A Gesture-based Tool for Sterile Browsing of Radiology Images

A PROJECT REPORT

Submitted by

TEAM ID: PNT2022TMID39878

LAKSHMI G V N	**511319104039 (Team Leader)**
ANTONY SOPHIA	**511319104005**
GODAVARI M	**511319104017**
HEENA SHAREEN N	**511319104023**

In partial fulfilment for the award of the degree

of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

KINGSTON ENGINEERING COLLEGE, VELLORE

ANNA UNIVERSITY: CHENNAI 600 025

Executed By:

IBM

CONTENT

ABSTRACT

The use of doctor-computer interaction devices in the operating 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. It is 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.

Humans are able to recognize body and sign language easily. This is possible due to the combination of vision and synaptic interactions that were formed along brain development. In order to replicate this skill in computers, some 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 project Gesture based Desktop automation , First the model is trained pre trained on the images of different hand gestures, such as a showing numbers with fingers as 1 ,2,3,4 . 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. If the gesture predicts is 1 then images is blurred;2, image is resized;3,image is rotated etc.

1. INTRODUCTION

Computer information technology is increasingly penetrating 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 methods 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.

1. In this paper, 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.

2. 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 objectNaturalness of expression, non-encumbered interaction, intuitiveness, and high sterility are all good reasons to replace the current interface technology with more natural interfaces.

This paper 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 of 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 were 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 how the surgeon controls medical information avoiding contamination of the patient, the OR, and the surgeon.

Numerous technologies can be used for the automation of a process like Artificial Intelligence (AI), Machine Learning (ML), RoboticProcess Automation (RPA), Business Process Automation (BPA), Industrial Robots, Virtual Assistance, etc. Allthesetechnologiesarecurrentlyusedfor automatedprocesses, eachonehavinga diverseandspecific reasonforusage. Butamongalloftheseapplicationsand technologies, the Robotic Process Automation technology is the most eminent one, or as one might say, it is the base for all other existingautomatedtechnologies. Forexample, ifweconsider Artificial Intelligence and RPA, AI is the ability of a machine to "think", whereas RPA is the ability of a machine to "do". AI will empower the machine to thinklike humans, whereas RPA will instruct the machine to do what humans do. The building of AI is on the conceptualization thatmachines could think betterthan humans, but it will only be possible if the machine is taught in a manner such that it exceeds human judgment. And even ifitwouldbeabletothink, whatnext? Wemightgetthe solutionsforourproblems, buttheimplementationofthese solutions is carried ahead by RPA. To implement RPA, the organization needs to first identify the processes that need to be automated. Without this, the implementation of RPA cannot be initiated.Unfortunately, no method can guide organizations to identify these processes, so this task needs to be done depending on the organizations' requirements. This paper aims to describe what the technology of RPA is, and how it can be beneficial to use. As Elon Musk once quoted "the robots will be able to do everything better than us", the technology of RPA is the best-suited one for the automation of processes.

The paper covers the total VI sections. Section II of the paper shows the Literature Surveyedfor RPA. Section III describesthe technologyof Robotic process automation, itsscalabilityfor use, along withits benefitsand drawbacks. Section IV inthe paperexplains the previous projects using RPA.Section V describes the proposed system, its architecture, and its explanation. Section VI definesthe Future Scopefor RPA, whereas Section VII of the paper presents the Conclusion of our project.

1.1 PROJECT OVERVIEW

The hand gesture control system "*Gestix*" 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.

1) surgeons kept their focus of attention between the patient and the surgical point of interest on the touch-screen navigation system.
2) a short distance between the surgeon and the patient was maintained during most of the surgery.
3) the surgeon had to move close to the main control wall to discuss and browse through the patient's MRI images.

1.2 Purpose

In this paper,we present animage-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 of command that is meant to represent it. For example, moving the hand left represents a “turn left” command.

2. LITERATURE SURVEY

2.1 Existing problem

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

2.2 References

1.Sahoo, J.P.; Prakash, A.J.; Pławiak, P.; Samantray, S. Real-Time Hand Gesture Recognition Using Fine-Tuned Convolutional Neural Network. Sensors 2022.

This study focuses on a literature review of hand gesture strategies and discusses their pros and limits in various situations. In addition, the performance of these methods is tabulated, with an emphasis on computer vision techniques that deal with similarity and difference points; hand segmentation techniques algorithms and limitations; number and types of gestures; dataset used detection range (distance); and camera type.

2.Gadepalli, T.R.; Srivastava, G.; Liyanage, M.; Aparejo, M.; Chowdhary, C.L.; Koppa, S.; Maddi Kunta, P.K.R. Hand gesture recognition based on harris hawks optimized convolution neural network. Compute. Electra. Eng. 2022.

Convolutional neural networks (CNN) are used to categorize images of hand gestures. A newly developed metaheuristic technique, Harris hawk's optimization (HHO) algorithm, is utilized to optimize CNN's hyperparameters.

Their extensive comparison research shows that the proposed HHO-CNN hybrid model outperforms current models by achieving 100 percent accuracy.

3.Amin, M.S.; Rizvi, S.T.H. Sign Gesture Classification and Recognition Using Machine Learning. Cybernet. Syst. 2022.

This article examines the flex, accelerometer, and gyroscope-based smart prototype developed to recognize sign language motions. These sensors are put on a glove to record and assemble alphabetic (i.e., 0–10, A–Z) and numeric (i.e., 0–10, and A–Z datasets). The primary purpose of the proposed model is to categorize sign gestures produced by deaf-mute deaf-mutantand identify the true meaning of movements performed.

4.Kong, F.; Deng, J.; Fan, Z. Gesture recognition system based on ultrasonic FMCW and Comvest model. Measurement 2022.

Based on the ultrasonic frequency modulated continuous wave (FMCW) and the Conv LSTM model, a system for gesture identification was suggested in this work. It uses a hardware configuration consisting of one transmitter and three spatially separated receivers.

5.Sabo, S.; Singha, J.; Laskar, R.H. Dynamic hand gesture recognition using a combination of two-level tracker and trajectory-guided features. Multimer. Syst. 2022.

In this research, the authors proposed a hand gesture detection system for a dataset of lowercase numbers and alphabets. The suggested method recognizes the hand using the information on skin color and mobility. Hand tracking is performed using a two-level tracking system and a modified Kanade–Lucas–Tomasi (KLT) tracking algorithm.

6.Oudah, M.; Al-Naji, A.; Chahal, J. Computer Vision for Elderly Care Based on Hand Gestures. Computers 2021.

Deaf-mute elderly folk use five distinct hand signals to seek a particular item, such as drink, food, toilet, assistance, and medication. Since older individuals cannot do anything independently, their requests were delivered to their smartphones.

Challenges

Microsoft Kinect v2 sensor's capability to extract hand movements in real-time keeps this study in a restricted area.

Gesture-based technology may assist the handicapped, as well as the general public, to maintain their safety and requirements. Due to the significant changeability of the properties of each motion about various persons, gesture detection from video streams is a complicated matter.

