GESTURE BASED TOOL FOR STERILE BROWSING AND RADIOLOGY IMAGES

A MINI PROJECT REPORT

Submitted by

MILIRNA G	**(AC19UCS068)**
MEGHANA SHRAVANI A	**(AC19UCS066)**
RANJITH KUMAR M	**(AC19UCS094)**
KUBERA SELVAN D	**(AC19UCS059)**

in partial fulfilment for the award of the degree

of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

ADHIYAMAAN COLLEGE OF ENGINEERING

DR. M.G.R NAGAR, HOSUR-635130

ANNA UNIVERSITY: CHENNAI 600 025

OCTOBER 2021

A GESTURE-BASED TOOL FOR STERILE BROWSING OF RADIOLOGY IMAGES

CONTENTS

CHAPTER 6 – PROJECT PLANNING & SCHEDULING

CHAPTER 7-CODING & SOLUTIONING(Explain the features added in the project along with code)

CHAPTER 8-TESTING

CHAPTER 9-RESULTS

CHAPTER 10-ADVANTAGES & DISADVANTAGES

CHAPTER 11-CONCLUSION

CHAPTER 12-FUTURE SCOPE

CHAPTER 13-APPENDIX

ABSTRACT

The use of doctor-computer interaction devices in the operation room (OR) requires new modalities that support medical imaging manipulation while allowing doctors' hands to remain sterile, supporting their focus of attention, and providing fast response times. This project presents " a vision-based hand gesture capture and recognition system that interprets in real-time the user's gestures for navigation and manipulation of images in an electronic medical record (EMR) database. Navigation and other gestures are translated to commands based on their temporal trajectories, through video capture. It was tested during a brain biopsy procedure. In the in vivo experiment, this interface prevented the surgeon's focus shift and change of location while achieving a rapid intuitive reaction and easy interaction. Data from two usability tests provide insights and implications regarding human-computer interaction based on nonverbal conversational modalities. Neural Network to first train the model on the images of different hand gestures, like showing numbers with fingers as 0,1,2,3,4,5. Then we made a web portal using Flask where user can input any image on which he wants to perform the operations. After uploading the image, our portal uses the integrated webcam to capture the video frame using OpenCV

Keywords: Sterile browsing

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

Computer information technology is increasingly penetrating into the hospital domain. A major challenge involved in this process is to provide doctors with efficient, intuitive, accurate and safe means of interaction without affecting the quality of their work. Keyboards and pointing devices, such as a mouse, are today's principal method of human—computer interaction. However, the use of computer keyboards and mice by doctors and nurses in intensive care units (ICUs) is a common method for spreading infections.

In this project, we suggest the use of hand gestures as an alternative to existing interface techniques, offering the major advantage of sterility. Even though voice control also provides sterility, the noise level in the operating room (OR) deems it problematic.

In this work we refer to gestures as a basic form of non-verbal communication made with the hands. Psychological studies showed that young children use gestures to communicate before they learn to talk. Manipulation, as a form of gesticulation, is often used when people speak to each other about some object. Naturalness of expression, non-encumbered interaction, intuitiveness and high sterility are all good reasons to replace the current interface technology (e.g., keyboard, mouse, and joystick) with more natural interfaces.

This project presents a video-based hand gesture capture and recognition system used to manipulate magnetic resonance images (MRI) within a graphical user interface. A hand gesture vocabulary of commands was selected as being natural in the sense that each gesture is cognitively associated with the notion or command that is meant to represent it. For example, moving the hand left represents a "turn left" command.

The operation of the gesture interface was tested at the Washington Hospital Center in Washington, DC. Two operations were observed in the hospital's neurosurgery department and insights regarding the suitability of a hand gesture system was obtained. To our knowledge, this is the first time that a hand gesture recognition system was successfully implemented in an "in vivo" neurosurgical biopsy. A sterile human—machine interface is of supreme importance because it is the means by which the surgeon controls medical information avoiding contamination of the patient, the OR and the surgeon.

In two brain surgeries at the Neurosurgery OR at the Washington Hospital Center, procedures were observed by the authors to gain insights about the use of current technologies and how they affect the quality of the surgeon’s performance.

We found that:

(a) surgeons kept their focus of attention between the patient and the surgical point of interest on the touch-screen navigation system;

(b) a short distance between the surgeon and the patient was maintained during most of the surgery;

(c) the surgeon had to move close to the main control wall to discuss and browse through the patient’s MRI images. The hand gesture control system is “developed by the authors helped the doctor to remain in place during the entire operation, without any need to move to the main control wall since all the commands were performed using hand gestures.

In this project we have used Convolutional Neural Network to first train the model on the images of different hand gestures, like showing numbers with fingers as 0,1,2,3,4,5. Then we made a web portal using Flask where user can input any image on which he wants to perform the operations. After uploading the image, our portal uses the integrated webcam to capture the video frame using OpenCV. The gesture captured in the video frame is compared with the Pre-trained model and the gesture is identified. If the prediction is 0 - then images is converted into rectangle, 1 - image is Resized into (200,200), 2 - image is rotated by -45°, 3 - image is blurred , 4 - image is Resized into (400,400) , 5 - image is converted into grayscale.

Many of these deficiencies may be overcome by introducing a more natural human-computer interaction mode into the hospital environment. The bases of human-human communication are speech, hand and body gestures, facial expression, and eye gaze.

Some of these concepts have been exploited in systems for improving medical procedures. In FAce MOUSe, a surgeon can control the motion of the laparoscope by simply making the appropriate face gesture, without hand or foot switches or voice input. Current research to incorporate hand gestures into doctor-computer interfaces appeared in Graetzel et al. They developed a computer vision system that enables surgeons to perform standard mouse functions (pointer movement and button presses) with hand gestures.

Another aspect of gestures is their capability to aid handicapped people by offering a natural alternative form of interface and serving as a diagnostic tool. 6 Wheelchairs, as mobility aids, have been enhanced as robotic vehicles able to recognize the user's commands through handgestures.

CHAPTER 2

LITERATURE REVIEW

2.1 REVIEWS ON AUTOMATED IRRIGATIONS

Mokhar M. Hasan, Pramod K. Mishra, (2012). "Robust Gesture Recognition Using Gaussian Distribution for Features Fitting" International Journal of Machine Learning and Computing,Vol. 2(3). Data from two usability tests provide insights and implications regarding human computer interaction based on nonverbal conversational modalities.

Simei G. Wysoski, Marcus V. Lamar, Susumu Kuroyanagi, Akira Iwata, (2002)."A Rotation Invariant Approach On Static Gesture Recognition Using Boundary Histograms And Neural International Journal of Artificial Intelligence & Applications (IJAIA), Vol.3, No.4, July 2012.Al Rock Paper Scissor with hand gesture is an AI based python project in which you can detect hand and fingers and with the help of your fingers co-ordinates you can figure out if it.

.

Joseph J. LaViola Jr., (1999). "A Survey of Hand Posture and Gesture Recognition Techniques and Technology", Master Thesis, Science and Technology Center for Computer Graphics and Scientific Visualization USA. Convolutional Neural Network to first train the model on the images of like showing numbers with fingers as 0.1.2.3.4.5.

E. Stergiopoulou, N. Papamarkos. (2009). "Hand gesture recognition using a neural network shape fitting technique," Elsevier Engineering Applications of Artificial Intelligence, vol. 22(8), pp. 1141-1158, doi: 10.1016/j.engappai.2009.03.008. G. R. S. Murthy, R. S. Jadon. (2009). Requires new modalities that support medical imaging manipulation while allowing doctors' hands to remain sterile, supporting their focus of attention, and providing fast response times.

CHAPTER 3

SYSTEM ANALYSIS

3.1 EXISTING SYSTEM

- Surgical imaging services allow physicians and surgeons to see internal images of the body, including bones and organs, prior to and during surgical procedures. These images help reduce risk and speed recovery.
- Imaging in intensive care unit (ICU) is integral to patient management. The portable chest radiograph is the most commonly requested imaging examination in ICU, and, despite its limitations, it significantly contributes to the decision-making process.

Medical Gesture Interfaces

- By the early 1990's scientists, surgeons and other experts were beginning to draw together state-of-the-art technologies to develop comprehensive image-guidance systems for surgery, such as the StealthStation.3 This is a free-hand stereo-tactic pointing device, in which a position is converted into its corresponding location in the image space of a high-performance computer monitor. In a setting like the OR, touch screen displays are often used, and must be sealed to prevent the buildup of contaminants. They should also have smooth surfaces for easy cleaning with common cleaning solutions. These requirements are often overlooked in the busy OR environment

- Many of these deficiencies may be overcome by introducing a more natural human-computer interaction mode into the hospital environment. The bases of human-human communication are speech, hand and body gestures, facial expression, and eye gaze. Some of these concepts have been exploited in systems for improving medical procedures. In FAce MOUSe,2 a surgeon can control the motion of the laparoscope by simply making the appropriate face gesture, without hand or foot switches or voice input.

- To incorporate hand gestures into doctor-computer interfaces appeared in Graetzel et al.4 They developed a computer vision system that enables surgeons to perform standard mouse functions (pointer movement and button presses) with hand gestures. Another aspect of gestures is their capability to aid handicapped people by offering a natural alternative form of interface and serving as a diagnostic tool.6 Wheelchairs, as mobility aids, have been enhanced as robotic vehicles able to recognize the user’s commands through hand gestures.

- The stereotype of the radiologist as a boring/faceless part of the medical team hiding in a dark room is not accurate. Radiologists are consultants - often at the centre of the clinical environment with a unique opportunity to provide critical information that frequently directs management.

- Lectures in radiology are often content driven, which can be boring in the absence of background knowledge. Diagnostic Radiology is intellectually stimulating, detail oriented, and can be very gratifying. It is perfect for those that love problem solving.

