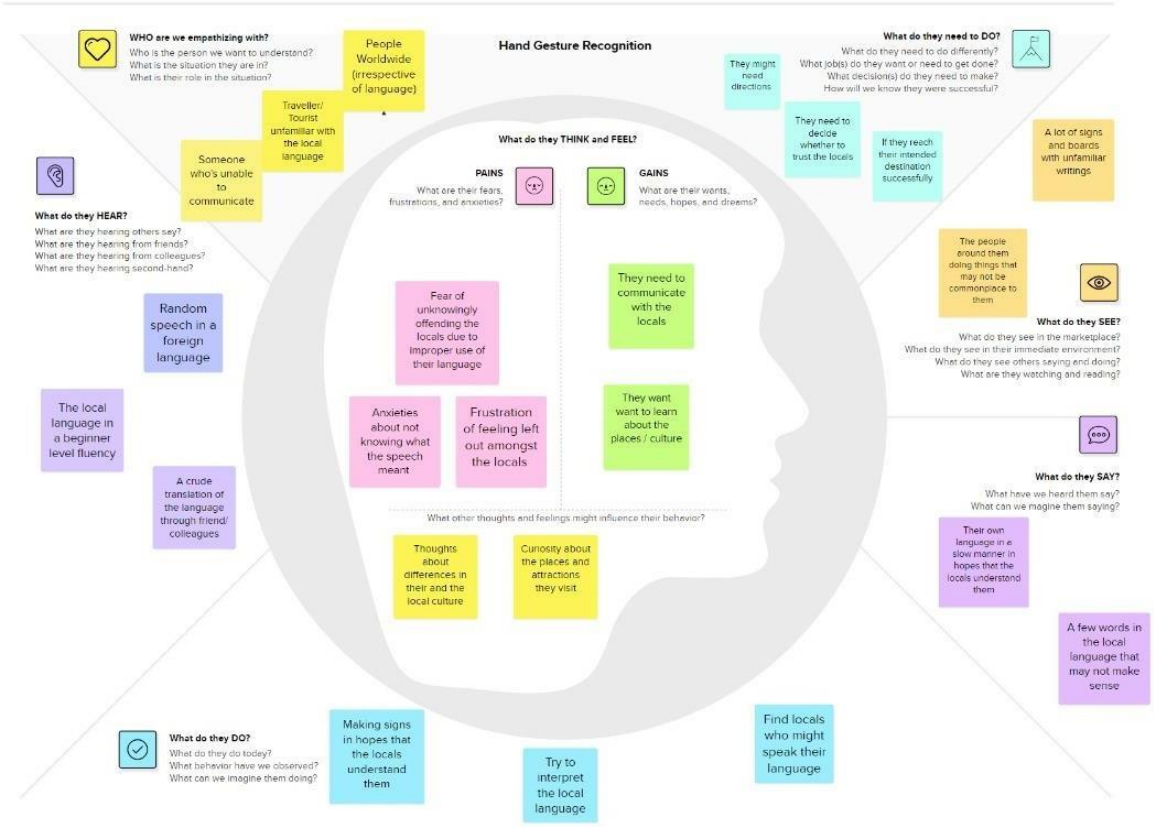
Project Report

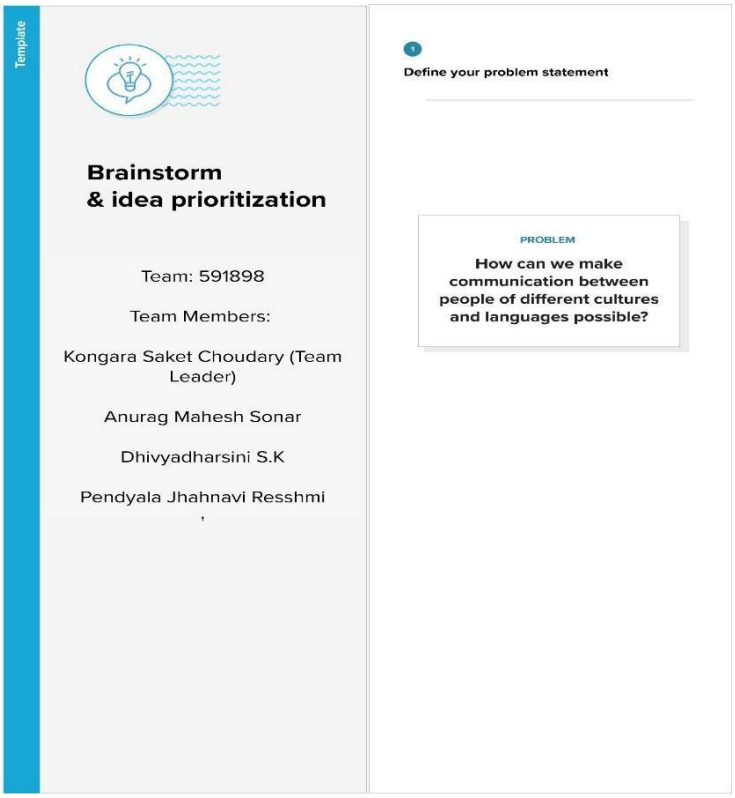
Date	18 November 2023
Team ID	Team-591898
Project Name	Hand Gesture Recognition

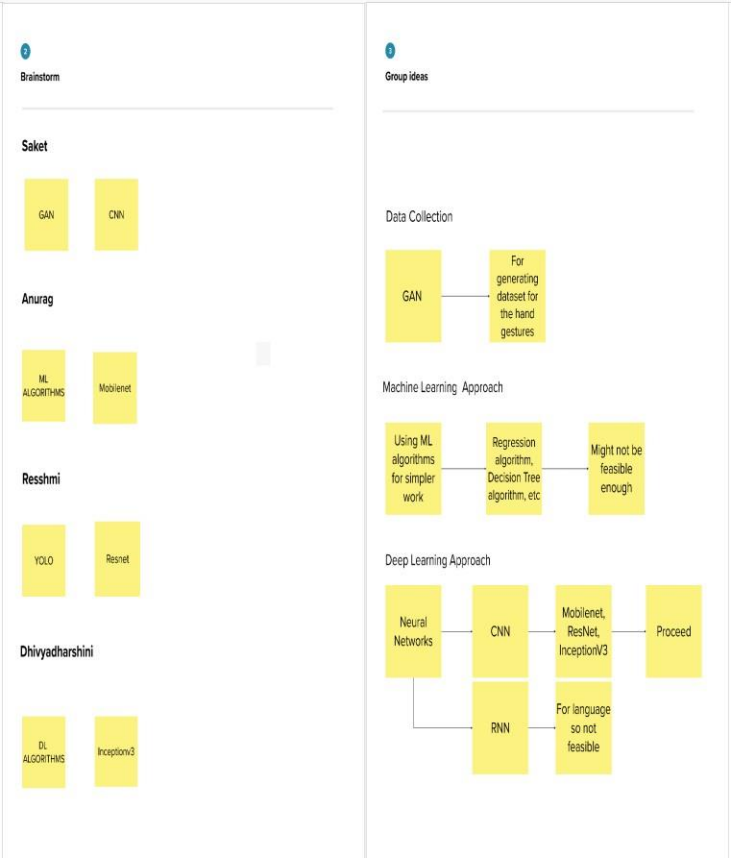
S.No.	***Parameter***	***Description***
1.	*Problem Statement (Problem to be solved)*	*How to make communication among people of different cultures and languages possible?*
2.	*Idea / Solution description*	*The program uses deep learning algorithms to detect and differentiate various hand gestures given to it through video input by the user with a webcam.*
3.	*Novelty / Uniqueness*	*The innovation that we propose is using a deep GAN (generative adversarial neural network) for collecting data and image classification algorithms for detecting hand gestures.*
4.	*Social Impact / Customer Satisfaction*	*To make technology such as Al and robotics accessible to everyone for efficient communication without language barriers through hand gestures.*
5.	*Business Model (Revenue Model)*	*Used in Industries like Healthcare & VR to help make resources accessible to everyone hassle-free.*