7.Mujahid, A.; Awan, M.J.; Yasin, A.; Mohammed, M.A.; Damaševi˘cius, R.; Maskeliunas, R.; Abdulkareem, K.H. Real-Time Hand ¯ Gesture Recognition Based on Deep Learning YOLOv3 Model. Appl. Sci. 2021.

This article offers a lightweight model based on the YOLO (You Look Only Once) v3 and the DarkNet-53 neural networks for gesture detection without further preprocessing, filtration of pictures, and image improvement. Even in a complicated context, the suggested model was very accurate, and even in low-resolution image mode motions were effectively identified. Rate of the high frame.

Challenges

The primary challenge of this application for the identification of gestures in real-time is the classification and recognition of gestures. Hand recognition is a method used by several algorithms and ideas of diverse approaches for understanding the movement of a hand, such as pictures and neural networks.

8.Bhuiyan, M.R.; Abdullah, D.; Hashim, D.; Farid, F.; Uddin, D.; Abdullah, N.; Samsudin, D. Crowd density estimation using deep learning for Hajj pilgrimage video analytics. F1000Research 2021.

The purpose of this study is to offer a method for Hajj applications that is based on a convolutional neural network model. They also created a technique for counting and then assessing crowd density. The model employs an architecture that recognizes each individual in the crowd, marks their head position with a bounding box, and counts them in their unique dataset (HAJJ-Crowd).

Challenges

There has been a growth in interest in the improvement of video analytics and visual monitoring to better the safety and security of pilgrims while in Makkah. It is mostly because Hajj is a one-of-a-kind event with hundreds of thousands of people crowded into a small area.

9.Bhuiyan, M.R.; Abdullah, J.; Hashim, N.; Al Farid, F.; Samsudin, M.A.; Abdullah, N.; Uddin, J. Hajj pilgrimage video analytics using CNN. Bull. Electr. Eng. Inform. 2021.

This study presents crowd density analysis using machine learning. The primary goal of this model is to find the best machine-learning method for crowd-density categorization with the greatest performance

Challenges

Crowd control is essential for ensuring crowd safety. Crowd monitoring is an efficient method of observing, controlling, and comprehending crowd behavior

10.Zamri, M.N.H.B.; Abdullah, J.; Bhuiyan, R.; Hashim, N.; Farid, F.A.; Uddin, J.; Husen, M.N.; Abdullah, N. A Comparison of ML and DL Approaches for Crowd Analysis on the Hajj Pilgrimage. In Proceedings of the International Visual Informatics Conference; Springer: Berlin/Heidelberg, Germany, 2021.

Data augmentation is a technique of expanding the data set by producing various picture shapes to increase model performance

It also helps to mitigate the over-fitting issue in the model during the training stage. The overcast issue arises when random noise or mistakes occur instead of when the underlying connection is there. Using an increase in data, additional images were produced for the model from each picture because some irrelevant patterns may occur throughout the model training process. Several methods were employed for data augmentation operations: rotational changes, vertical and horizontal rotations, and intensity disorder, including light disturbances.

A classical ANN involves a local minimal issue, which typically ends with a local optimization process rather than a globally optimal state. More overfitting issues often complicate general machine learning models. Intensive network structure optimization may address the issues of the local minima and override by DNNs [80,81]. Deep learning is a machine learning-based approach that educates computers to accomplish tasks similar to those performed by humans.

11.Bari, B.S.; Islam, M.N.; Rashid, M.; Hasan, M.J.; Razman, M.A.M.; Musa, R.M.; Ab Nasir, A.F.; Majeed, A.P.A. A real-time approach to diagnosing rice leaf disease using deep learning-based faster R-CNN framework. PeerComput. Sci. 2021.

1A classical ANN involves a local minimal issue, which typically ends with a local optimization process rather than a globally optimal state. More overfitting issues often complicate general machine learning models. Intensive network structure optimization may address the issues of the local minima and override by DNNs.

Deep learning is a machine learning-based approach that educates computers to accomplish tasks similar to those performed by humans. For example, deep learning is the underlying technology that enables driverless automobiles to detect traffic lights and people. It is also the underlying principle of audio and speech recognition in a variety of devices, such as mobile phones and tablets. Deep learning is gaining popularity because it is capable of performing previously impossible tasks.

12Alonso, D.G.; Teyseyre, A.; Soria, A.; Berdun, L. Hand gesture recognition in real-world scenarios using approximate string matching. Multimer. Tools Appl. 2020.

investigated dynamic hand movements using spatial-temporal algorithm convolutional networks. Three types of graph edges associated with the activity of hand joints were proposed in a skeleton-based model. A deep neural network was utilized to pick semantic characteristics to provide an accurate output.

13.Zhang, T.; Lin, H.; Ju, Z.; Yang, C. Hand Gesture recognition in the complex background based on convolutional pose machine and fuzzy Gaussian mixture models. Int. J. Fuzzy Syst. 2020

created a string-matching technique for understanding hand motions in real-time situations. The k-means technique was used to create an approximation

string matching to capture the features of hand joints. It was done to enhance the precision of various motions by specifying the number of clusters.

14.Köpüklü, O.; Gunduz, A.; Kose, N.; Rigoll, G. Online dynamic hand gesture recognition including efficiency analysis. IEEE Trans. Biom. Behav. Identity Sci. 2020

The gesture matching was performed by analyzing the fused gesture dataset, where the gesture frames were categorized. The SFM was then utilized to accelerate the analysis processing. To improve the efficiency of gesture analysis

15.Alnaim, N. Hand Gesture Recognition Using Deep Learning Neural Networks. Ph.D. Thesis, Brunel University, London, UK, 2020.

In this study, image processing techniques such as wavelets and empirical mode decomposition were suggested to extract picture functionalities to identify 2D or 3D manual motions. Classification of artificial neural networks (ANN), which was utilized for the training and classification of data in addition to CNN (CNN).

Challenges

Three-dimensional gesture disparities were measured utilizing the left and right 3D gesture videos

16.Min, Y.; Zhang, Y.; Chai, X.; Chen, X. An efficient pointlet for point clouds-based gesture recognition. In Proceedings of the IEEE/CVF

Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020

This work formulates the recognition of gestures as an irregular issue of sequence identification and aims to capture long-run spatial correlations in points of the cloud. To spread information from past to future while maintaining its spatial structure, a new and effective Pointlet is suggested.

Challenges

The underlying geometric structure and distance information for the object surfaces are accurately described in dot clouds as compared with RGB data, which offer additional indicators of gesture identification

17.Al-Hammadi, M.; Muhammad, G.; Abdul, W.; Alsulaiman, M.; Bencherif, M.A.; Alrayes, T.S.; Mathkour, H.; Mekhtiche, M.A. Deep learning-based approach for sign language gesture recognition with efficient hand gesture representation. IEEE Access 2020

A new system is presented for dynamic recognition of hand gestures utilizing various architectures to learn how to partition hands, local and global features, and globalization and recognition features of the sequence

Challenges

To create an efficient system for recognition, hand segmentation, local representation of hand forms, global corporate configuration, and gesture sequence modeling need to be addressed.

18.Neethu, P.; Suguna, R.; Sathish, D. An efficient method for human hand gesture detection and recognition using deep learning convolutional neural networks. Soft Comput. 2020

This article detects and recognizes the gestures of the human hand using the method of classification for neural networks (CNN). This process flow includes hand area segmentation using mask image, finger segmentation, segmented finger image normalization, and CNN classification finger identification

Challenges

SVM and the naive Bayes classification were used to recognize the conventional gesture technique and needed a large amount of data for the identification of gesture patterns.

19.Zoph, B.; Cubuk, E.D.; Ghiasi, G.; Lin, T.Y.; Shlens, J.; Le, Q.V. Learning data augmentation strategies for object detection. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020

It also helps to mitigate the over-fitting issue in the model during the training stage. The overcast issue arises when random noise or mistakes occur instead of when the underlying connection is there. Using an increase in data, additional images were produced for the model from each picture because some irrelevant patterns may occur throughout the model training process. Several methods were employed for data augmentation operations: rotational changes, vertical and horizontal rotations, and intensity disorder, including light disturbances

20.Mungra, D.; Agrawal, A.; Sharma, P.; Tanwar, S.; Obaidat, M.S. PRATT: A CNN-based emotion recognition system using histogram equalization and data augmentation. Multimed. Tools Appl. 2020

Convolutional neural networks (CNN) are used to categorize images of hand gestures. A newly developed metaheuristic technique, Harris hawk's

optimization (HHO) algorithm, is utilized to optimize CNN's hyperparameters. Their extensive comparison research shows that the proposed HHO-CNN hybrid model outperforms current models by achieving 100 percent accuracy

21.Rashid, M.; Sulaiman, N.; PP Abdul Majeed, A.; Musa, R.M.; Bari, B.S.; Khatun, S. Current status, challenges, and possible solutions of EEG-based brain-computer interface: A comprehensive review. Front. Neurorobotics 2020

A classical ANN involves a local minimal issue, which typically ends with a local optimization process rather than a globally optimal state. More overfitting issues often complicate general machine learning models. Intensive network structure optimization may address the issues of the local minima and override by DNNs

22.Mathew, A.; Amudha, P.; Sivakumar, S. Deep Learning Techniques: An Overview. In Proceedings of the International Conference on Advanced Machine Learning Technologies and Applications, Manipal, India, 13–15 February 2020;

Deep learning is a machine learning-based approach that educates computers to accomplish tasks similar to those performed by humans. For example, deep learning is the underlying technology that enables driverless automobiles to detect traffic lights and people. It is also the underlying principle of audio and speech recognition in a variety of devices, such as mobile phones and tablets. Deep learning is gaining popularity because it is capable of performing previously impossible tasks. A deep learning model is constructed by layering data, which may be images, text, or audio, into distinct and discrete categorization layers.

2.3 Problem Statement Definition

1. Visual-based mouse design receives wearable models instead of visual mice.

2. Basically for this project use hand gestures recorded by a webcam.

3. The camera captures and detects hand movements and performs mouse functions.

In this project Gesture based Desktop automation. First, the model is trained pre-trained on the images of different hand gestures, such as a showing number with fingers 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

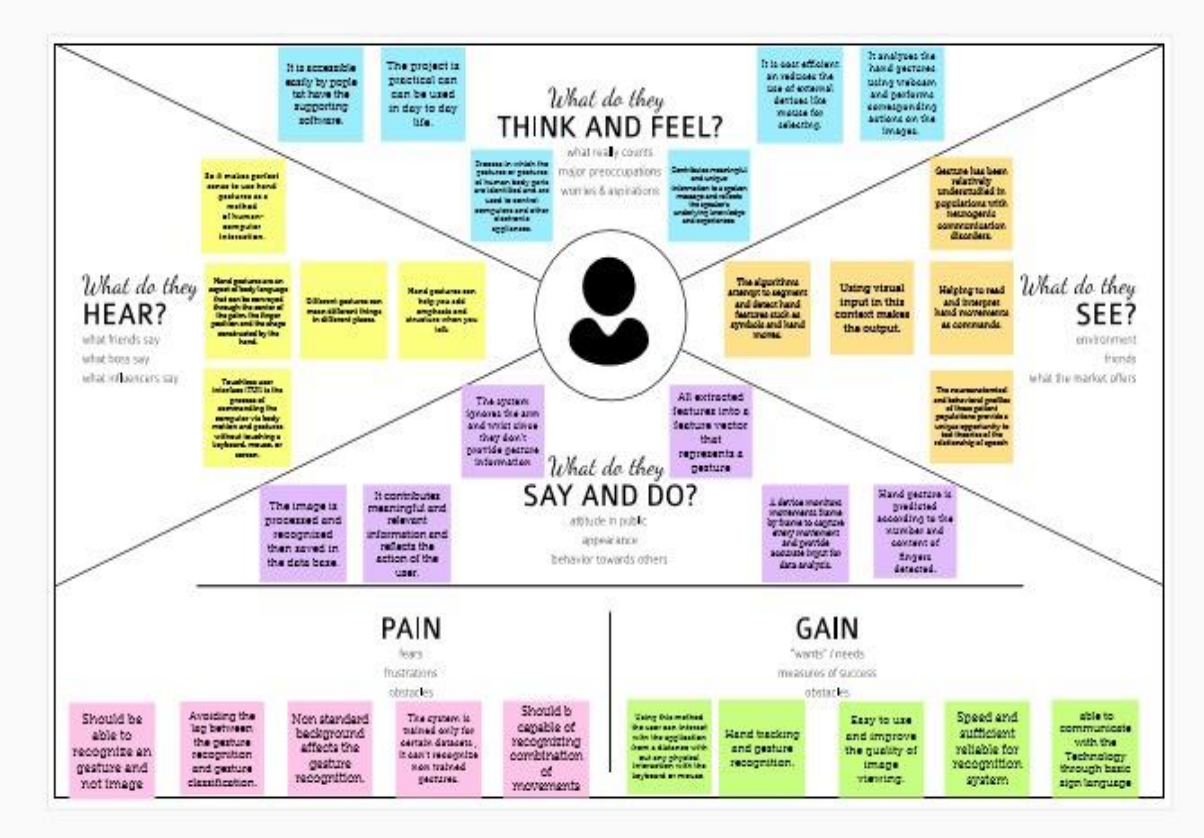
3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas

An empathy map canvas is a more in-depth version of the original empathy map, which helps identify and describe the user's needs and pain points. And this is valuable information for improving the user experience.

Teams rely on user insights to map out what is important to their target audience, what influences them, and how they present themselves. This information is then used to create personas that help teams visualize users and empathize with them as individuals, rather than just as a vague marketing demographic or account number.

An empathy map canvas helps brands provide a better experience for users by helping teams understand the perspectives and mindsets of their customers. Using a template to create an empathy map canvas reduces the preparation time and standardizes the process so you create empathy map canvases of similar quality.



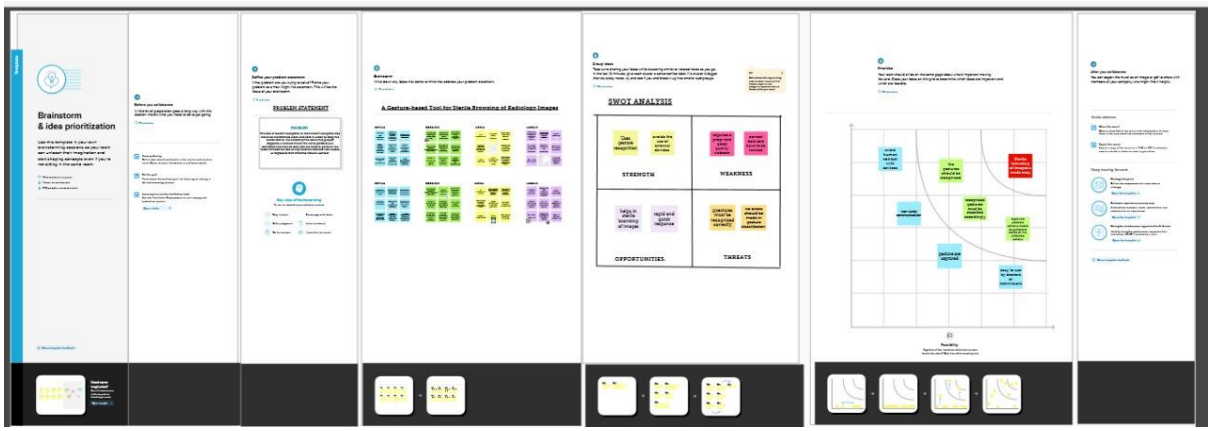
3.2 Ideation & Brainstorming

Ideation is often closely related to the practice of brainstorming, a specific technique that is utilized to generate new ideas. A principal difference between ideation and brainstorming is that ideation is commonly more thought of as being an individual pursuit, while brainstorming is almost always a group activity. Brainstorming is usually conducted by getting a group of people together to come up with either general new ideas or ideas for solving a specific problem or dealing with a specific situation.

Participants in a brainstorming session are encouraged to freely toss out whatever ideas may occur to them. The thinking is that by generating a large number of ideas, the brainstorming group is likely to come up with a suitable solution for whatever issue they are addressing.

The lines between ideation and brainstorming have become a bit more blurred with the development of several brainstorming software programs, such as Brightidea and Ideawake. These software programs are designed to encourage employees of companies to generate new ideas for improving the companies' operations and, ultimately, bottom-line profitability.

The programs often combine the processes of ideation and brainstorming in that individual employees can use them, but companies may simulate brainstorming sessions by having several employees all utilize the software to generate new ideas intended to address a specific purpose.



3.3 Proposed Solution

RPA canbeusedinan education organization in the account section, examination section, and admission depart it and thus reducing thenumber of humans required to carry out the work. RPA can be usedto handledata inanyorganization. RPA canbe usedtoautomateotherpaperworkformalitiesthus saving paper.

RPA in education can help educators and administrators automate tasks and processes which will otherwise become cumbersome for them.

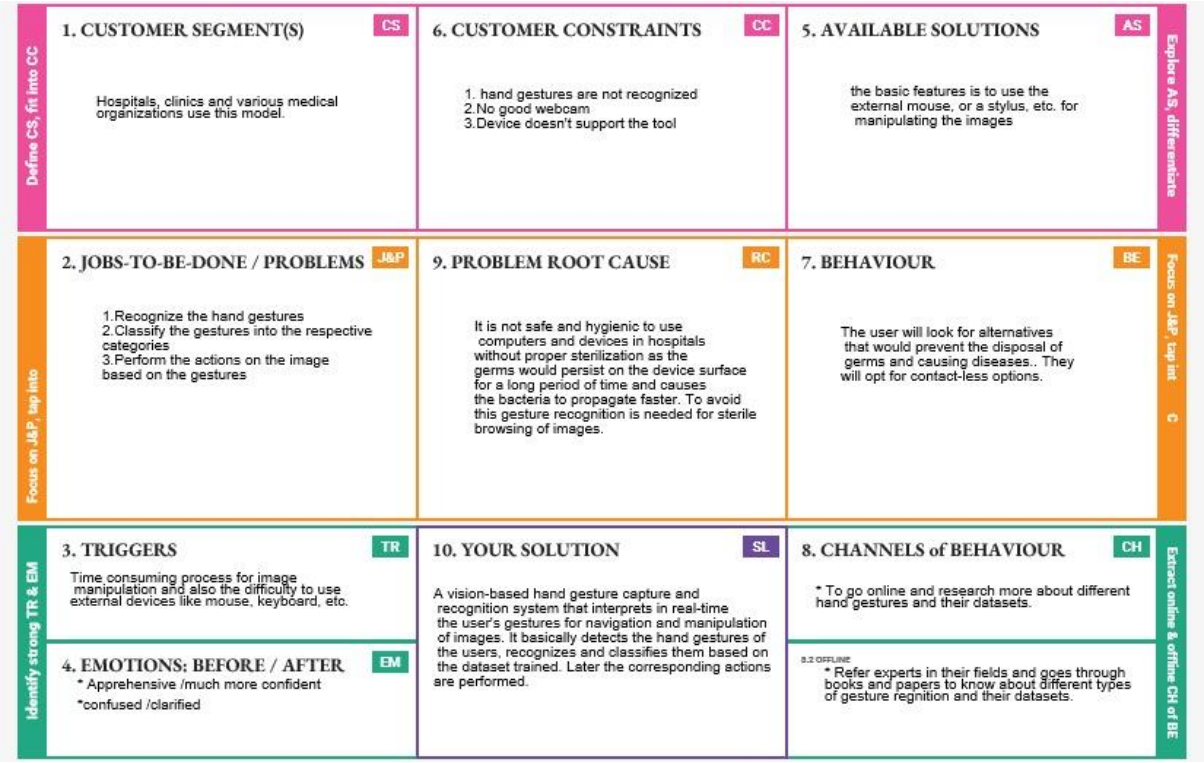
Helping educational institutes to seamlessly handle high-volume tasks, robotic process automation tools can bring in the much-desired efficiency in the administrative process.

Robotic Process Automation, as the definition suggests, makes it a cakewalk for you to execute tedious, heavy paperwork, or repetitive tasks. It does so with more efficiency and zero error rate. So, what kind of tasks can you automate at your educational institution? Be its administrative tasks, academic ones, or finance & HR tasks, you can easily automate them all with robotic process automation. At least 60% of these tasks can be efficiently done by RPA tools and that way educators will have more time to do what they do best i.e. teaching.

3.4 Problem Solution fit

The Problem-Solution Fit simply means that you have found a problem with your customer and that the solution you have realized for it solves the customer's problem.

What are important go/no-go parts of your business, product, or service? Map these out and, as much as possible, Once you have found a quantifiable problem with the customer, it is fairly easy to measure the effectiveness of your solution. Make sure you have sufficient evidence that your product or service solves the customer's problem.



4. REQUIREMENT ANALYSIS

4.1 Functional requirement

Functional requirements specify the main technical functionalities and specifications that the system should incorporate.The overview of the hand gesture recognition system consists of the following stages. The first stage is the hand gesture image capture stage where the images are taken using a digital camera under different conditions such as scaling, translation, and rotation. The second stage is a pre-processor stage in which edge detection, smoothing, and other filtering processes occur. In the next stage, the features of the images of hand gestures are extracted using two methods, namely, hand contour and complex moments. The last stage is the classification using an Artificial Neural Network (ANN), where the recognition rate is calculated for both hand contour-based ANN and complex moments-based ANN, and a comparison are carried out.

This software shall utilize a face detection system to filter out faces from the video-capturing device. By applying face detection, the system can disregard the region where the face is located and thus reducing the amount of calculation needed to perform hand detection.Once the program has filtered out most of the unwanted parts of the picture after using the skin detection module, the software shall read and recognize “clusters” of skin-colored objects also known as “blobs”.

4.2 Non-Functional requirements

Non-functional requirements specify the criteria in the operation and the architecture of the system.This software shall minimize the use of the Central Processing Unit (CPU) and memory resources on the operating system. When HGR is executing, the software shall utilize less than 80% of the system’s CPU resource and less than 100 megabytes of system memory.

The software shall be extensible to support future developments and add-ons to the HGR software. The gesture control module of HGR shall be at least 50% extensible to allow new gesture recognition features to be added to the system.

The HGR software shall be 100% portable to all operating platforms that support Java Runtime Environment (JRE). Therefore, this software should not depend on different operating systems.

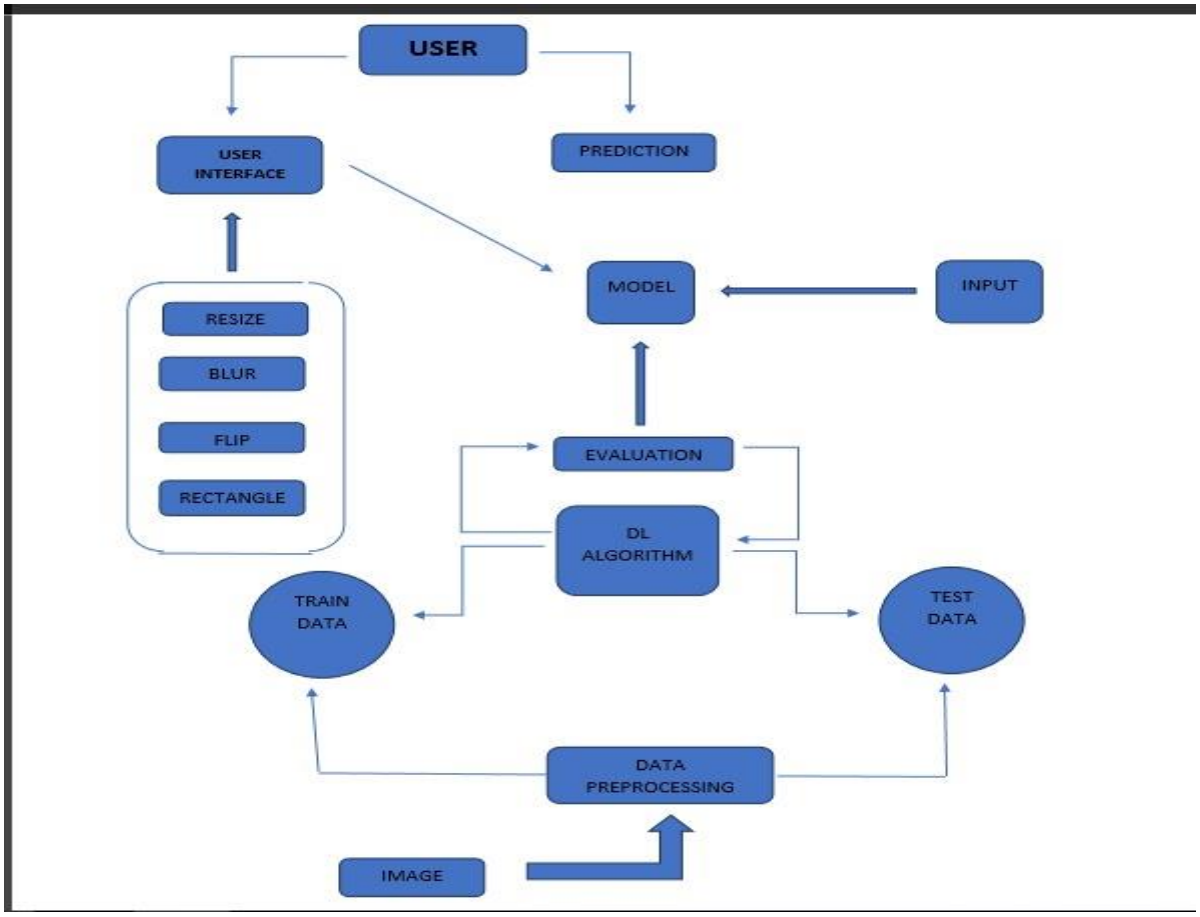
This software shall minimize the number of calculations needed to perform image processing and hand gesture detection. Each captured video frame shall be processed within 350 milliseconds to achieve 3 frames-per-secondperformance.

5. PROJECT DESIGN

The dependencies from open-source image processing libraries, software logic cycles, software constraints, and design components of the HGR software are discussed in detail.

5.1 Data Flow Diagrams

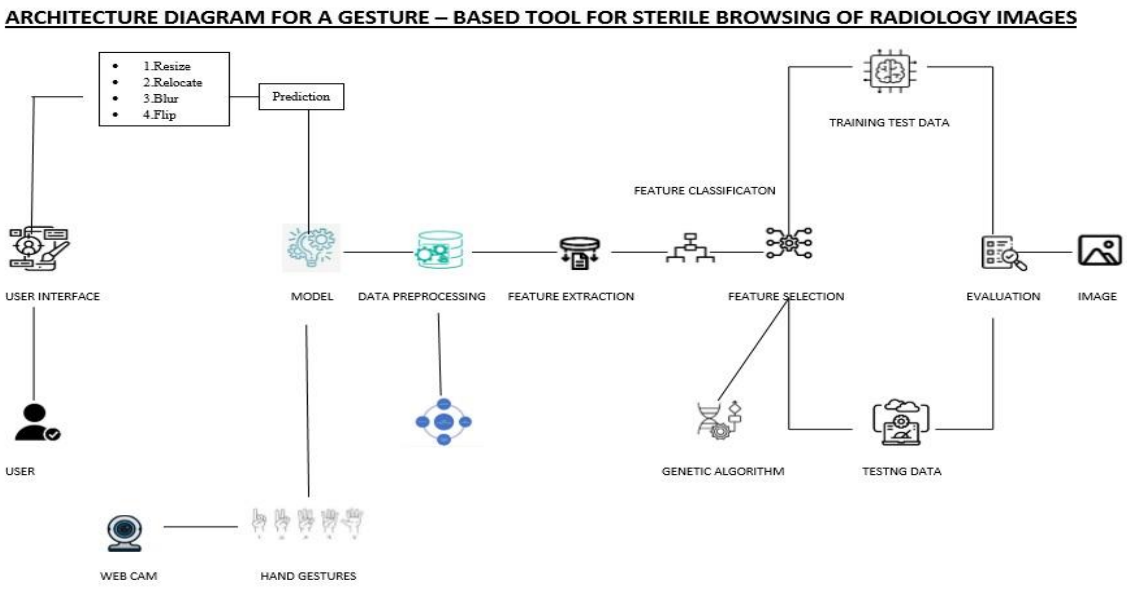
Data Flow Diagram (DFD) is part of the SSADM method (Structured Systems Analysis and Design Methodology), intended for analysis and information systems projection. Data Flow Diagrams are intended for the graphical representation of data flows in the information system and for analysis of data processing during the structural projection. Using data flow diagrams, it is possible to show visually the work of the information system and the results of this work. Data Flow Diagram visualizes processes, data depositories, and external entities in information systems and the data flow connecting these elements.



5.2 Solution & Technical Architecture

Gesture operations are initiated by a calibration mode in which a skin color model of the user's hand or glove, and under locallights under local trusted. In a browse mode, superimposed over the image of the camera's scene is a rectangularframe 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 neighbor image. To evoke a zoom mode, the open palm is rotated within the "neutral area" clockwise/counterclockwise (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 notepad pages up/down. The rotation gesture (zoom-in/zoom-out commands) reminds one of a radio knob to increase or decrease volume. Dropping the hand (stop-tracking command) is associated with the idea of 'stop-play-ing', while the waving gesture ("wake-up" command) is associated with 'greeting a new person.



5.3 User Stories

After creating the cross-validation sets, three user-independent classification methods were used to classify the datasets obtained. Two of these methods were briefly introducedand include the Adaptive Least Squares Support Vector Machines (LS-SVM) and the Bilinear Model-based classification method. The third classification method was implemented using a Multilayer Perceptron (MLP) Network, which is a type of artificial neural network (ANN). This method has been used in the pastto classify unseen data that are not easily linearly separable, e.g., EMG data. The decision to include a simple MLP Network in this study was based on the results from a previous pilot study.

It is important to mention that each method has its minimum requirements of repetitions per gesture to work correctly. In the case of the Adaptive LS-SVM and the Bilinear Models-based classifiers, the minimum number of repetitions per gesture is one. However, having 10 repetitions per gesture allows some of the user-independent classification models used in this study to perform better. For example, because the working principle of the adaptive LS-SVM is based on updating a predictive model that has similar data distributions with new data, having more repetitions per gesture increases the chances of finding said predictive model. Similarly, for the Bilinear Models-based classifier, having more repetitions per gesture allows the extraction of better motion-dependent factors during the creation of the bilinear models. The implementation of these classification methods is explained.

6. PROJECT PLANNING & SCHEDULING

6.1SprintPlanning&Estimation

IDENTIFYING THE PROBLEM STATEMENT:

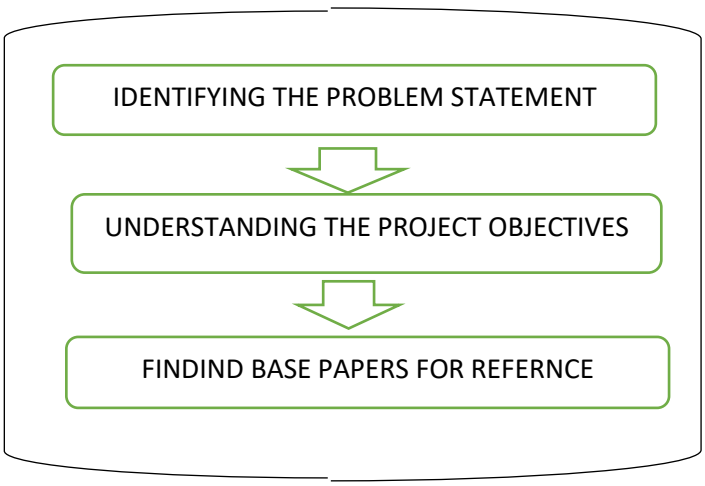
It is necessary to identify the problem for any problem. In our project the problem we are trying to overcome is using hand gestures to perform actions on the radiology images, this is useful in the medical field.

UNDERSTANDING THE PROJECT OBJECTIVES:

Identifying the method to solve the problem. Selecting the best algorithms, software, and technology to develop an application that enables hand gesture recognition.

FINDING BASE PAPERS FOR REFERENCE:

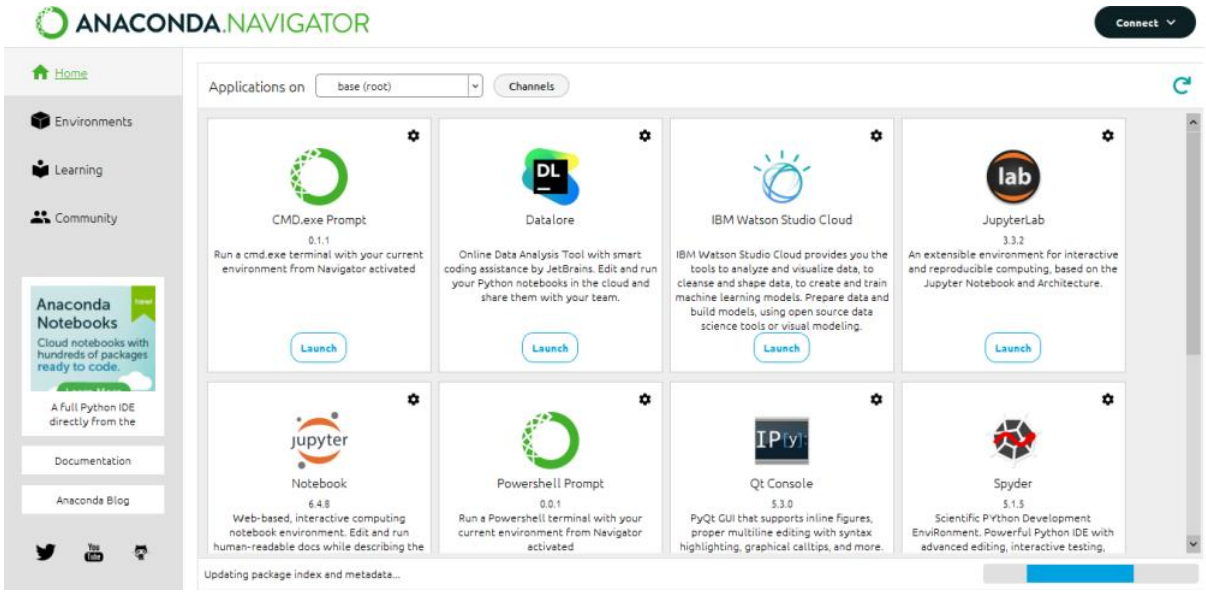
Look for base papers that were previously published that aimed at solving similar problems. Try to take adaptations from it and makes modifications for the drawbacks of the previous papers.



