language Translator

Dataset:

We have chosen Indian Sign Language data set [9] for our model, which consists of numbers from 1 to 9, each having 1200 images and a total of 10800 images. We will split the data for training and testing.

Input:

The input will be taken in RGB format using openCV

Image pre-processing:

The input will be converted to grayscale and by using threshold method, we will make the image easier to analyze. Then finally we apply Gaussian blur filter to our image which helps us extract various features of our image.

Convolutional Neural Network Layers:

After extracting features, we will train our model. Convolutional Neural Network (CNN) model takes image as an input, process it and classify it. The main objective of convolution is to extract features such as edges, colours, corners from the input. The proposed model has convolution layers and dense layers with ReLu (Rectified Linear Unit) activation. Then the accuracy of the model will be checked by predicting for the testing set of data.

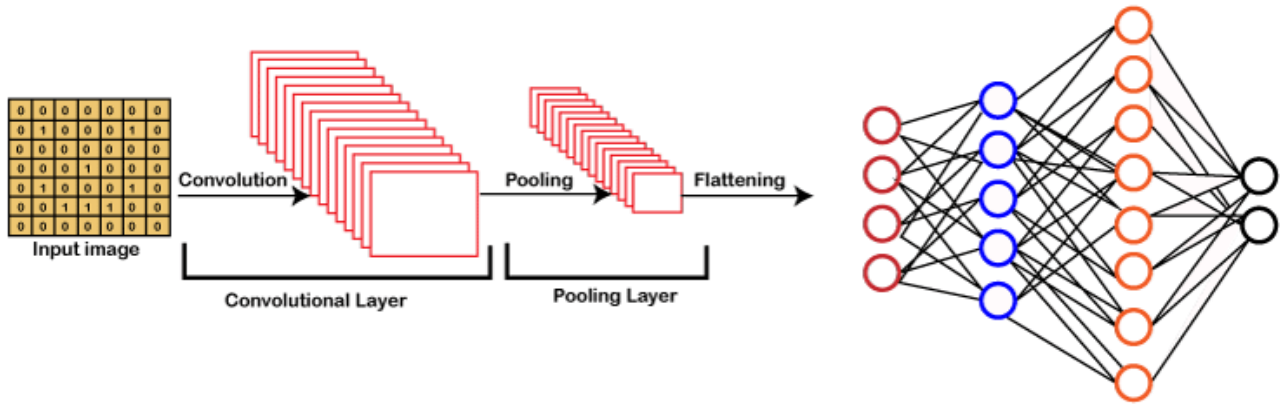

Fig 4.3 Working of Convolutional Neural Network Layers

Chapter 5

Conclusion and Future Scope

Conclusion and Future Scope:

This project will be identifying numbers in Sign Language from the corresponding gestures. Gesture recognition and sign language recognition has been a well researched topic for American Sign Language(ASL), but few research works have been published regarding Indian Sign Language(ISL). But instead of using high-end technology like gloves or motion detector, we aim to solve this problem using state of the art computer vision and machine learning algorithms. With the help of right technology we can make this world a better place for everyone. Scope of this project can be expanded to recognition of alphabets in Indian sign language by not restricting it to only numbers. We also wish to expand our project to apps and websites so that anyone can easily access it.

References

[1] Ujan Ziaie, Thomas M¨uller , Mary Ellen Foster , and Alois Knoll“A Na¨ive Bayes Munich,Dept. of Informatics VI, Robotics and Embedded Systems,Boltzmannstr. 3, DE-85748 Garching, Germany

[2] K. Tiku, J. Maloo, A. Ramesh and I. R., "Real-time Conversion of Sign Language to Text and Speech," *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, 2020,

[3] Ashikur Rahman, Nazmus Saquib “Application of Machine Learning Techniques for Real-Time Sign Language Detection using Wearable Sensors." 2020.

[4] Sayali Gore, Namrata Salvi, Swati Singh, “Conversion of Sign Language into Text Using Machine Learning Technique”

[5] Omkar Vedak, Prasad Zavre, Abhijeet Todkar, Manoj Patil “Sign Language Interpreter using Image Processing and Machine Learning.” 2019

[6] Sharmila Gaikwad, Akanksha Shetty, Akshaya Satam, Mihir Rathod, Pooja Shah. IJCSMC, Vol. 8, Issue. 3, March 2019 “Recognition of American Sign Language using Image Processing and Machine Learning”

[7] Hand Gesture Recognition in Images and Video Ilan Steinberg, Tomer M. London, Dotan Di Castro Department of Electrical Engineering, Technion- Israel Institute of Technology,” Hand Gesture Recognition in Images and Video Ilan Steinberg,”

[8] Sign Language Recognition Using Modified Convolutional Neural Network Model, 2018. Shubhendu Apoorv, Sudharshan Kumar Bhowmick and R Sakthi Prabha.

[9]Dataset reference: https://www.kaggle.com/vaishnaviasonawane/indian-sign-language-dataset

Acknowledgement

Success of a project like this involving high technical expertise, patience and massive support of guides, is possible when team members work to-gather. We take this opportunity to express our gratitude to those who have been instrumental in the successful completion of this project. We would like to appreciate the constant interest and support of our mentor **Ms. Snehal Nikalje** in our project and aiding us in developing a flair for the field of Machine Learning. We would always cherish the journey of transforming the idea of our project into reality. We would like to show our appreciation to **Ms. Snehal Nikalje** for her tremendous support and help, without whom this project would have reached nowhere. We would also like to thank our project coordinator **Mrs. Dakshayani R** for providing us with regular inputs about documentation and project timeline. A big thanks to our HOD **Dr. Lata Ragha** for all the encouragement given to our team. We would also like to thank our principal, **Dr. S. M. Khot**, and our college, **Fr. C. Ro-drigues Institute of Technology, Vashi**, for giving us the opportunity and the environment to learn and grow.

Project Group Members:

1. Sahil Raj, 1020129

2. Anushka Gaat, 1020140

3. Saatvik Inampudi, 1020148

4. Vithi Kulkarni, 1020164

Appendix A: Timeline Chart

Timeline Chart: