PROJECT REPORT

GESTURE BASED TOOL FOR STERILE BROWSING OF RADIOLOGY IMAGES

Team ID: PNT2022TMID53623

TEAM MEMBERS:

PRAPTI D[TL] – 2127190701078

PRABUMANOJ S J[TM] – 2127190701075

RAGHU YOGESH [TM] – 2127190701083

OVIYA G[TM] – 2127190701071

1. INTRODUCTION

1.1 Project Overview

In this project we use hand gestures to browse images obtained from x-rays.
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, mouse or touchscreen are the most popular input methods today, but these depend on physical touch to work. This leads to higher possibilities of touch surfaces being contaminated. This is a very risky scenario in hospital where a sterile environment needs to be maintained for patient safety.

In order overcome this point of failure, we are going to use gesture identification to handle the xray images. To achieve this goal certain problems need to be solved: how to separate objects of interest in images and which image capture technology and classification technique are more appropriate, among others.
In this Gesture based operation project, first the model is trained on the images of different hand gestures, such as showing numbers with fingers as 1,2,3,4,5. This model uses the integrated webcam in the laptop to capture the hand images.

The image of the gesture captured in the video frame is compared with the Pre-trained model and the gesture is identified. If the gesture predicts is,

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 etc.

1.2 Purpose

It is used to browse and manipulate the images obtained using X-ray using hand gestures rather than using mouse,keyboard,etc thereby maintaining sterility.

2. LITERATURE SURVEY

2.1 Existing Problem

- Limited number of options.
- Recognising the gestures
- Personalization
- Application style
- Fine control over image manipulation

2.2 References:

https://www.researchgate.net/publication/5401674_A_Gesture-based_Tool_for_Sterile_Browsing_of_Radiology_Images

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 paper presents "Gestix," 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.

https://www.mdpi.com/2072-4292/13/3/527/htm

Human–Computer Interfaces (HCI) deals with the study of interfaces between humans and computers. The use of radar and other RF sensors to develop HCI based on Hand Gesture Recognition (HGR) has gained increasing attention over the past decade. Today, devices have built-in radars for recognizing and categorizing hand movements. In this article, we present the first ever review related to HGR using radar sensors. We review the available techniques for multi-domain hand gestures data representation for different signal processing and deep-learning-based HGR algorithms. We classify the radars used for HGR as pulsed and continuous-wave radars, and both the hardware and the algorithmic details of each category are presented in detail.

https://www.degruyter.com/document/doi/10.1515/cdbme-2021-2068/pdf

This work examines how a touchless interaction concept contributes to an efficient, direct, and sterile interaction workflow during CT-guided interventions. Two hand gesture sets were designed specifically under consideration of the clinical workflow and the hardware capabilities. These were used to change the position of an X-Ray tube and detector of a CT scanner without breaking sterility and are compared regarding usability and performance in a user study with 10 users. The user study revealed that it is possible to change the angle of the gantry within 10 seconds on average in an experimental setup.

https://journals.sagepub.com/doi/pdf/10.1177/1460458217748342

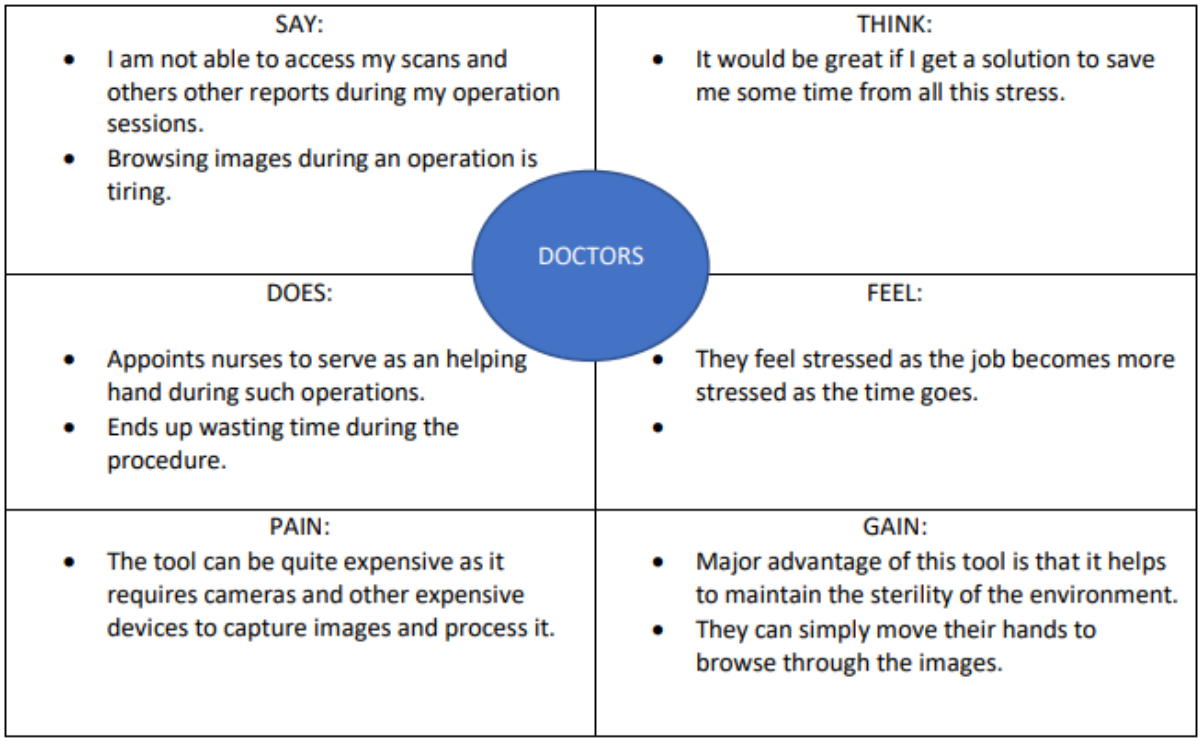
The widespread use of technology in hospitals and the difficulty of sterilising computer controls has increased opportunities for the spread of pathogens. This leads to an interest in touchless user interfaces for computer systems. We present a review of touchless interaction with computer equipment in the hospital environment, based on a systematic search of the literature. Sterility provides an implied theme and motivation for the field as a whole, but other advantages, such as hands-busy settings, are also proposed.Gestures have been implemented for input, system and content control.

2.3 Problem Statement Definition

The aim of this project is to provide a touch-free solution for browsing radiology images in the hospital to avoid contamination.

3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas



3.2 Ideation & Brainstorming

Using image processing to recognise hand gestures and execute the corresponding actions, thus providing a sterile work environment.

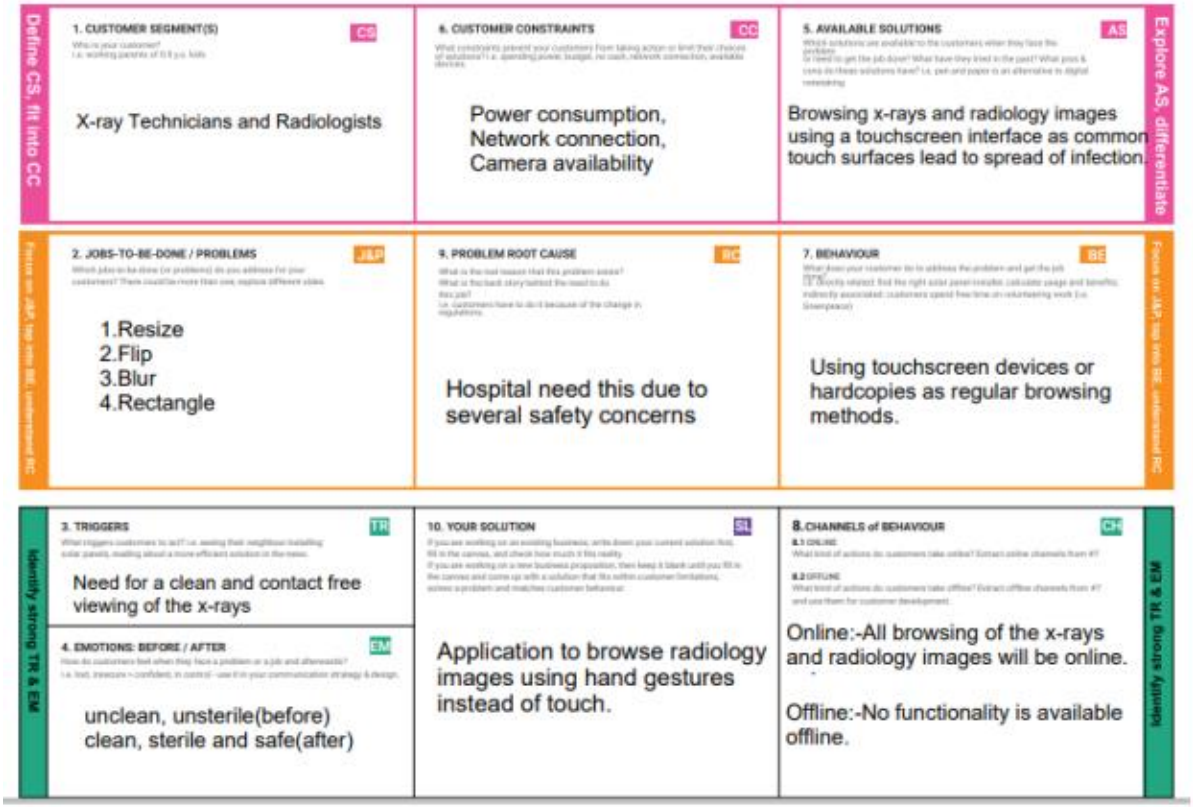
This model uses the integrated webcam to capture the video frame. The image of the gesture captured in the video frame is compared with the pre-trained model and the gesture is identified.