The CT scanner is capable of imaging several parts of the human body which include the skeleton, internal organs and blood vessels. Moreover, CT scanning is mostly used for the brain, neck, spine and chest. Another important use of the CT scanner is to locate tumours and broken bones within the patient.

However, an X-ray scan is sometimes not conclusive and therefore the CT scanner is used to aid further diagnosis. The CT scanning process involves the patient resting on a movable horizontal bed, which travels through the entrance of the scanner, with the gantry containing the X-ray units rotating around cross-sections of the body. At this stage, the patient is advised to remain still in order to obtain clear CT images.

During the scanning procedure, the radiologist will be in an adjoining room to operate the scanner and be shielded from the radiation. The CT scan can take up to 20 minutes – depending on the condition of the patient – and the scans are subsequently analysed by a computer. Also, CT scanning is a non-invasive, pain-free technique.

DISADVANTAGES:

- There are few excuses to give the surgeon if you cannot screen the area of interest because of poor positioning of the patient and equipment.

- This equipment does not support sterile browsing as there is a need of using system input devices which might be contaminated.

3.2. REFERENCES

Robust Gesture Recognition Using Gaussian Distribution

Mokhar M. Hasan, Pramod K. Mishra, (2012). “Robust Gesture Recognition Using Gaussian Distribution for Features Fitting” International Journal of Machine Learn and Computing,Vol. 2(3)

A Review of Vision Based Hand Gestures Recognition

Simei G. Wysoski, Marcus V. Lamar, Susumu Kuroyanagi, Akira Iwata, (2002)."A Rotation Invariant Approach On Static Gesture Recognition Using Boundary Histograms And Neural International Journal of Artificial Intelligence & Applications (IJAIA), Vol.3, No.4, July 2012

Divide the scaled normalized hand image into Blocks of intensity features.

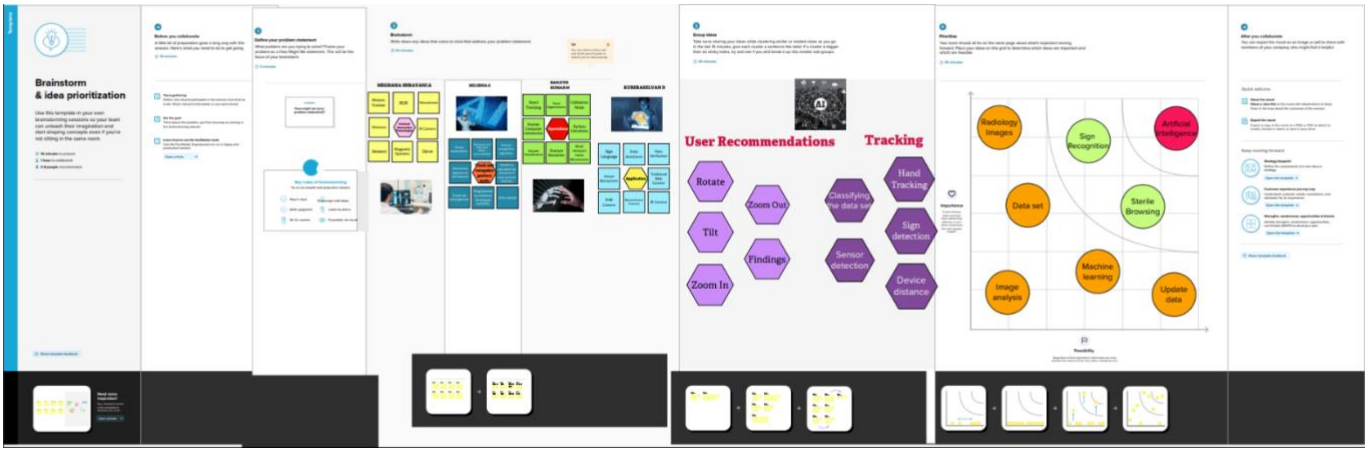
Joseph J. LaViola Jr., (1999). “A Survey of Hand Posture and Gesture Recognition Techniques and Technology”, Master Thesis, Science and Technology Center for Computer Graphics and Scientific Visualization USA

Robot control application

E. Stergiopoulou, N. Papamarkos. (2009). “Hand gesture recognition using a neural network shape fitting technique,” Elsevier Engineering Applications of Artificial Intelligence, vol. 22(8), pp. 1141-1158, doi: 10.1016/j.engappai.2009.03.008. G. R. S. Murthy, R. S. Jadon. (2009). .

3.3 EMPATHY MAP CANVAS

3.4 IDEATION AND BRAINSTROMING



3.2 PROPOSED SYSTEM

A hand gesture system for MRI manipulation in an EMR image database was tested during a brain biopsy surgery. This system is a real-time hand-tracking recognition technique based on color and motion fusion. In an in vivo experiment, this type of interface prevented surgeon's focus shift and change of location while achieving, rapid intuitive interaction with an EMR image database. In addition to allowing sterile interaction with EMRs, the hand gesture interface provides:

(i) ease of use—the system allows the surgeon to use his/her hands, their natural work tool.

(ii) rapid reaction—nonverbal instructions by hand gesture commands are intuitive and fast (In practice, the system can process images and track hands at a frame-rate of 150 Hz, thus, responding to the surgeon's gesture commands in real-time),

(iii) an unencumbered interface—the proposed system does not require the surgeon to attach a microphone, use head-mounted (body contact) sensing devices or to use foot pedals, and

(iv) distance control—the hand gestures can be performed up to 5 meters from the camera and still be recognized accurately. The results of two usability tests (contextual and individual interviews) and a satisfaction questionnaire indicated that

The system provided a versatile method that can be used in the OR to manipulate medical images in real-time and in a sterile manner. We are now considering the addition of a body posture recognition system to increase the functionality of the system, as well as visual tracking of both hands to provide a richer set of gesture commands.

For example, pinching the corners of a virtual image with both hands and stretching the arms would represent an image zoom-in action. In addition, we wish to assess whether a stereo camera will increase the gesture recognition accuracy of the system. A more exhaustive comparative experiment between our system and other human–machine interfaces, such as voice, is also left for future work.

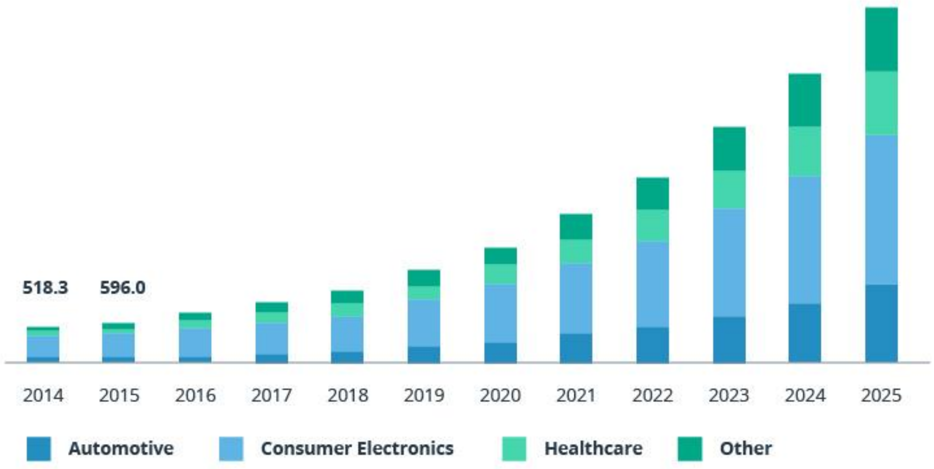
Gesturing is a natural and intuitive way to interact with people and the environment. So it makes perfect sense to use hand gestures as a method of human-computer interaction (HCI).

But there are quite a few challenges, starting from needing to wave your hands in front of your small smartphone screen and ending with the complex machine learning algorithms needed to recognize more than a simple thumbs up. Is the juice worth the squeeze? Let’s find out, starting from definitions and moving to the technical details.

The need for gesture recognition technology

Markets and Markets that the gesture recognition market will reach $32.3 billion in 2025, up from $9.8 billion in 2020. Today’s top producers of gesture interface products are, unsurprisingly, Intel, Apple, Microsoft, and Google. The key industries driving mass adoption of touchless tech are automotive, healthcare, and consumer electronics.

Gesture recognition market in China, 2014–2025



Keep in mind that hand tracking and gesture recognition are not the same things. Both technologies are supposed to use hands for human-machine interaction (HMI) without touching, switching, or employing controllers. Sometimes, systems for hand tracking and gesture recognition require the use of markers, gloves, or sensors, but the ideal system requires nothin but a human hand.

Systems employing gesture recognition technology are only capable of distinguishing specific gestures: thumbs up, wave, peace sign, rock sign, etc. Hand tracking is more complex: it provides more variability in the HMI, since it tracks hand size, finger position, and other characteristics. The number of potential interactions with digital objects is limitless, but overlapping, occlusion, and interpretation issues occur. While AI in the gesture recognition system is only trained to identify a limited number of gestures and is less flexible than hand tracking technology, it doesn’t suffer from the same issues.

Why may people want to use gestures instead of just touching or tapping a device? A desire for contactless sensing and hygiene concerns are the top drivers of demand for touchless technology. Gesture recognition can also provide better ergonomics for consumer devices. Another market driver is the rise of biometric systems in many areas of people’s lives, from cars to homes to shops.

During the coronavirus pandemic, it’s not surprising that people are reluctant to use touchscreens in public places. Moreover, for drivers, tapping a screen can be dangerous, as it distracts them from the road. In other cases, tapping small icons or accidentally clicking on the wrong field increases frustration and makes people look for a better customer experience.

Real-time hand gesture recognition for computer interactions is just the next step in technological evolution, and it’s ideally suited for today’s consumer landscape. Besides using gestures when you cannot conveniently touch equipment, hand tracking can be applied in augmented and virtual reality environments, sign language recognition, gaming, and other use cases.

The high cost of touchless sensing products is one of the major challenges of this technology, along with the complexity of software development for HGR. To create a robust system that detects hand positions, a hand tracking solution requires the implementation of advanced machine learning and deep learning algorithms, among other things.

Gesture recognition provides real-time data to a computer to make it fulfill the user's commands. Motion sensors in a device can track and interpret gestures, using them as the primary source of data input. A majority of gesture recognition solutions feature a combination of 3D depth-sensing cameras and infrared cameras together with machine learning systems. Machine learning algorithms are trained based on labeled depth images of hands, allowing them to recognize hand fingerpositions.

ADVANTAGES

- Major advantage of this tool is that it helps to maintain the sterility of the environment.
- It is also easy to use and is quicker than the existing methods to browse images.
- It can also be performed even if the surgeon is a bit far away from the system, this helps to save time.
- The tool does not need the person using it to have an apparatus or any devices on them to use it.
- They can simply move their hands to browse through the images. ease of use—the system allows the surgeon to use his/her hands, their natural work tool.
- rapid reaction—nonverbal instructions by hand gesture commands are intuitive and fast (In practice, the system can process images and track hands at a frame-rate of 150 Hz, thus, responding to the surgeon's gesture commands in real-time),
- an unencumbered interface—the proposed system does not require the surgeon to attach a microphone, use head-mounted (body contact) sensing devices or to use foot pedals, and
- distance control—the hand gestures can be performed up to 5 meters from the camera and still be recognized accurately. The results of two usability tests (contextual and individual interviews) and a satisfaction questionnaire indicated that

The system provided a versatile method that can be used in the OR to manipulate medical images in real-time and in a sterile manner. We are now considering the addition of a body posture recognition system to increase the functionality of the system, as well as visual tracking of both hands to provide a richer set of gesture commands.

3.3 ARCHITECTURE DESIGN

The Tracking Algorithm After a short calibration process, where a probability color model of the doctor's hand is built, images of the surgeon's hand gesturing are acquired by video-camera and each image is back-projected using a color model. The hand is then tracked by an algorithm which segments it from the background using the color model back-projection and motions.This is followed by black/white thresholding, and a sequence of opening and closing morphological operations resulting in a set of components (blobs) in the image.

The location of the hand is represented by the 2D coordinates of the centroid of the biggest blob in the current image. "Gibson" Image Browser The "Gibson" image browser is a 3D visualization medical tool that enables examination of images, such as: MRIs, CT scans and X-rays. The images are arranged over a multiple layer 3D cylinder. The image of interest is found through rotating the cylinder in the four cardinal directions. To interface the gesture recognition routines with the "Gibson" system, information such as the centroid of the hand, its size, and orientation are used to enable screen operations in the graphical user interface.

Hand Tracking and Operation Modes Gesture operations are initiated by a calibration mode in which a skin color model of the user's hand or glove, under local lighting, is constructed. In a browse mode, superimposed over the image of the camera's scene is a rectangular frame called the "neutral area." Movements of the hand across its boundary constitute directional browser commands. When a doctor/surgeon wishes to browse the image database, the hand is moved rapidly out of the "neutral area" toward any of four directions, and then back again. When such a movement is detected, the displayed image is moved off the screen and replaced by a neighbour image.

To evoke a zoom mode, the open palm of the hand is rotated within the "neutral area" clockwise/counter clockwise (zoom-in/zoom-out). To avoid the tracking of unintentional gestures, the user may enter a "sleep mode" by dropping the hand. To re-arouse the system the user waves the hand in front of the camera. The selection of these gestures was designed to be intuitive, expressing the "natural" feeling of the user.

For example, the left/right/up/down gestures evoke the actions used to turn pages in a book left/right, or flip notepad pages up/down. The rotation gesture (zoomin/zoom-out commands) reminds one of a radio knob to increase or decrease volume. Dropping the hand (stoptracking command) is associated to the idea of 'stop-playing', while the waving gesture ("wake-up" command) is associated with 'greeting a new person'

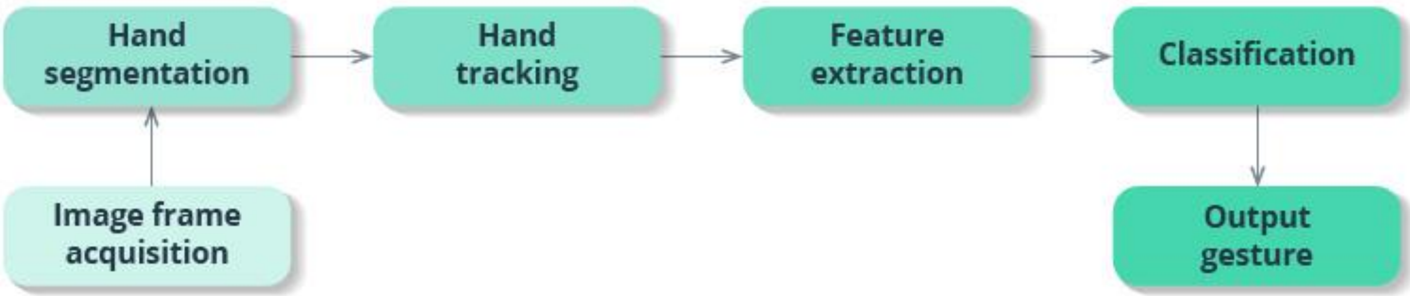
Gesture recognition provides real-time data to a computer to make it fulfill the user's commands. Motion sensors in a device can track and interpret gestures, using them as the primary source of data input. A majority of gesture recognition solutions feature a combination of 3D depth-sensing cameras and infrared cameras together with machine learning systems. Machine learning algorithms are trained based on labeled depth images of hands, allowing them to recognize hand and finger positions.

Gesture recognition consists of three basic levels:

- **Detection.** With the help of a camera, a device detects hand or body movements, and a machine learning algorithm segments the image to find hand edges and positions.

- **Tracking.** A device monitors movements frame by frame to capture every movement and provide accurate input for data analysis.

- **Recognition.** The system tries to find patterns based on the gathered data. When the system finds a match and interprets a gesture, it performs the action associated with this gesture. Feature extraction and classification in the scheme below implements the recognition functionality.



Many solutions use vision-based systems for hand tracking, but such an approach has a lot of limitations. Users have to move their hands within a restricted area, and these systems struggle when hands overlap or aren't fully visible. With sensor-based motion tracking, however, gesture recognition systems are capable of recognizing both static and dynamic gestures in real time.

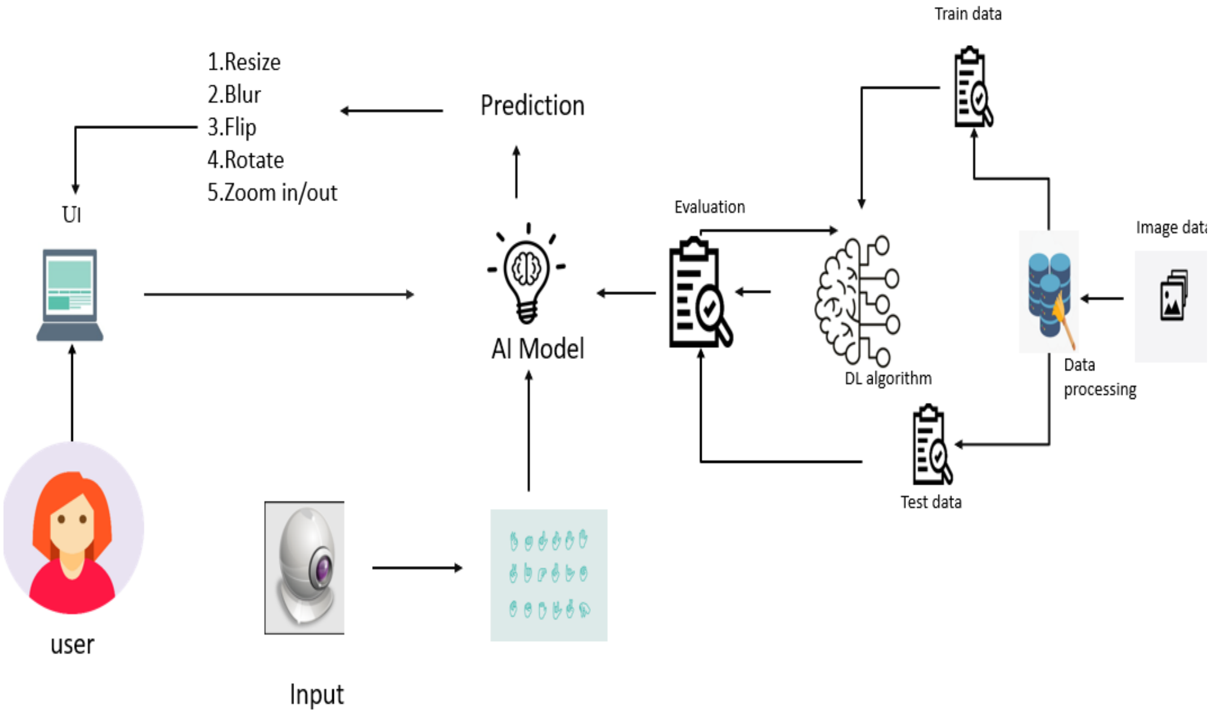
In sensor-based systems, depth sensors are used to align computer-generated images with real ones. Leap motion sensors are also used in hand tracking to detect the number and three-dimensional position of fingers, locate the center of the palm, and determine hand orientation. Processed data provides insights on fingertip angles, distance from the palm center, fingertip elevation, coordinates in 3D space, and more. The hand gesture recognition system using image processing looks for patterns using algorithms trained on data from depth and leap motion sensors:

1. The system distinguishes a hand from the background using color and depth data. The hand sample is further divided into the arm, wrist, palm, and fingers. The system ignores the arm and wrist since they don't provide gesture information.
2. Next, the system obtains information about the distance from the fingertips to the center of the palm, the elevation of the fingertips, the shape of the palm, the position of the fingers, and so on.
3. Lastly, the system collects all extracted features into a feature vector that represents a gesture. A hand gesture recognition solution, using AI, matches the feature vector with various gestures in the database and recognizes the user's gesture.

Depth sensors are crucial for hand tracking technology since they allow users to put aside specialized wearables like gloves and make HCI more natural.

Intel has recently released a suite of depth and tracking technologies called RealSense, providing the developer community with open-source tools for a variety of languages and platforms. The Intel RealSense Depth Camera D455 with Lidar, stereo depth, tracking, and coded light capabilities provides a high level of gesture recognition and a longer range for HMI. With the help of a camera like this, dynamic hand gesture recognition systems can be applied to various use cases, from robotics and drones to 3D scanning and people tracking.

SOLUTION ARCHITECHTURE



Hand gesture perception, while being natural for people, is quite a challenge for computer vision. Hands often get in the way of each other as seen by a camera (think of a fist or handshake) and lack high-contrast patterns.

To develop an HGR system, AI algorithms are trained to recognize labeled data and predict unknown data based on the developed model. A hand tracking database is the first step in AI training. To create a training data set, depth cameras are used to segment a specific element from the background. High-quality segmentation helps AI distinguish between left and right hands, individual fingers, etc. The higher the quality of data sets and the more annotations they include, the higher the accuracy of dynamic hand gesture recognition with computer vision.

At CVPR 2019, Google announced a new approach to hand perception implemented in MediaPipe — a cross-platform framework for building multimodal machine learning pipelines. With this new method, real-time performance can be achieved even on mobile devices.

- Palm detector
- Hand landmark
- Gesture recognizer

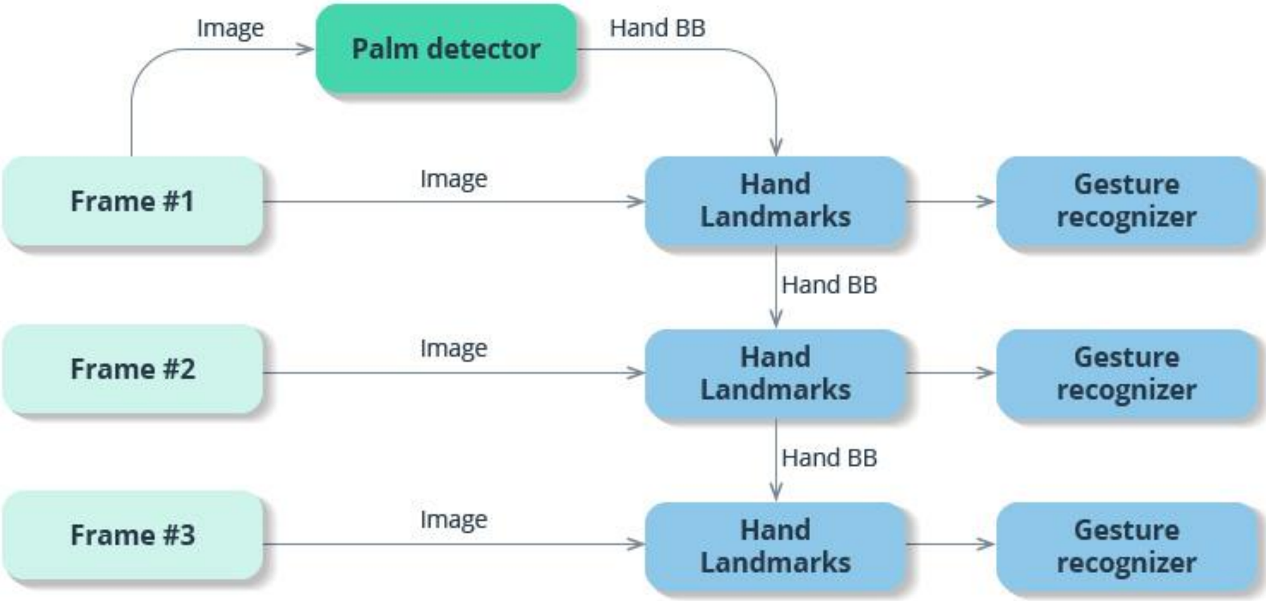


Image
Palm detector
Hand BB
Frame #1
Image
Hand Landmarks
Gesture recognizer
Hand BB
Frame #2
Image
Hand Landmarks
Gesture recognizer
Hand BB
Frame #3
Image
Hand Landmarks
Gesture recognizer

Source: Google AI Blog

Since this hand tracking and gesture recognition pipeline is open-source, developers have a complete stack for prototyping and innovating on top of Google's model. Extensive datasets for AI training will help increase the number of gestures recognized accurately and make the system more robust.

EXPERIMENTAL INVESTIGATIONS :

We found that many hospitals rely on mouse and keyboard to browse the images that are obtained during different surgeries, scans, etc. This can contaminate the environment with various infections thus compromising the sterility.

Various technologies have been developed to overcome this issue and one such technology this hand gesture system for MRI manipulation in an EMR image database called "Gestix" wa s tested during a brain biopsy surgery.

This system is a real-time hand-tracking recognition technique based on color and motion fusion. In an in vivo experiment, this type of interface prevented the surgeon's focus shift and change of location while achieving rapid intuitive interaction with an EMR image database.

In addition to allowing sterile interaction with EMRs, the " hand gesture interface provides:

1. ease of use—the system allows the surgeon to use his/her hands, their natural work tool;

2. rapid reaction—nonverbal instructions by hand gesture commands are intuitive and fast

3. an unencumbered interface—the proposed system does not require the surgeon to attach a microphone, use head-mounted (body-contact) sensing devices or to use foot pedals

4. distance control—
the hand gestures can be performed up to 5 meters from the camera and still be recognized accurately.

3.4 DESCRIPTION OF MODULES

In this project we have used the following modules

- Data processing module
- Trained data module
- Dataset classification and prediction module
- Gesture detection and relevant task Module

3.5 EXPLANATION OF MODULES

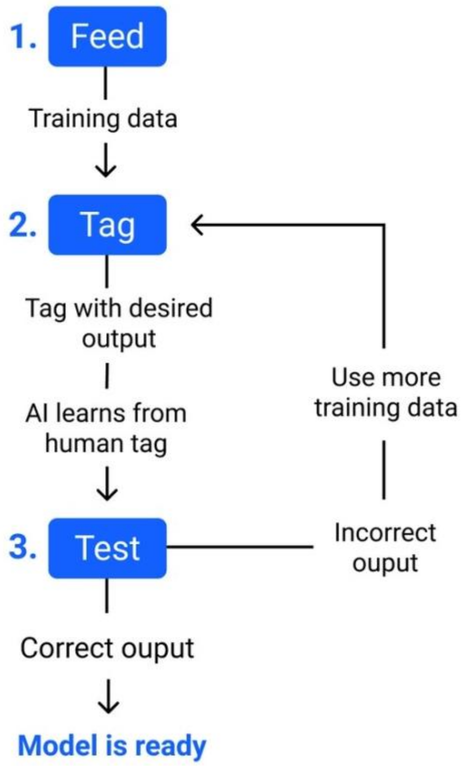
Data processing module

Artificial intelligence (AI)-driven big data processing technologies based on pattern recognition, machine learning, and deep learning, are intensively applied to deal with the large-scale heterogeneous data. However, challenges still exist in the development of AI-driven big data processing.

In computer vision and image processing, increasingly more databases and data streams have been transmitted and collected. One of the biggest challenges in the massive image/video data analysis is to develop energy efficient and real-time methods to extract useful information out of the colossal amount of data being generated every second. In speech signal processing, benefitting from the help of 'Big Data' and new AI technology, a lot of progress has also been made in speech processing area. How to build a condition robustness speech processing system using limited labeled data is still a direction to emphasize studying in the future.

Trained data module

Training data (or a training dataset) is the initial data used to train machine learning models.Training datasets are fed to machine learning algorithms to teach them how to make predictions or perform a desired task.



Dataset classification and prediction module

Classification is to identify the category or the class label of a new observation. First, a set of data is used as training data. The set of input data and the corresponding outputs are given to the algorithm. So, the training data set includes the input data and their associated class labels. Using the training dataset, the algorithm derives a model or the classifier. The derived model can be a decision tree, mathematical formula, or a neural network. In classification, when unlabeled data is given to the model, it should find the class to which it belongs. The new data provided to the model is the test data set.

Another process of data analysis is prediction. It is used to find a numerical output. Same as in classification, the training dataset contains the inputs and corresponding numerical output values. The algorithm derives the model or a predictor according to the training dataset. The model should find a numerical output when the new data is given. Unlike in classification, this method does not have a class label. The model predicts a continuous-valued function or ordered value.

Gesture detection and relevant task Module

Human-to-Robot handovers are useful for many Human-Robot Interaction scenarios. It is important to recognize when a human intends to initiate handovers, so that the robot does not try to take objects from humans when a handover is not intended. We pose the handover gesture recognition as a binary classification problem in a single RGB image. Three separate neural network modules for detecting the object, human body key points and head orientation, are implemented to extract relevant features from the RGB images, and then the feature vectors are passed into a deep neural net to perform binary classification.

CHAPTER 4
REQUIREMENTS ANALYSIS

4.1 FUNCTIONAL REQUIREMENTS

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	Identifying User Gestures	The user gestures are identified using the images of gestures capturedby the camera
FR-2	Deployment in Cloud	The trainedDeep Learning Model is deployed in cloud, which could be accessed anywhere around the world

FR-3	User Interface	The user interface, which helps in the Human Computer Interaction is designed
FR-4	Gestures related to the Application Domain	The model should be trained with the gestures related to the application domain.

4.2. NON-FUNCTIONAL REQUIREMENTS

FR No.	**Non-Functional Requirement**	**Description**
NFR-1	Usability	The user interface which acts as an intermediate between the user and the DL Model which is deployed in the cloud
NFR-2	Security	The model deployed in the cloud should be accessible only by the approved users and it should be inaccessible by the attackers or the terrorists
NFR-3	Reliability	The tool or the systemis 95% reliability for a year
NFR-4	Performance	The tool or the system should respond with the accurate response within 4-5 seconds
NFR-5	Availability	The model deployed in the cloud must be available to 99.8% of the people over a month during working hours

NFR-6	Scalability	The model deployed in the cloud must be accessible by over 10,00,000 people trying to access it using the user interface

CHAPTER 5

PROJECT DESIGN

5.1. DATA FLOW DIAGRAM

Dataflow is a model which defines how the data i.e. values gets travelled across the ML prediction model, which is shown in fig 11. In this project, the data flow starts from in-person data collection from thefarming fields. Here in-person could be considered as wireless sensor network which will collection data from field from listening to changes in the environment, which is farm field.

The collected data is then processed to build the dataset. At this stage, the dataset is inconsistent with the data integrity that could be later used in ML algorithm. Hence, data cleaning is performed, in order to make the data consistent with the ML prediction model. The cleaned data in the dataset is then fitted into Machine Learning Algorithm. The fitted dataset is splitted into trainingand testing dataset with the required proportion.

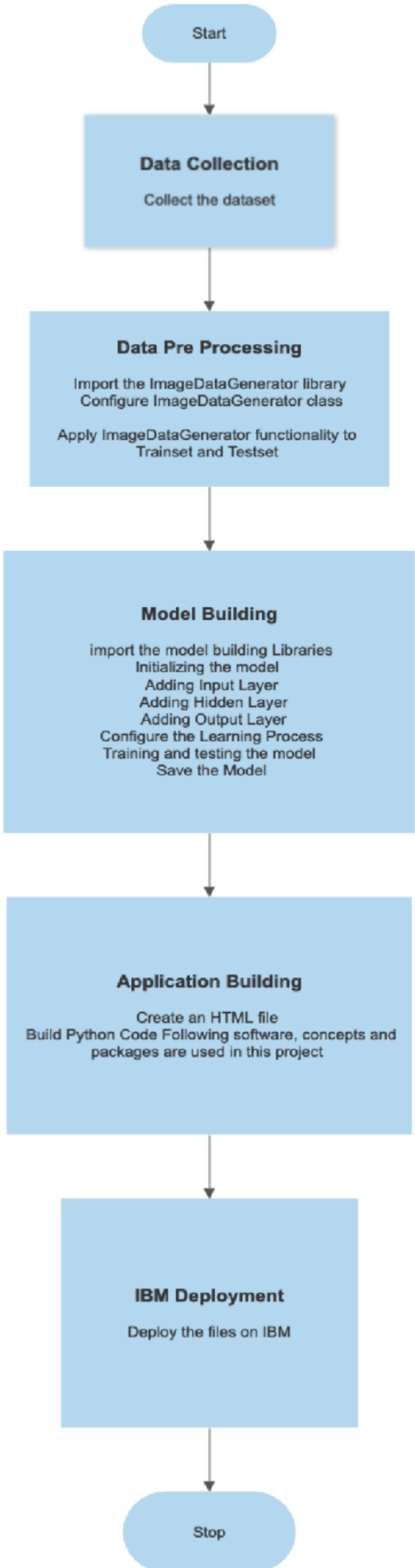
Training dataset is used to train the ML model

User interacts with the UI (User Interface) to upload the image as input.

Depending on the different gesture inputs different operations are applied to the input image.

Once model analyses the gesture, the prediction with operation applied on image is showcased on the UI. To accomplish this, we have to complete all the activities and tasks listed below:

- Data Collection.
  Collect the dataset or Create the dataset
- Data Pre processing
  - Import the ImageDataGenerator library
  - Configure ImageDataGenerator class
  - Apply ImageDataGenerator functionality to Trainset and Testset
- Model Building ○ Import the model building Libraries
  - Initializing the model
  - Adding Input Layer
  - Adding Hidden Layer
  - Adding Output Layer
  - Configure the Learning Process
  - Training and testing the model
  - Save the Model
- Application Building
  - Create an HTML file
  - Build Python Code Following software, concepts and packages are used in this project
- Anaconda navigator
- Python packages:
  - open anaconda prompt as administrator
  - Type “pip install TensorFlow” (make sure you are working on python 64 bit)
  - Type “pip install opencv-python”
  - Type “pip install flask”


Start
Data Collection
Collect the dataset
Data Pre Processing
Import the ImageDataGenerator library
Configure ImageDataGenerator class
Apply ImageDataGenerator functionality to Trainset and Testset
Model Building
import the model building Libraries
Initializing the model
Adding Input Layer
Adding Hidden Layer
Adding Output Layer
Configure the Learning Process
Training and testing the model
Save the Model
Application Building
Create an HTML file
Build Python Code Following software, concepts and packages are used in this project
IBM Deployment
Deploy the files on IBM
Stop

5.2. USER STORIES

TITLE: A GESTURE-BASED TOOL FOR STERILE BROWSING OF RADIOLOGY IMAGES
USER REQUIREMENTS
A sterile human—machine interface is of supreme importance because it is the means by which the surgeon controls medical information avoiding contamination of the patient, the OR and the surgeon.
ACCEPTANCE CRITERIA
Surgeons kept their focus of attention between the patient and the surgical point of interest on the touch-screen navigation system. A short distance between the surgeon and the patient was maintained during most of the surgery. The surgeon had to move close to the main control wall to discuss and browse through the patient's MRI images.

CHAPTER 6
PROJECT PLANNING & SCHEDULING

6.1. SPRINT PLANNING AND ESTIMATION

S.NO	ACTIVITY TITLE	ACTIVITY DESCRIPTION	DURATION
1	Understanding the project requirement	Assign the team members and create repository in the Github, Assign the task to each members and teach how to use and open and class the Github and IBM career education.	1 WEEK
2	Starting of project	Advice students to attend classes of IBM portal create and develop an rough diagram based on project description and gather of information on AI and IBM project and team leader assign task to each member of the project.	1 WEEK
3	Attend class	Team members and team lead must watch and learn from classes provided by IBM and NALAYATHIRAN and must gain access of MIT license for their project.	4 WEEK
4	Budget and scope of project	Budget and analyze the use of AI in the project and discuss with team for budget prediction to predict the favorability for the customer to buy.	1 WEEK

6.2. SPRINT DELIVERY SCHEDULE

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
		USN-1	To download the	2	High	MILIRNA G

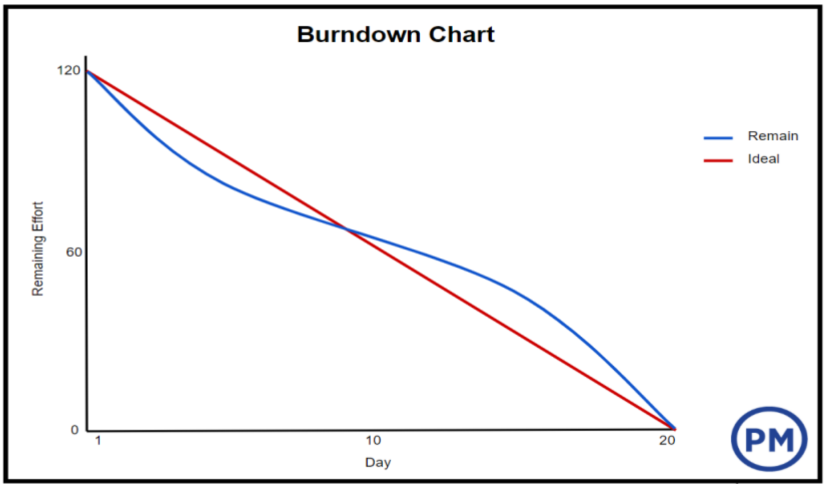
Sprint-1	Data Collection		data and apply image preprocessing on the data. Import the library,			MEGHANA SHRAVANI A RANJITH KUMAR M KUBERA SELVAN D
Sprint-2	Model Building	USN-2	Importing the necessary Model Building Libraries and initialize the model	1	Medium	MILIRNA G MEGHANA SHRAVANI A RANJITH KUMAR M KUBERA SELVAN D
Sprint-3	Application Building	USN-3	Create HTML Pages and build a python code and then run the application	2	High	MILIRNA G MEGHANA SHRAVANI A RANJITH KUMAR M KUBERA SELVAN D
Sprint-4	Train The Model on IBM	USN-4	To train the model on IBM and integrate it with the flask Application.	2	High	MILIRNA G MEGHANA SHRAVANI A RANJITH KUMAR M KUBERA SELVAN D

Velocity:

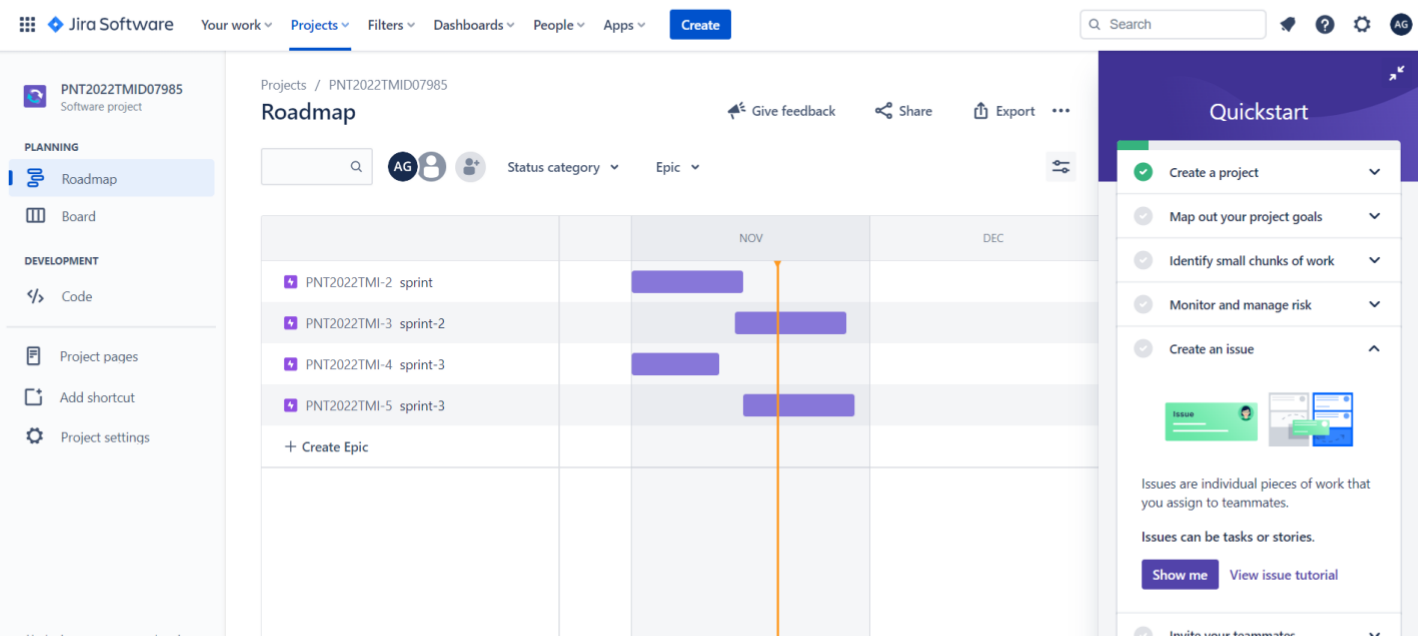
$Team\ Velocity$ = Σ Sprint 1 + Sprint 2 + ⋯ $Total\ Sprint$ = 20 + 15 + 10 + 5 4 = 12.5 AV = $Team\ Velovity\ Duration$ = 12.5 6 = 2.08

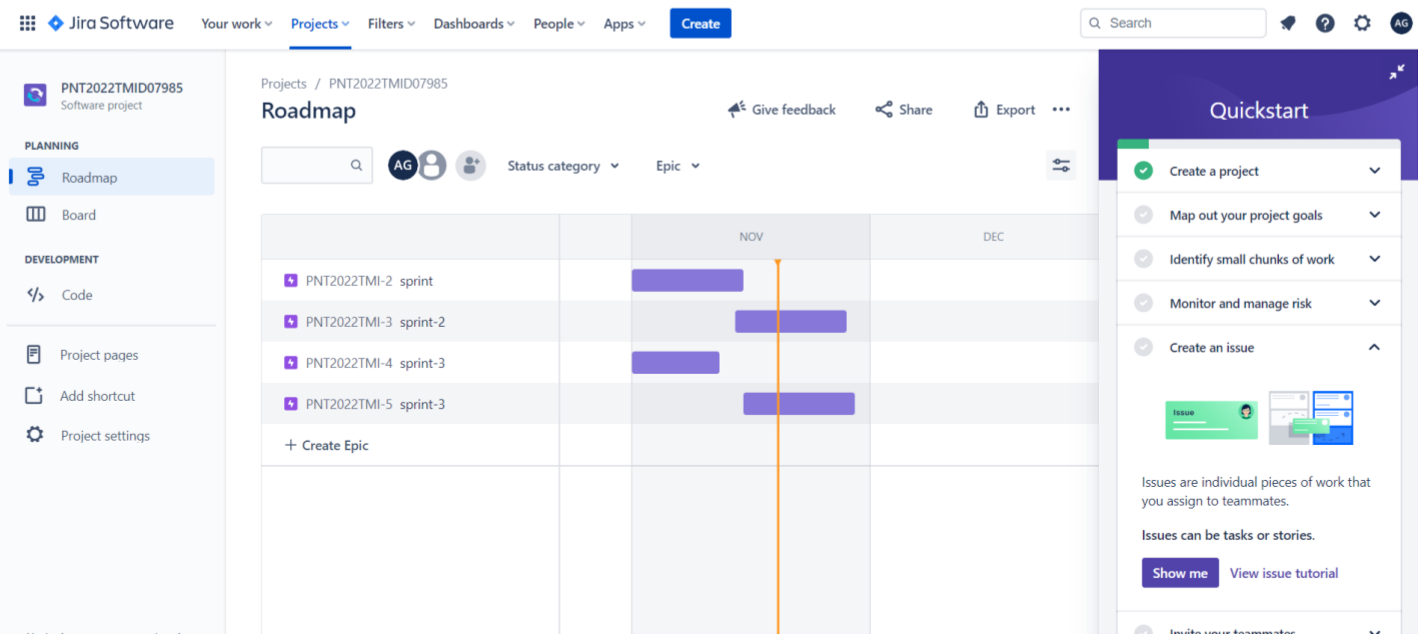
Burndown Chart:

A burn down chart is a graphical representation of work left to do versus time. It is often used in agile software development methodologies such as Scrum. However, burn down charts can be applied to any project containing measurable progress over time



6.3. Reports from JIRA



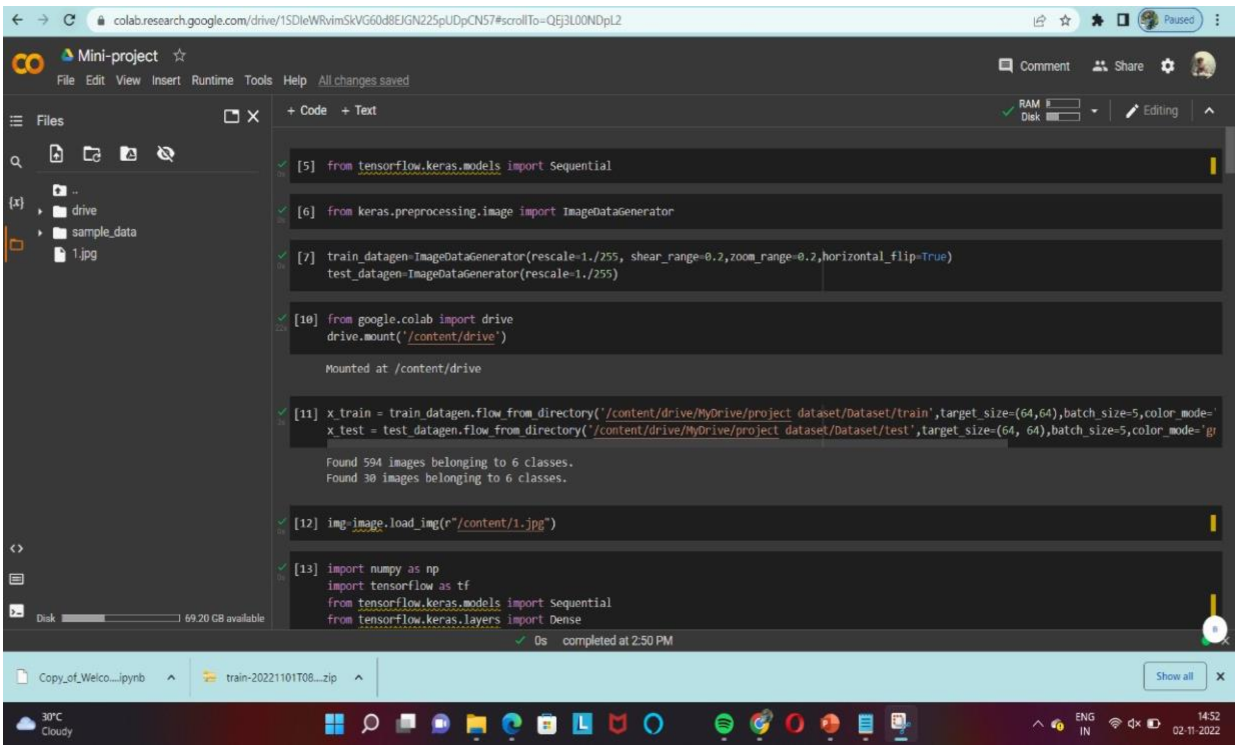


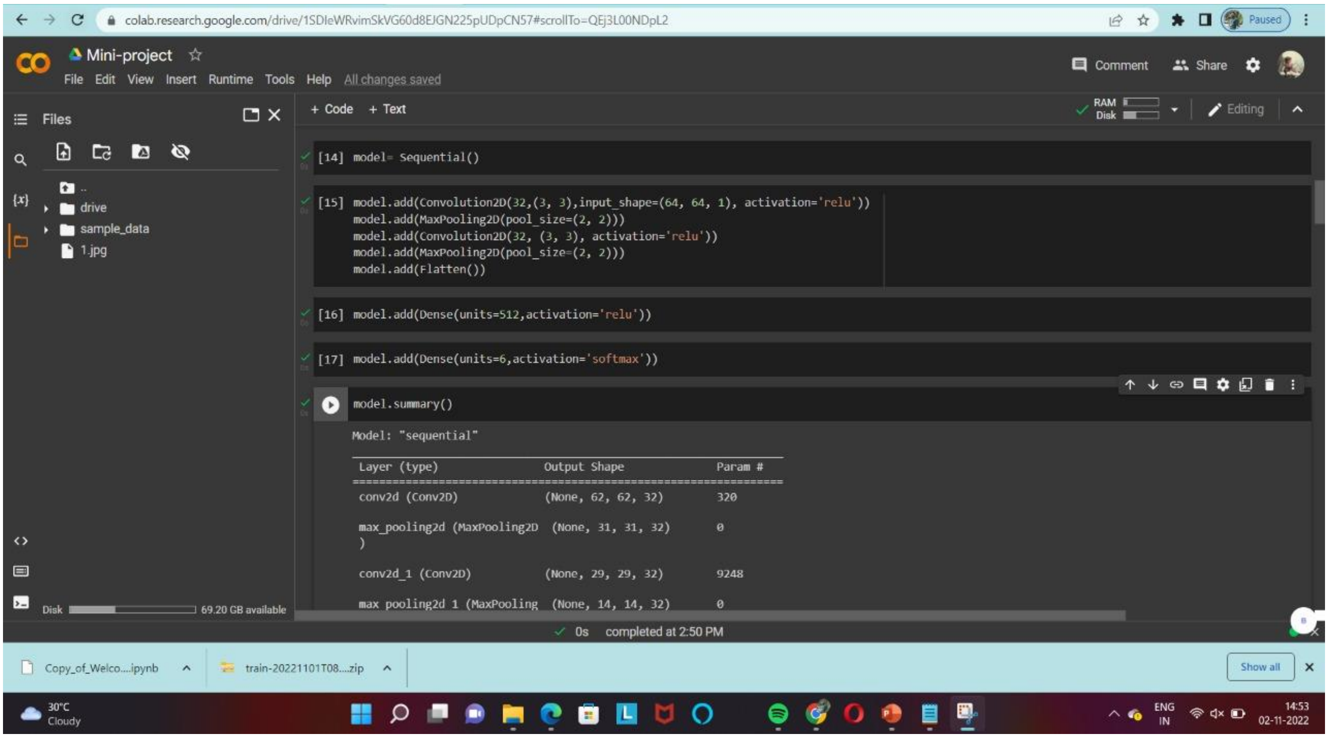
CHAPTER 7

CODING & SOLUTIONING (Explain the features added in the project along with code)

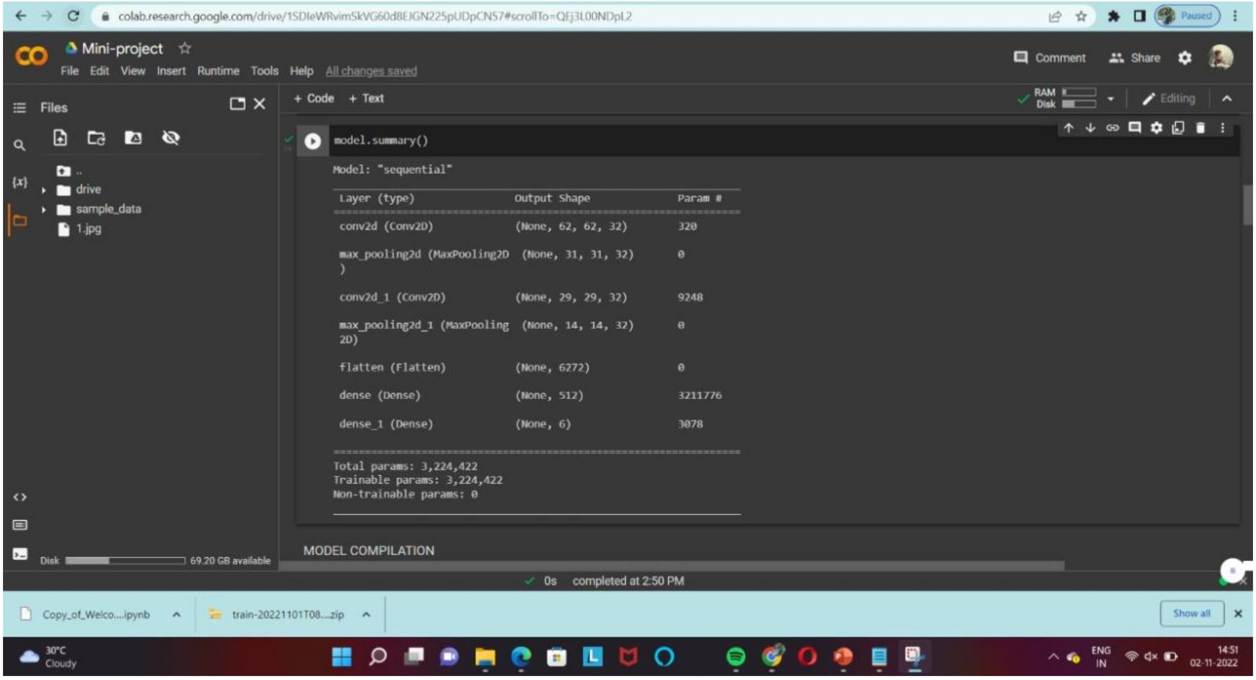
7.1. FEATURE 1

In this step we build Convolutional Neural Networking which contains a input layer along with the convolution, maxpooling and finally a output layer.

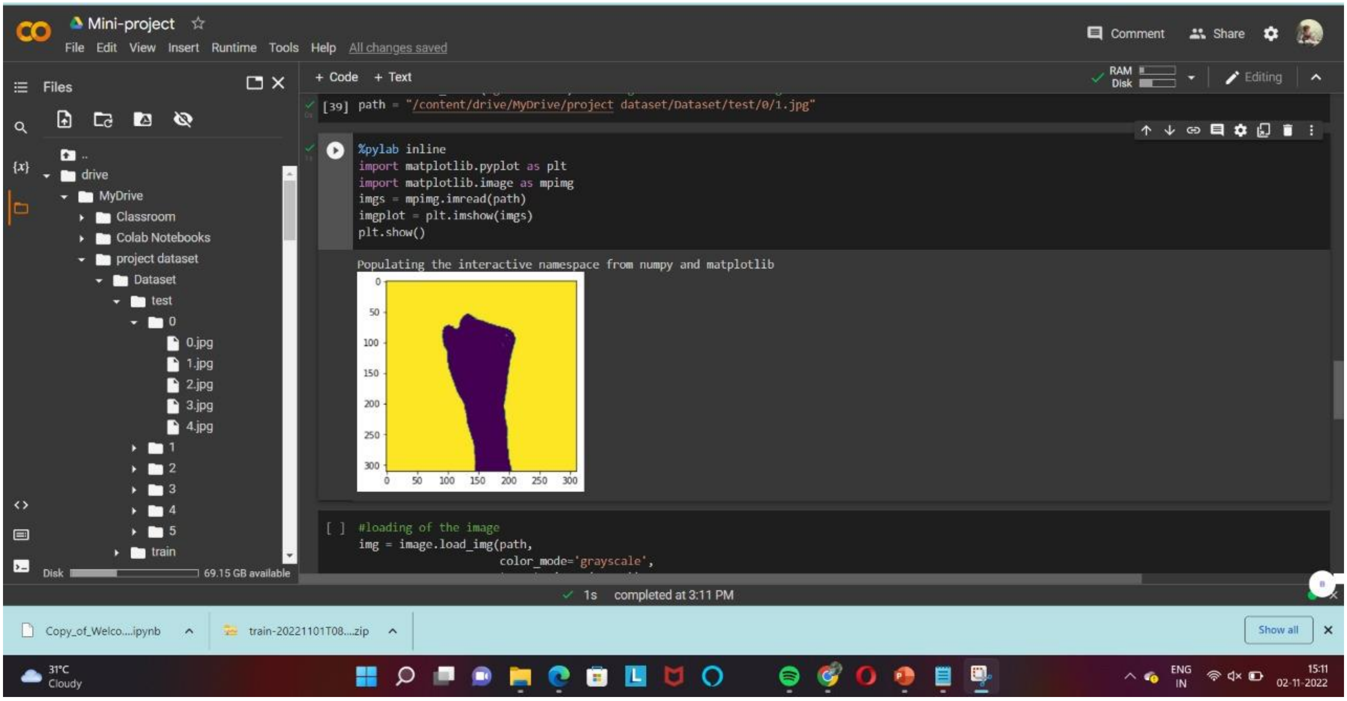



colab.research.google.com/drive/1SDIeWRvimSkVG60d8EJGN225pUDpCN57#scrollTo=QEj3L00NDpL2
Mini-project
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
Files
+ Code + Text
RAM Disk
Editing
drive
sample_data
1.jpg
[14] model= Sequential()
[15] model.add(Convolution2D(32,(3, 3),input_shape=(64, 64, 1), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Convolution2D(32, (3, 3), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Flatten())
[16] model.add(Dense(units=512,activation='relu'))
[17] model.add(Dense(units=6,activation='softmax'))
model.summary()
Model: "sequential"
Layer (type) Output Shape Param #
conv2d (Conv2D) (None, 62, 62, 32) 320
max_pooling2d (MaxPooling2D (None, 31, 31, 32) 0
)
conv2d_1 (Conv2D) (None, 29, 29, 32) 9248
max_pooling2d_1 (MaxPooling (None, 14, 14, 32) 0
Disk 69.20 GB available
0s completed at 2:50 PM
Copy_of_Welco....ipynb
train-20221101T08....zip
Show all
30°C Cloudy
ENG IN
14:53 02-11-2022

7.2. FEATURE 2



Mini-project
model.summary()
Model: "sequential"
Layer (type) Output Shape Param #
conv2d (Conv2D) (None, 62, 62, 32) 320
max_pooling2d (MaxPooling2D) (None, 31, 31, 32) 0
conv2d_1 (Conv2D) (None, 29, 29, 32) 9248
max_pooling2d_1 (MaxPooling2D) (None, 14, 14, 32) 0
flatten (Flatten) (None, 6272) 0
dense (Dense) (None, 512) 3211776
dense_1 (Dense) (None, 6) 3078
Total params: 3,224,422
Trainable params: 3,224,422
Non-trainable params: 0
MODEL COMPILATION
completed at 2:50 PM



Mini-project
path = "/content/drive/MyDrive/project dataset/Dataset/test/0/1.jpg"
%pylab inline
import matplotlib.pyplot as plt
import matplotlib.image as mpimg
imgs = mpimg.imread(path)
imgplot = plt.imshow(imgs)
plt.show()
Populating the interactive namespace from numpy and matplotlib
#loading of the image
img = image.load_img(path,
color_mode='grayscale',
completed at 3:11 PM

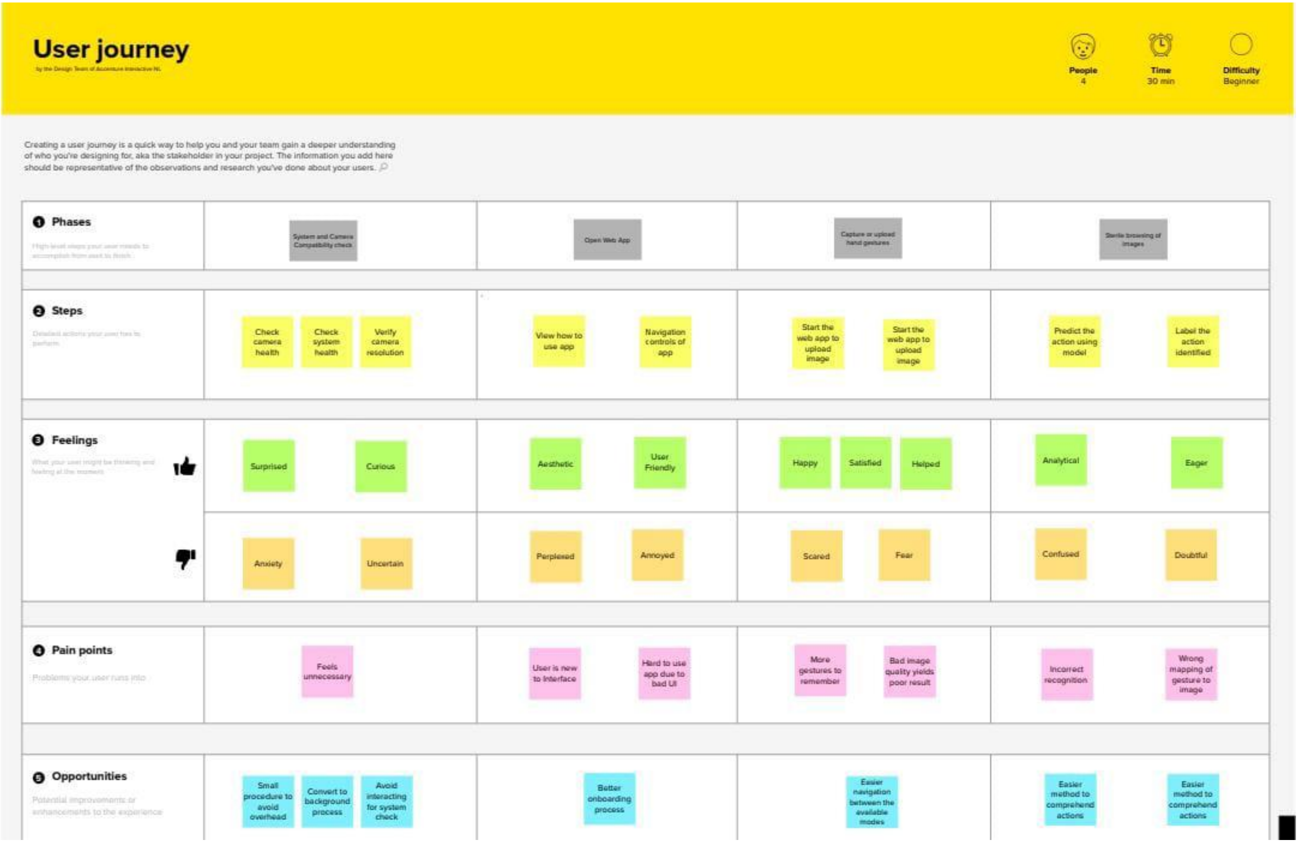
CHAPTER 8

TESTING

8.1 TEST CASE

5		Steps To Execute	Test Data	Expected Result	Actual Result	Status	C
11		1.Enter URL 2.Click on upload image utility 3. Choose the image thet needs to be manipulated	http://127.0.0.1:5000/launch Upload the image to be manipulated	The name of the image uploaded should be displayed	Working as expected	Pass	
12		1.Enter URL 2.Click on upload image utility 3. Choose the image thet needs to be manipulated 4. Click Submit	http://127.0.0.1:5000/launch Upload the image to be manipulated Click Enter	The web-camera starts	Working as expected	Pass	
13		1.Enter URL 2.Click on upload image utility 3. Choose the image thet needs to be manipulated 4. Click Submit	http://127.0.0.1:5000/launch Upload the image to be manipulated Click Enter	The image is stored in a folder in the backend	Working as expected	Pass	
14		1.Enter URL 2.Click on upload image utility 3. Choose the image thet needs to be manipulated 4. Click Submit	http://127.0.0.1:5000/launch Upload the image to be manipulated Click Enter	The webcam starts	Working as expected	Pass	
15		1.Enter URL 2.Click on upload image utility 3. Choose the image thet needs to be manipulated 4. Click Submit 5. Show gesture in the webcam	http://127.0.0.1:5000/launch Upload the image to be manipulated Click Enter Show Gesture 0 in the webcam	The gesture shown in camera should be captured ad the region of interest should be found	Working as expected	Pass	
16		1.Enter URL 2.Click on upload image utility 3. Choose the image thet needs to be manipulated 4. Click Submit 5. Show gesture in the webcam	http://127.0.0.1:5000/launch Upload the image to be manipulated Click Enter Show Gesture 0 in the webcam	The uploaded image should be resized and should be shown	Working as expected	Pass	
17		1.Enter URL 2.Click on upload image utility 3. Choose the image thet needs to be manipulated 4. Click Submit 5. Show gesture in the webcam	http://127.0.0.1:5000/launch Upload the image to be manipulated Click Enter Show Gesture 1 in the webcam	The uploaded image should be resized as rectangle and should be shown	Working as expected	Pass	
18		1.Enter URL 2.Click on upload image utility 3. Choose the image thet needs to be manipulated 4. Click Submit 5. Show gesture in the webcam	http://127.0.0.1:5000/launch Upload the image to be manipulated Click Enter Show Gesture 2 in the webcam	The uploaded image should be rotated 45 degrees and should be shown	Working as expected	Pass	
19		1.Enter URL 2.Click on upload image utility 3. Choose the image thet needs to be manipulated 4. Click Submit 5. Show gesture in the webcam	http://127.0.0.1:5000/launch Upload the image to be manipulated Click Enter Show Gesture 3 in the webcam	The uploaded image should be blurred and should be shown	Working as expected	Pass	
20		1.Enter URL 2.Click on upload image utility 3. Choose the image thet needs to be manipulated 4. Click Submit 5. Show gesture in the webcam	http://127.0.0.1:5000/launch Upload the image to be manipulated Click Enter Show Gesture 4 in the webcam	The uploaded image should be zoomed and should be shown	Working as expected	Pass	
21		1.Enter URL 2.Click on upload image utility 3. Choose the image thet needs to be manipulated 4. Click Submit 5. Show gesture in the webcam	http://127.0.0.1:5000/launch Upload the image to be manipulated Click Enter Show Gesture 5 in the webcam	The uploaded image should be negated and should be shown	Working as expected	Pass	

8.2. USER ACCEPTANCE TESTING



CHAPTER 9

RESULTS

Final findings (Output) of the project along with screenshots.

Through this project we found that we can maintain the sterility of an operation theater, etc by using hand based gesture tools to browse the images obtained.

CHAPTER 10

ADVANTAGES & DISADVANTAGES

Advantages:

- Major advantage of this tool is that it helps to maintain the sterility of the environment. It is also easy to use and is quicker than the existing methods to browse images.

- It can also be performed even if the surgeon is a bit far away from the system, this helps to save time.

- The tool does not need the person using it to have an apparatus or any devices on them to use it.

- They can simply move their hands to browse through the images.

Disadvantages:

- The tool can be quite expensive as it requires cameras and other expensive devices to capture images and process it.

CHAPTER 11

CONCLUSION

- In this project we developed a tool which recognises hand gestures and enables doctors to browse through radiology images using these gestures. This enables doctors and surgeons to maintain the sterility as they would not have to touch any mouse or keyboard to go through the images.

- This tool is also easy to use and is quicker than the regular method of using mouse/keyboard. It can be used regardless of the users location since they don't have to be in contact with any device.

- It also does not require the user to have any device on them to use it.

- Further this technology can be extended to other industries like it can be used by presenters, by teachers for show images in the classroom, etc.

CHAPTER 12

FUTURE SCOPE

The tool can be made quicker by increasing the recognition speed.More number of gestures can be added thereby increasing this tool's functionality and useability for different purposes.Tracking of both hands can be added to increase the set of commands. Voice commands can also be added to further increase the functionality.

CHAPTER 13

APPENDIX

Research papers:

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2410001/

https://pubmed.ncbi.nlm.nih.gov/18451034/

https://www.researchgate.net/publication/5401674 A Gesture- based Tool for Sterile Browsing of Radiology Images

Smartinternz Website:

https://careereducation.smartinternz.com/Student/guided_project_info/9890#

Appendix source code

https://github.com/smartinternz02/Gesture-based-Tool-for-Sterile-Browsing-of-Radiology-Images-Using-IBM-Watson