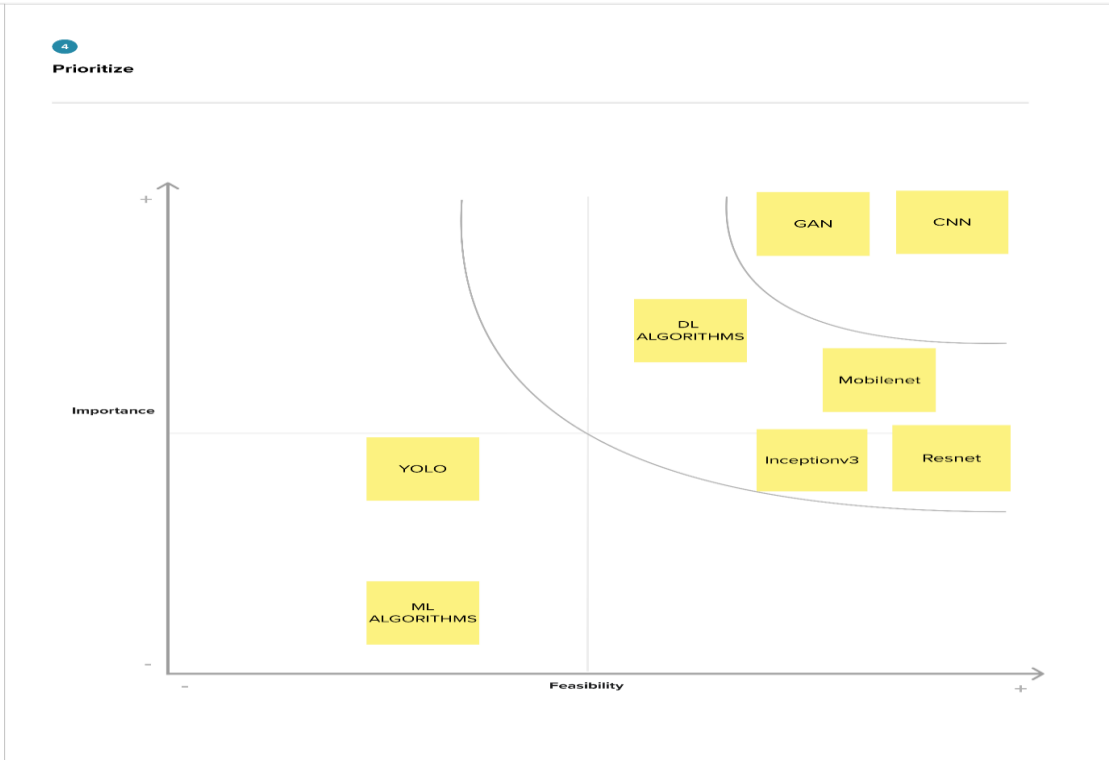
Empathy Map


Hand Gesture Recognition
WHO are we empathizing with?
Who is the person we want to understand?
What is the situation they are in?
What is their role in the situation?
People Worldwide (irrespective of language)
Traveller/ Tourist unfamiliar with the local language
Someone who's unable to communicate
What do they need to DO?
What do they need to do differently?
What job(s) do they want or need to get done?
What decision(s) do they need to make?
How will we know they were successful?
They might need directions
They need to decide whether to trust the locals
If they reach their intended destination successfully
A lot of signs and boards with unfamiliar writings
What do they THINK and FEEL?
PAINS
What are their fears, frustrations, and anxieties?
GAINS
What are their wants, needs, hopes, and dreams?
What do they HEAR?
What are they hearing others say?
What are they hearing from friends?
What are they hearing from colleagues?
What are they hearing second-hand?
Random speech in a foreign language
Fear of unknowingly offending the locals due to improper use of their language
They need to communicate with the locals
The people around them doing things that may not be commonplace to them
What do they SEE?
What do they see in the marketplace?
What do they see in their immediate environment?
What do they see others saying and doing?
What are they watching and reading?
The local language in a beginner level fluency
Anxieties about not knowing what the speech meant
Frustration of feeling left out amongst the locals
They want want to learn about the places / culture
A crude translation of the language through friend/ colleagues
What do they SAY?
What have we heard them say?
What can we imagine them saying?
What other thoughts and feelings might influence their behavior?
Thoughts about differences in their and the local culture
Curiosity about the places and attractions they visit
Their own language in a slow manner in hopes that the locals understand them
A few words in the local language that may not make sense
What do they DO?
What do they do today?
What behavior have we observed?
What can we imagine them doing?
Making signs in hopes that the locals understand them
Try to interpret the local language
Find locals who might speak their language

Brainstorming



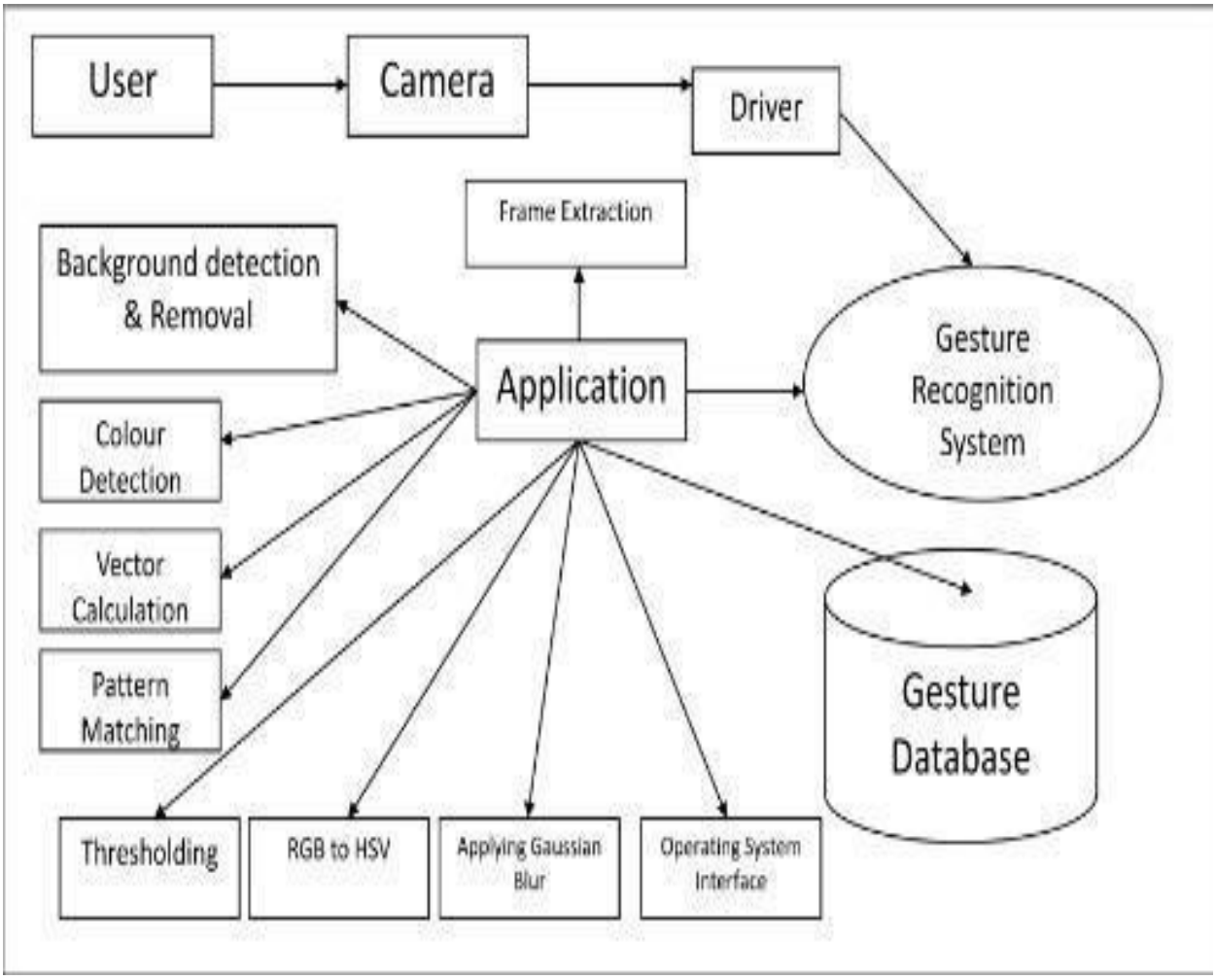




Requirements

- mediapipe 0.8.1
- OpenCV 3.4.2 or Later
- Tensorflow 2.3.0 or Later
- tf-nightly 2.5.0.dev or later (Only when creating a TFLite for an LSTM model)
- scikit-learn 0.23.2 or Later (Only if you want to display the confusion matrix)
- matplotlib 3.3.2 or Later (Only if you want to display the confusion matrix)

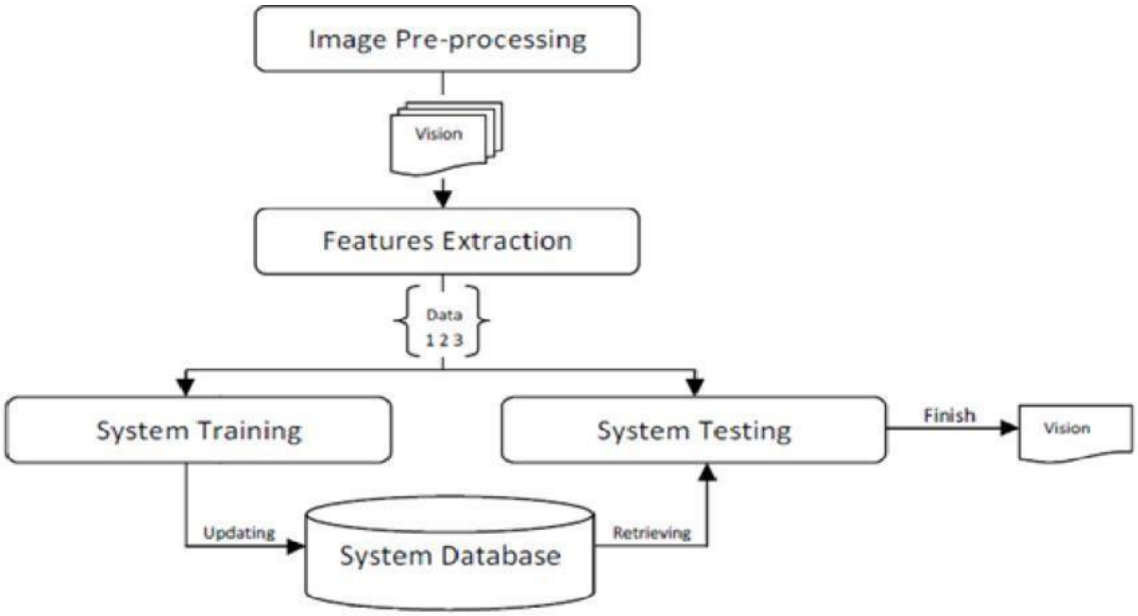
Data Flow Diagram:



User Stories:

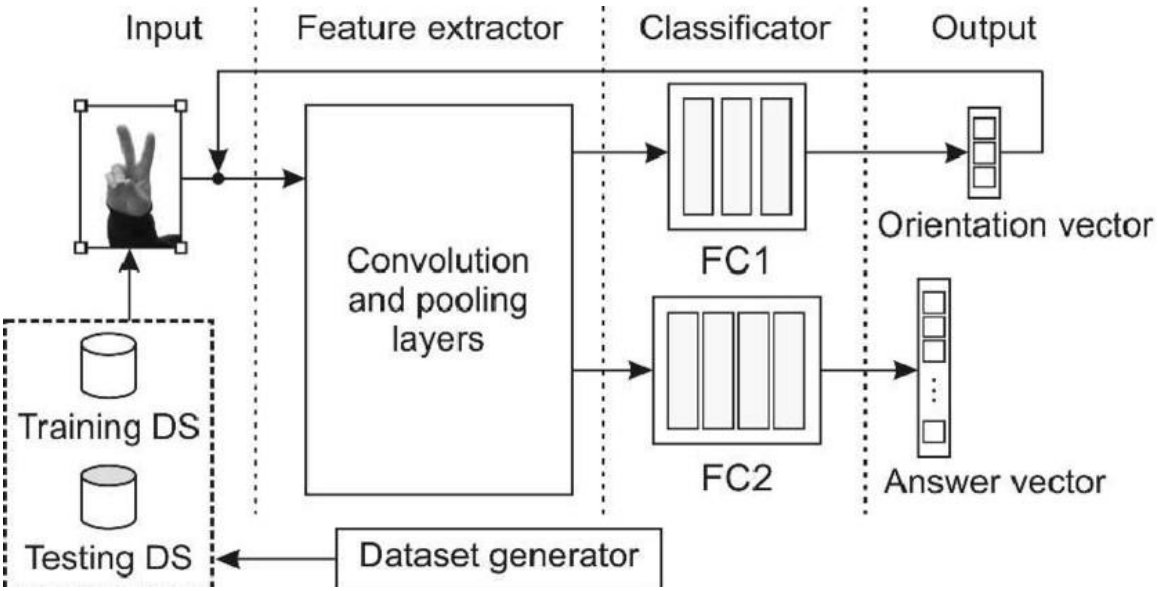
User Type	*User Story Number*	*User Story*	*Acceptance Criteria*
Deaf User	*USN-1*	*I want to be able to control my smart TV with hand gestures so that I can watch TV without having to use a remote control.*	*I can turn on/off and change the channels using hand gestures.*
Gamer	*USN-2*	*I want to be able to control my video games with hand gestures so that I can have a more immersive and engaging gaming experience.*	*I can move my character and attack by hand gestures only.*
Surgeon	*USN-3*	*I want to be able to control my surgical robots with hand gestures so that I can operate on patients more precisely and efficiently.*	*I can move the robot's arm and scalpels precisely using hand gestures.*
Teacher	*USN-4*	*I want to be able to use hand gestures to control my presentation software so that I can interact with my students more effectively.*	*I can advance, change the view, and zoom in/out of the slides using hand gestures.*
Smart Home User	*USN-5*	*I want to be able to turn on and off lights, adjust the thermostat, and lock and unlock doors with hand gestures.*	*I can use multiple appliances in my house using hand gestures.*
VR User	*USN-6*	*I want to interact with the virtual world using hand gestures, such as picking up and throwing objects, and opening and closing doors.*	*I can interact with the virtual environment using hand gestures.*

Solution Architecture:



Reference:

Technical Architecture:



Components and Technology:

S.No	*Component*	*Description*	*Technology*
1.	*User Interface*	*How user interacts with application e.g.Web UI*	*HTML, CSS, JavaScript / Angular Js /React Js etc.*
2.	*Application Logic-1*	*Logic for a process in the application*	*Java / Python*
3.	*Application Logic-2*	*Logic for a process in the application*	*IBM Watson STT service*
4.	*Application Logic-3*	*Logic for a process in the application*	*IBM Watson Assistant*
5.	*Database*	*Data Type, Configurations etc.*	*MySQL, NoSQL, etc.*
7.	*File Storage*	*File storage requirements*	*IBM Block Storage or Other StorageService or Local Filesystem*
8.	*Machine Learning Model*	*Purpose of Machine Learning Model*	*Object Recognition Model, etc.*
9.	*Infrastructure (Server)*	*Application Deployment on Local System Local Server Configuration:*	*Local, Cloud Foundry, Kubernetes, etc.*

Product Backlog, Sprint Schedule, and Estimation (4 Marks)