Humans will find it more convenient to navigate using gestures than touching unclean surfaces in a hospital.

3.3 Proposed Solution

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	A solution for sterile browsing of radiology images by using a contact free method
2.	Idea / Solution description	An application that can detect hand gestures using a camera and execute the appropriate action for browsing images by using a machine learning model, thereby providing a contact free experience.
3.	Novelty / Uniqueness	There are similar solutions in the internet.
4.	Social Impact / Customer Satisfaction	This solution can help society by decreasing the amount of touch surfaces which can spread diseases in a hospital environment. It would also provide a comfortable experience for the customer by being user-friendly to operate.
5.	Business Model (Revenue Model)	We can sell this application as a whole or setup a server which customers can use and setup a subscription system for the API keys for using the server providing 2 opportunities for revenue at different levels.
6.	Scalability of the Solution	In the case of a server we can upgrade the server to handle more requests at a time which ensures scalability from the server side. In the case of an application , we can configure the application to handle inputs from multiple cameras at a time.

3.4 Problem Solution fit



4. REQUIREMENT ANALYSIS

4.1 Functional requirement

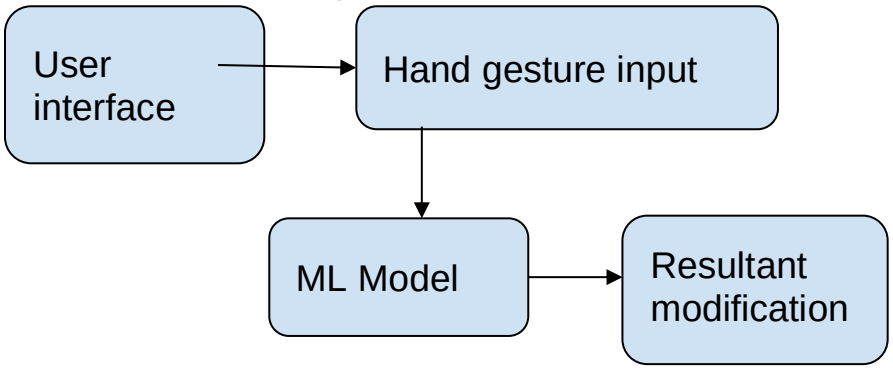
FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Experience	A responsive UI/UX shall be designed to help users interact with the system by using a cohort of hand gestures
FR-2	Cloud Deployment	The system shall deploy the trained CNN on the cloud.
FR-3	Hand Gesture Identification	The system shall be able to classify the images of hand gestures captured by a camera.
FR-4	Application Domain Pertinence	The CNN used by the system shall be trained on data that is relevant to the application domain.