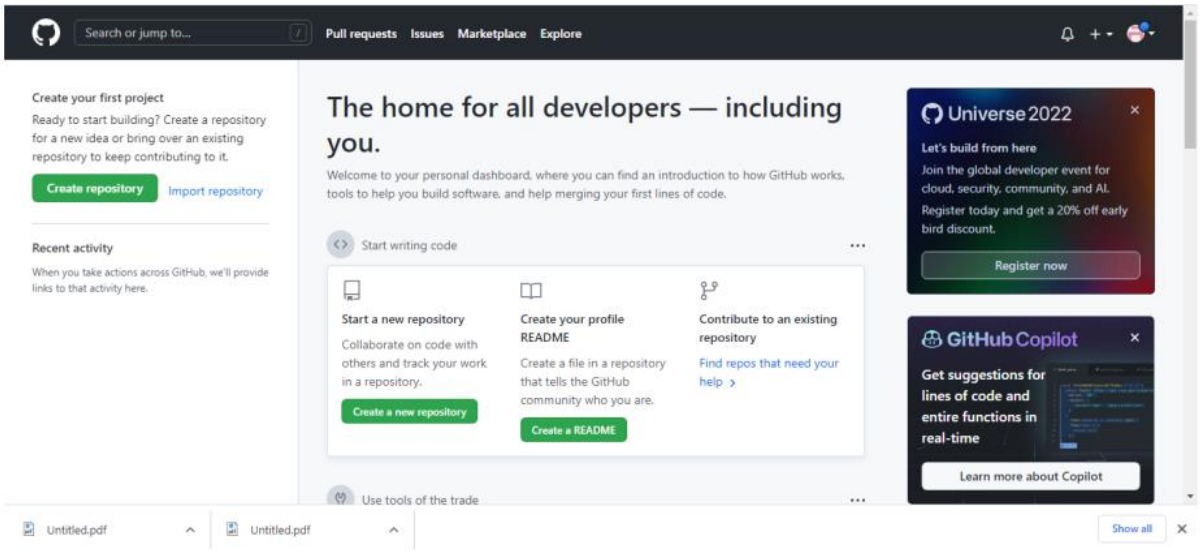
6.2 Sprint Delivery Schedule

PREREQUISITES:

For this project, we must download and install anaconda navigator, python, Jupyter notebook, and pip libraries



CREATING A GIT HUB ACCOUNT:



PRIOR KNOWLEDGE:

Understand and learn about the deep learning concepts such as

1. CNN
2. OpenCV
3. Flask

7. CODING & SOLUTIONING

7.1 Feature

In addition to allowing sterile interaction with EMRs, the "*Gestix*" 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

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

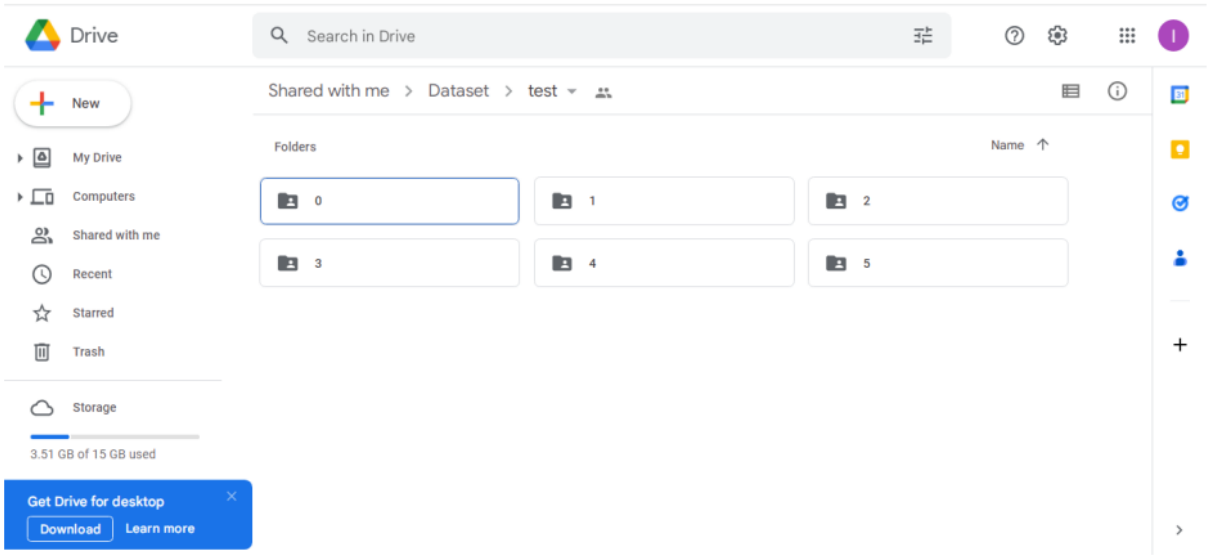
(iv) distance control—the hand gestures can be performed up to 5 meters from the camera and still be recognized accurately.

8.TESTING

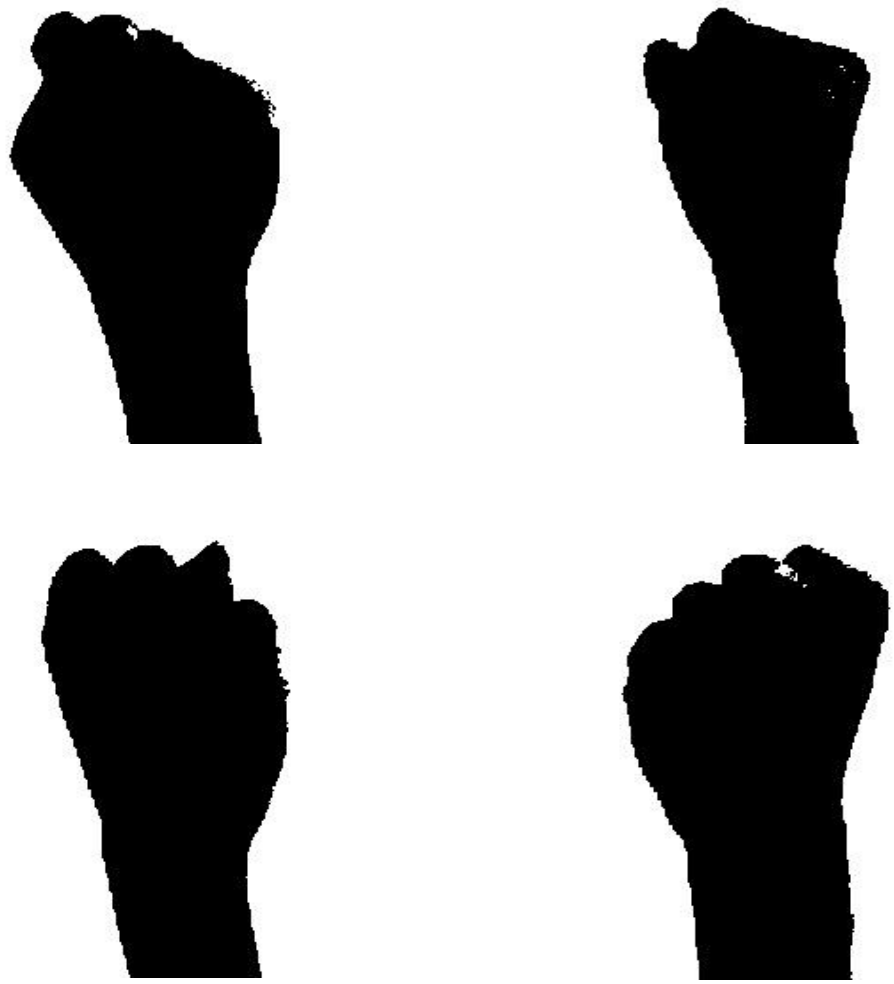
8.1Test Cases

the RGB algorithm allowed the system to find the red indicators on the user's fingers by varying the color threshold and comparing pixel by pixel from the captured video frame.