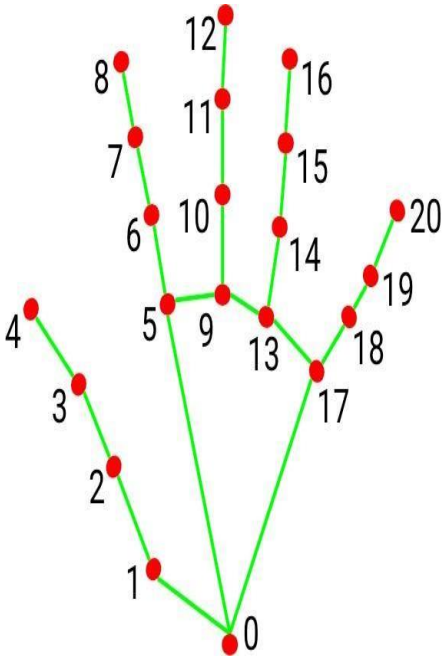
Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1		USN-1	Creation of Code	2	High	Saket
Sprint-1		USN-2	Creation of Functionality	2	High	Saket
Sprint-2		USN-3	Collection of Data	2	High	Anurag
Sprint-2		USN-4	Run Tests	1	Medium	Anurag

Project Tracker, Velocity & Burndown Chart: (4 Marks)

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	20	5 Days	10 Nov 2023	15 Nov 2023	20	15 Nov 2023
Sprint-2	20	5 Days	15 Nov 2023	20 Nov 2023	20	18 Nov 2023

Working Principle

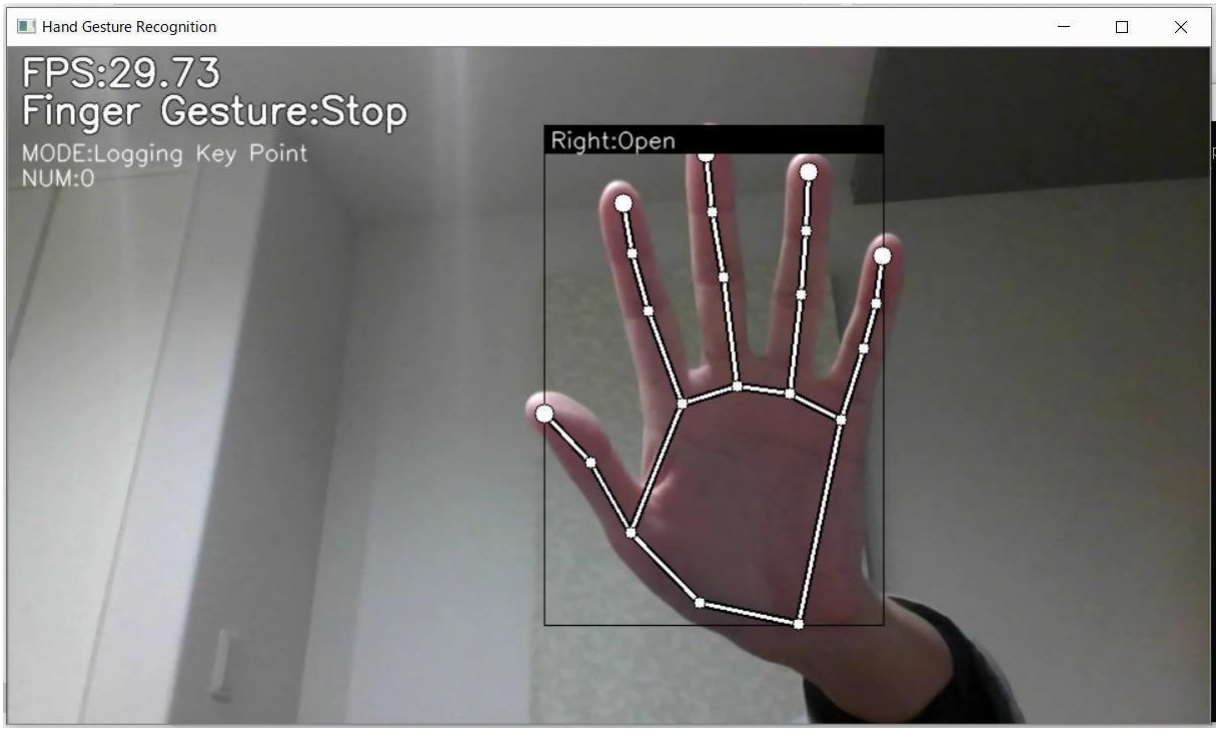
The code assigns key points to different points of the hand and for each hand gesture, it records the positions of the key points for each label (registered hand gesture). The code then reads the key point positions to recognize which hand gesture is being shown.