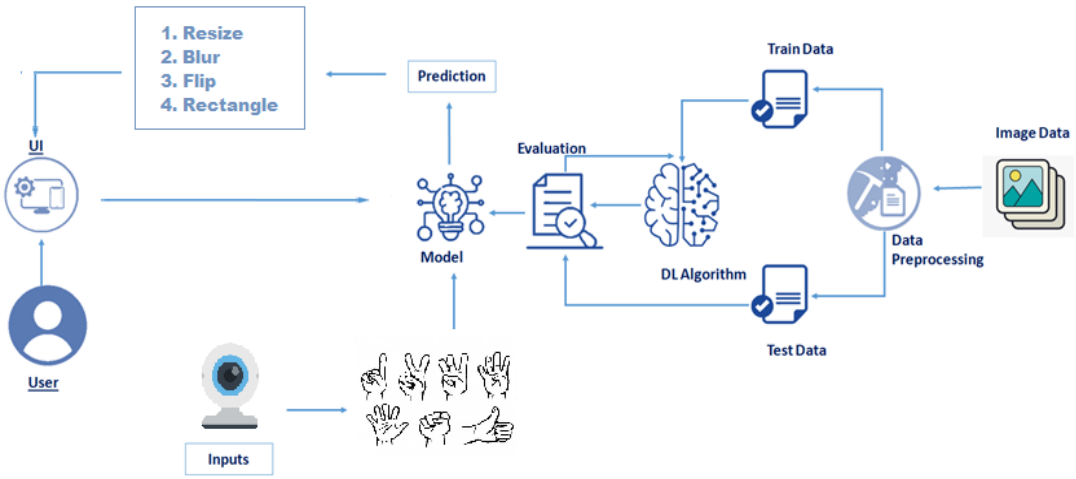
4.2 Non-Functional requirements

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	The system shall act as an arbitrator between the user and the deep learning model.
NFR-2	**Security**	The system shall only permit authorized users to access the system. The system shall prevent unauthorized users from entering the system.
NFR-3	**Reliability**	The system shall be in the operational mode for at least 361 days in a year (Uptime of 99%). In case of failure, the system shall be able to recover in under 3-5 seconds.
NFR-4	**Performance**	The system shall be able to respond to a user gesture in under 3-5 seconds.
NFR-5	**Availability**	The model shall be available for public use as long as it remains operational.
NFR-6	**Scalability**	The system shall be accessible to over 5,00,000 - 1,000,000 concurrent users without any loss of performance.

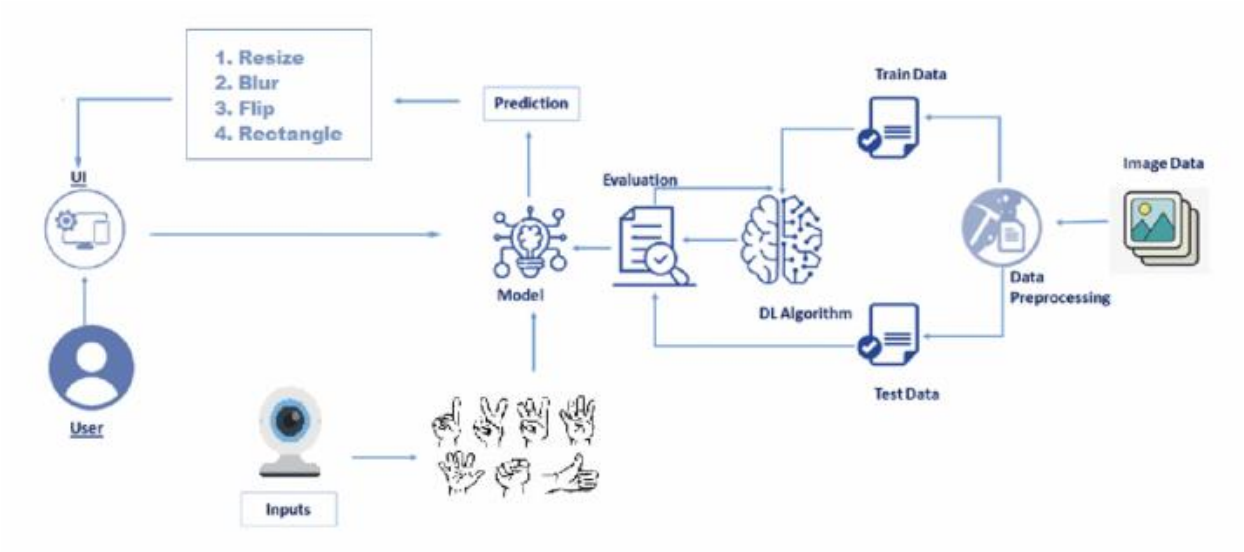
5. PROJECT DESIGN

5.1 Data Flow Diagrams





5.2 Solution & Technical Architecture



5.3 User Stories

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer	Image input	USN-1	As a user, I can upload images of x-rays	Image is successfully uploaded	High	Sprint-1
	Hand gesture input	USN-2	As a user, I can input a hand gesture using the camera	Hand gesture is recognised	High	Sprint-1
	Image Result	USN-3	As a user, I can view the updated image based on hand gesture	Correctly updated image is shown	High	Sprint-2

6. PROJECT PLANNING & SCHEDULING

6.1 Sprint Planning & Estimation

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Launch application	USN-1	As a user, I can download and launch the application.	1	Medium	**Raghu yogesh Prabumanoj S J Prapti D Oviya G**
Sprint-2	Camera interface	USN-2	As a user, I i need the camera interface for gesture input	1	Medium	**Raghu yogesh Prabumanoj S J Prapti D Oviya G**

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-2	AI model training	USN-3	As a user, I need the AI model ready for recognising the gestures	2	High	**Raghu yogesh Prabumanoj S J Prapti D Oviya G**
Sprint-3	Deployment to cloud	USN-4	As a user, I need the app to be accessible anywhere ,	1	Medium	**Raghu yogesh Prabumanoj S J Prapti D Oviya G**
Sprint-3	Prediction	USN-5	As a user, I need the predictions from the model	2	High	**Raghu yogesh Prabumanoj S J Prapti D Oviya G**
Sprint-4	Image manipulation	USN-6	As a user, I need the image to be altered based on the prediction of the model.	2	High	**Raghu yogesh Prabumanoj S J Prapti D Oviya G**

6.2 Sprint Delivery Schedule

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	6 Days	24 Oct 2022	29 Oct 2022		29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022		
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022		
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022		

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

7.1 Gesture recognition using model

The hand gesture of the user is captured using the webcam of the laptop. The image is then passed through the gesture identification model which returns the probability of the gesture being one of the 5 gestures. The gesture with the highest probability is returned as the prediction.

The process can be repeated on demand using the ESC key and stopped by holding the key for 10 secs.

```
basepath = os.path.dirname(__file__)
file_path = os.path.join(basepath, 'uploads', secure_filename(f.filename))
f.save(file_path)
print(file_path)
cap = cv2.VideoCapture(0)
while True:
    _, frame = cap.read() #capturing the video frame values
    # Simulating mirror image
    frame = cv2.flip(frame, 1)

    # Got this from collect-data.py
    # Coordinates of the ROI
    x1 = int(0.5*frame.shape[1])
    y1 = 10
    x2 = frame.shape[1]-10
    y2 = int(0.5*frame.shape[1])
    # Drawing the ROI
    # The increment/decrement by 1 is to compensate for the bounding box
    cv2.rectangle(frame, (x1-1, y1-1), (x2+1, y2+1), (255,0,0) ,1)
    # Extracting the ROI
    roi = frame[y1:y2, x1:x2]

    # Resizing the ROI so it can be fed to the model for prediction
    roi = cv2.resize(roi, (64, 64))
    roi = cv2.cvtColor(roi, cv2.COLOR_BGR2GRAY)
    _, test_image = cv2.threshold(roi, 120, 255, cv2.THRESH_BINARY)
    cv2.imshow("test", test_image)
    # Batch of 1
    result = model.predict(test_image.reshape(1, 64, 64, 1))
    prediction = {'ZERO': result[0][0],
                  'ONE': result[0][1],
                  'TWO': result[0][2],
                  'THREE': result[0][3],
                  'FOUR': result[0][4],
                  'FIVE': result[0][5]}
    # Sorting based on top prediction
    prediction = sorted(prediction.items(), key=operator.itemgetter(1), reverse=True)
```