0. WRIST
1. THUMB_CMC
2. THUMB_MCP
3. THUMB_IP
4. THUMB_TIP
5. INDEX_FINGER_MCP
6. INDEX_FINGER_PIP
7. INDEX_FINGER_DIP
8. INDEX_FINGER_TIP
9. MIDDLE_FINGER_MCP
10. MIDDLE_FINGER_PIP
11. MIDDLE_FINGER_DIP
12. MIDDLE_FINGER_TIP
13. RING_FINGER_MCP
14. RING_FINGER_PIP
15. RING_FINGER_DIP
16. RING_FINGER_TIP
17. PINKY_MCP
18. PINKY_PIP
19. PINKY_DIP
20. PINKY_TIP

Training

Press "k" to enter the mode to save key points（displayed as 「MODE: Logging Key Point」）



If you press "0" to "9", the key points will be added to "**model/keypoint_classifier/keypoint.csv**"

Model training

Open "**keypoint_classification.ipynb**" in Jupyter Notebook and execute from top to bottom.

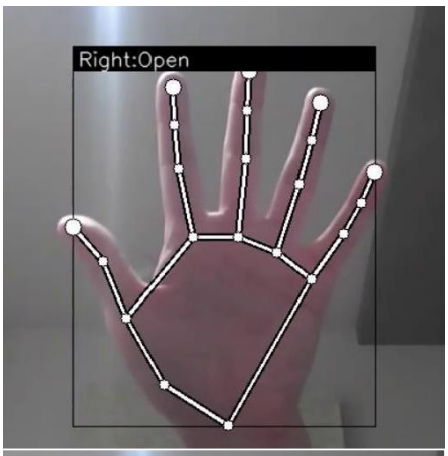
To change the number of training data classes, change the value of "NUM_CLASSES = 3"

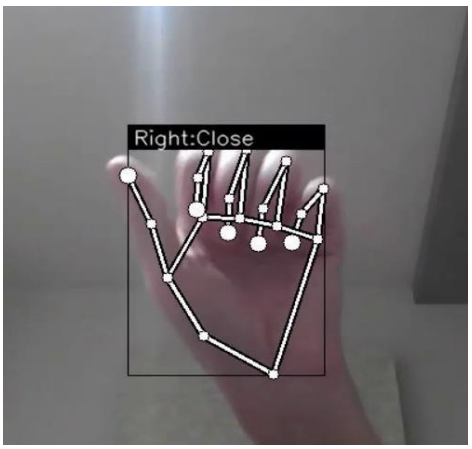
and modify the label of "**model/keypoint_classifier/keypoint_classifier_label.csv**" as appropriate.

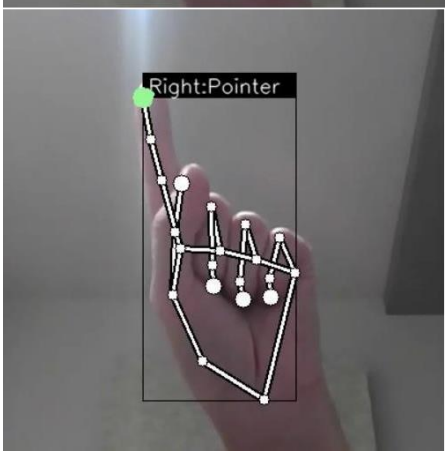
Model Performance Testing:

S.No.	Parameter	Values
1.	Metrics	**Mediapipe** **Tensorflow**
2.	Tune the Model	Collect data by saving keypoints using the webcam

Results







Future Scope

It can provide options for integrationinto existing products by collaborating with established companies in the relevant industry.

Appendix

Dataflow:https://www.researchgate.net/figure/Architecture-of-gesture-recognition-system-5_fig4_284626785

Source code: https://github.com/kinivi/hand-gesture-recognition-mediapipe

GitHub Repo: https://github.com/smartinternz02/SI-GuidedProject-609262-1698050431

Project Demo Link: https://youtu.be/7sywpZ7o2gg?si=Az_RwwuXpvynKlVy