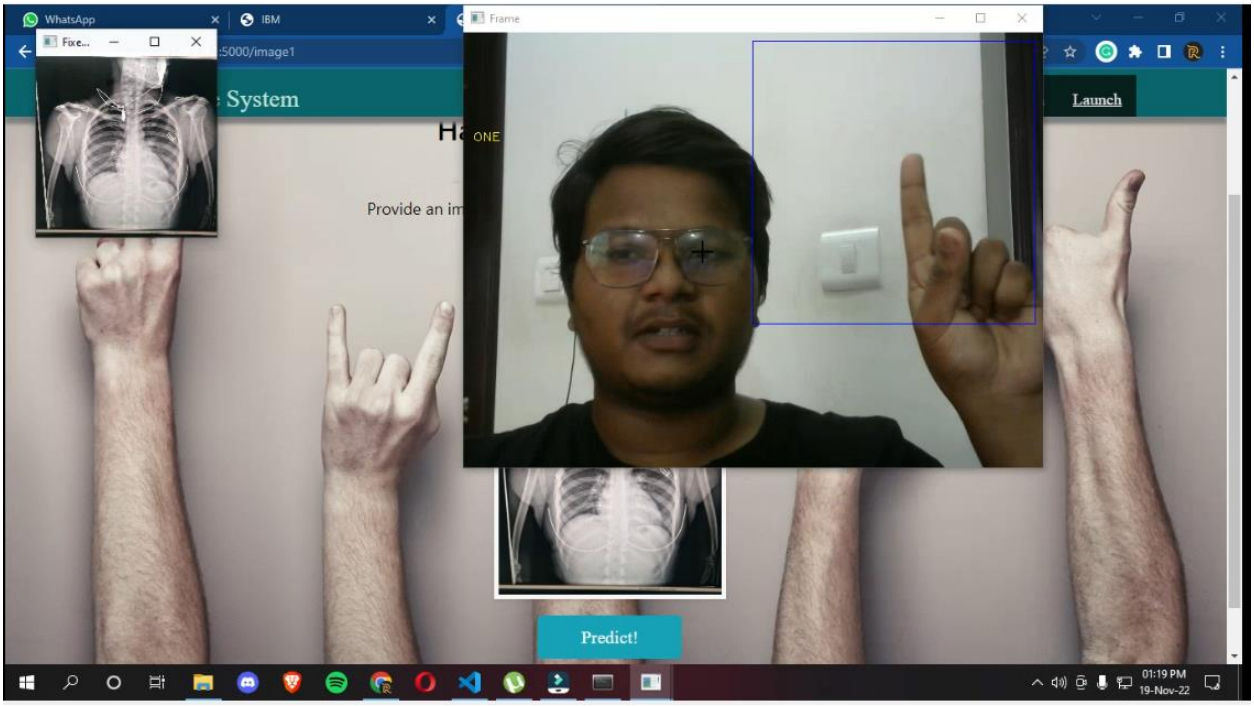
7.2 Rotating image

Each of the gestures correspond to one manipulation which can be done to the image. In this case we have taken the rotation operation. The image is altered using the opencv2 package and the resultant image is shown in a new window.

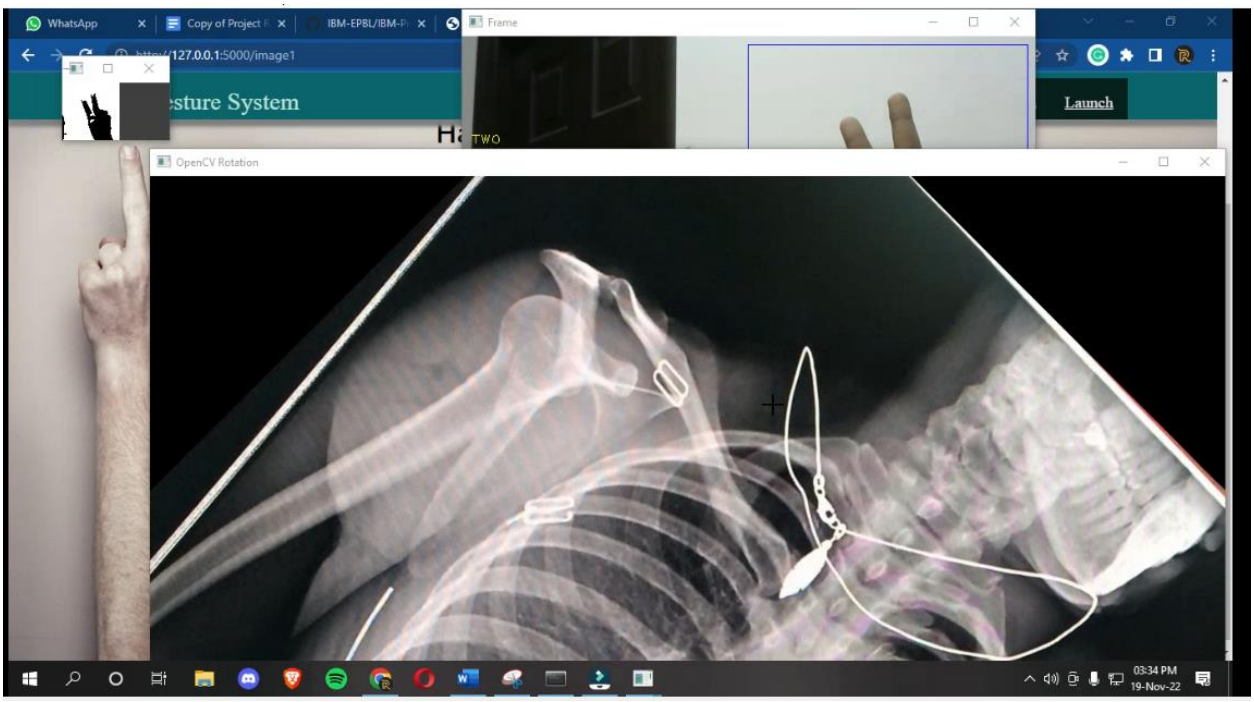
```
(h, w, d) = image1.shape
center = (w // 2, h // 2)
M = cv2.getRotationMatrix2D(center, -45, 1.0)
rotated = cv2.warpAffine(image1, M, (w, h))
cv2.imshow("OpenCV Rotation", rotated)
key=cv2.waitKey(3000)
if (key & 0xFF) == ord("2"):
    cv2.destroyWindow("OpenCV Rotation")
```

8. TESTING

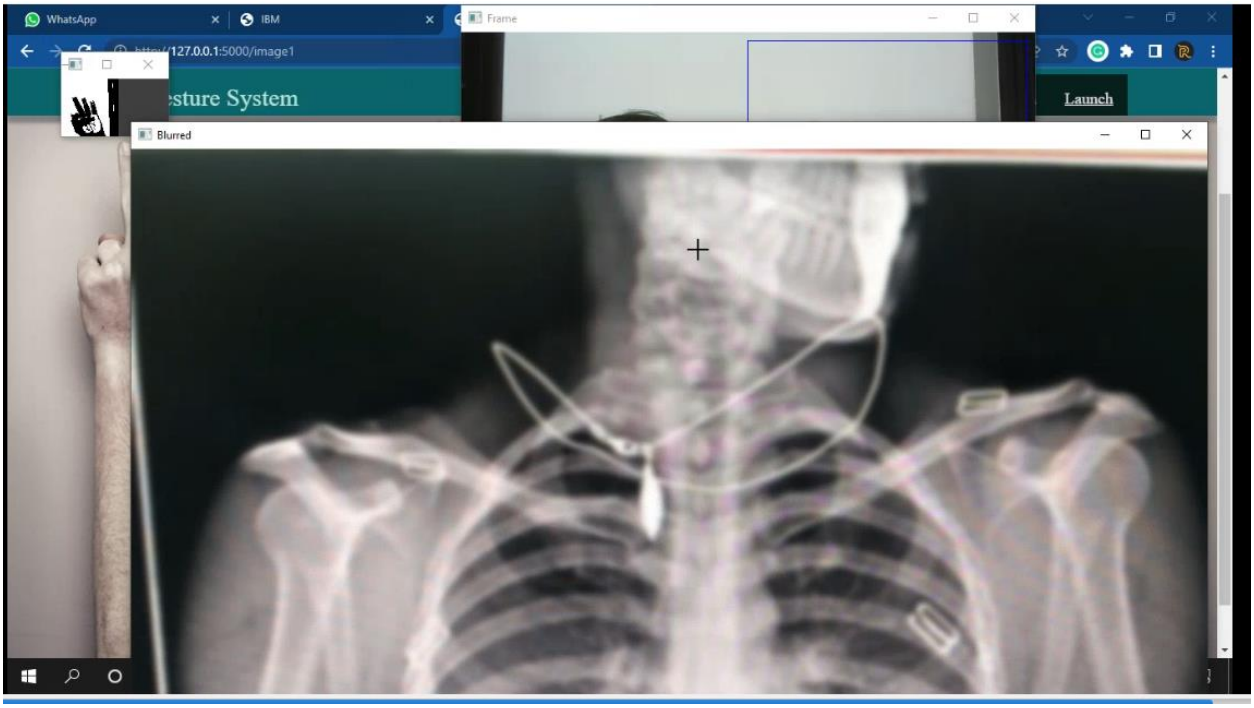
8.1 Test Cases



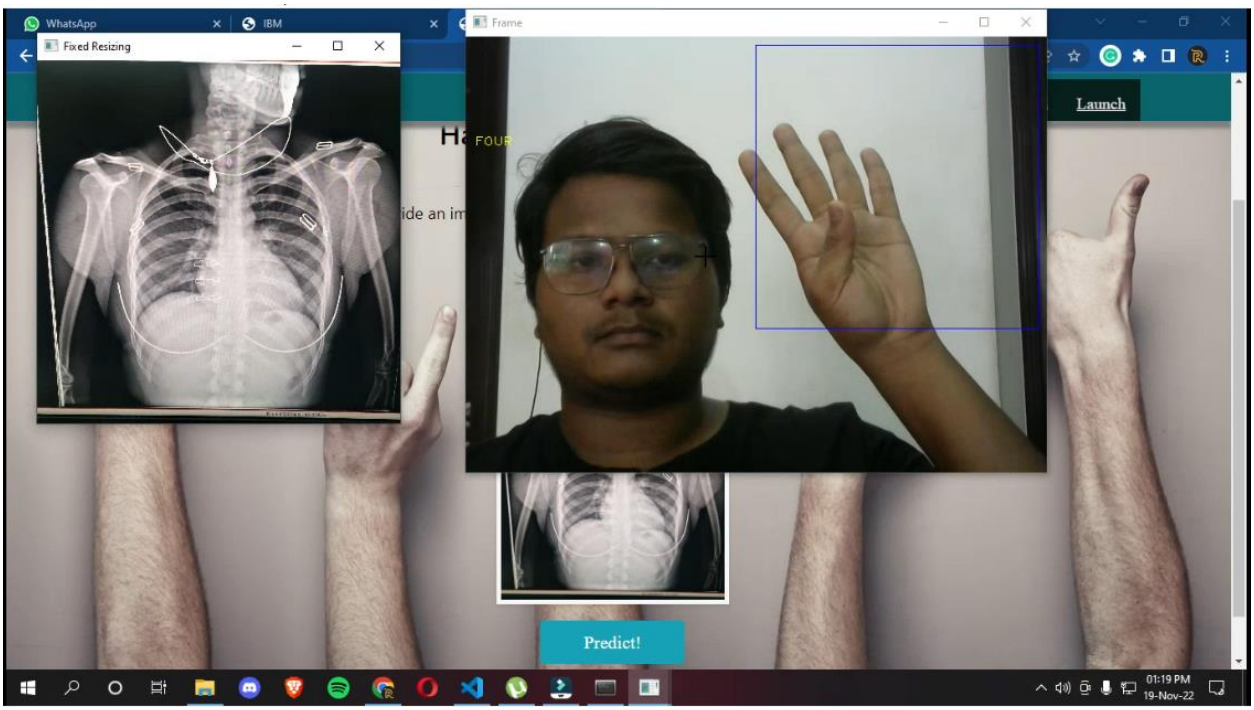
Gesture - 1 , resizing the image to make it smaller



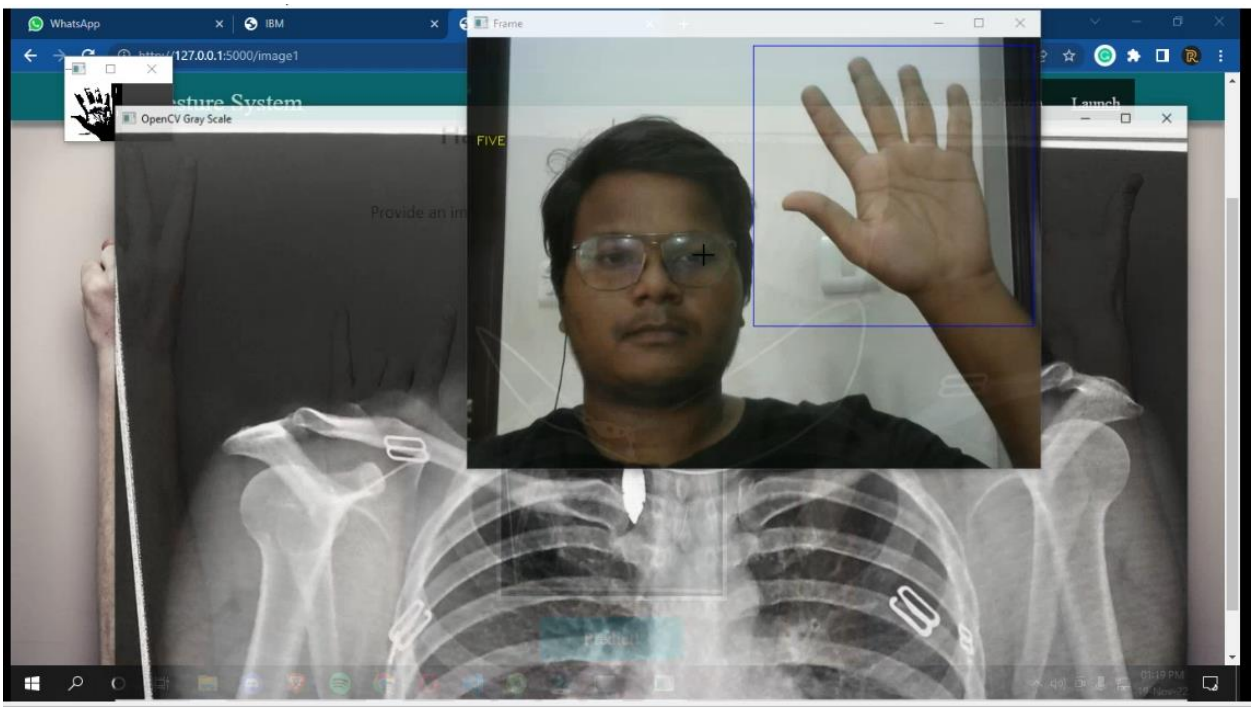
Gesture - 2 , rotating the image by 45 degrees



Gesture - 3 Making the image blurred



Gesture - 4 Resizing the image to make it bigger



Gesture - 5 Showing the image in grayscale (b/w)

9. RESULTS

9.1 Performance Metrics

loss: 0.0925

accuracy: 0.9781

val_loss: 0.2777

val_accuracy: 0.9000

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.

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 sterility by using a hands free interface. 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 user's 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 to show images in the classroom, etc.

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.

13. APPENDIX

Source Code

GitHub link : https://github.com/IBM-EPBL/IBM-Project-11403-1659326443

Demolink:https://drive.google.com/drive/folders/1mc1M7Qp63nUgYR0RS0TsxGTt_GNjycTZ?usp=sharing