TEST DATA:



the user can use his/her finger to control the mouse movement by waving the finger with the red indicator in front of the webcam. To perform mouse clicks, the user needs to raise the second finger with the red indicator. This method of detection performed very well; however, it relies on the fact that the user needs to wear indicators on one's fingers for the detection to work. Since the indicator color is arbitrary chosen, if it coincides with other background objects of the same color, the detection system will be affected and thus fail to recognize the controller file



1.1 User Acceptance Testing

Three different types of usability tests were conducted with the "*Gestix*" system in the OR:

(i) a contextual interview;
(ii) an individual interview;
(iii) a subjective satisfaction questionnaire.

The main result of the contextual interview which was based on watching and listening to the users while they work indicated that

the main surgeon had to remain sterile, close to the patient, and avoid distractions (change in focus of attention), which could be life-threatening. The main issues found in a 20-minute interview with the main surgeon were

(a) the need of replacing the plastic adhesive cover from the touch-screen monitor for every new surgery to keep it sterile

(b) the delay was caused by frequent visits of the surgeon to the main control wall and to return to the patient's side

(c) the surgeon preferred hand gesture control since it is based on interaction with hands, in which he/she is most proficient

9. RESULTS

9.1 Performance Metrics

As outlined in the project proposal, the purpose of this project is to allow users to execute hand gestures that a regular consumer PC webcam can recognize. The primary objective is to have hand detections without using indicators and special webcams. This is achieved during the second release of HGR where the software uses color filtering techniques to obtain hand recognition. However, after version 3.0, when the skin color detection method is added, the usability of the software is further enhanced. Users no longer need to manually calibrate the system at each start-up for hand recognition and do not require special colored indicators for tracking.

The secondary objective for this project is to convert hand detection results intouser-beneficial operations. This is achieved by having two built-in control modes: mouse control and browsing control modes. Users can control mouse cursor movements by just moving one's hands in front of the webcam. By raising two hands, the user can simulate a click on the mouse. In addition, in browsing mode, users can switch, scroll up and scroll down pages in Internet browsers. Through three different versions of the software, both objectives discussed in the proposal are accomplished. The software is also built in many modules where future add-ons to the project can be easily attained.

10. ADVANTAGES & DISADVANTAGES

Advantages:

The 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, th is helps to save time.
The tool does not need the person using it to have an apparatus or any devices o n 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 devi ces to capture images and process them.

The tool can be quite expensive as it requires cameras and other expensive devi ces to capture images and process them.

11. CONCLUSION

This report investigated the alternative solution to using integrated hardware components with the personal computer. The project presented a program that allowed users to perform hand gestures for easy software control. Through various algorithms and detection techniques, the program achieved simple hand detection and identified the accurate gestures executed by the user.

In achieving the hand detection module, the program utilized the combination of RSB, HSB, skin color, and face detection methods to determine user hand locations. After the hand location was defined, geometry point conversion and gesture recognition modules were called upon to execute user gesture commands.

Under extensive testing done in different locations and lighting conditions, the program showed that:

- Environment lighting changes could dramatically decrease the robustness of the software system in detecting hands

- The addition of a face detection module could decrease the performance speed of image processing but increase the accuracy of hand detection

12. FUTURE SCOPE

For future developments, it would be ideal to research advanced mathematical materials for image processing and investigate different hardware solutions that would result in more accurate hand detections. Not only did this project show the different gesture operations that could be done by the users but it also demonstrated the potential in simplifying user interactions with personal computers and hardware systems.

A feasibility study is an analysis to determine ifthe project idea is viable technically and medically feasible and is also applicable to real-time scenarios. This project is feasible becausethe resource is accessibleto the user.

13. APPENDIX

Source Code:

Home.html:

```
<html>
<script>


</script>


<style>
.header {       position: relative;
                        top:0;
                        margin:0px;
                        z-index: 1;
                        left: 0px;
                        right: 0px;
                        position: fixed;
                        background-color:rgb(10, 102, 109) ;
                        color: rgb(181, 228, 236);
                        box-shadow: 0px 8px 4px rgb(10, 102, 109);
                        overflow: hidden;
                        padding-left:20px;
                        font-family: 'Times New Roman';
                        font-size: 2vw;
                        width: 100%;
                        height:8%;
                        text-align: center;
                }
                .topnav {
  overflow: hidden;
  background-color: #84d39e;
}
```

```
.topnav-right a {
  float: left;
  color: black;
  text-align: center;
  padding: 14px 16px;
  text-decoration: none;
  font-size: 18px;
}
.topnav-right a.active {
  background-color: #07201e;
  color: rgb(238, 226, 234);
}
.topnav-right a:hover {
  background-color: rgb(181, 228, 236);
  color: rgb(6, 27, 36);
}

.topnav-right {
  float: right;
  padding-right:100px;
}

body {

  background-color: rgb(88, 129, 123) ;
  background-repeat: no-repeat;
  background-size:cover;
  background-position: 0px 0px;
  }
  .button {
  background-color: #091425;
  border: none;
  color: rgb(181, 228, 236);
  padding: 15px 32px;
  text-align: center;
  text-decoration: none;
  display: inline-block;
  font-size: 12px;
  border-radius: 16px;
}
.button:hover {
  box-shadow: 0 12px 16px 0 rgba(0,0,0,0.24), 0 17px 50px 0
rgba(0,0,0,0.19);
}
form {border: 3px solid #f1f1f1; margin-left:400px;margin-right:400px;}

input[type=text], input[type=password] {
  width: 100%;
  padding: 12px 20px;
  display: inline-block;
  margin-bottom:18px;
  border: 1px solid #ccc;
  box-sizing: border-box;
}

button {
  background-color: #091425;
  color: rgb(181, 228, 236);
  padding: 14px 20px;
  margin-bottom:10px;
  border: none;
```

```
  cursor: pointer;
  width: 17%;
  border-radius:4px;
  font-family:Montserrat;
}

button:hover {
  opacity: 0.8;
}

.cancelbtn {
  width: auto;
  padding: 10px 18px;
  background-color: rgb(181, 228, 236);
}

.imgcontainer {
  text-align: center;
  margin: 24px 0 12px 0;
}

img.avatar {
  width: 30%;
  border-radius: 50%;
}

.container {
  padding: 16px;
}

span.psw {
  float: right;
  padding-top: 16px;
}

/* Change styles for span and cancel button on extra small screens */
@media screen and (max-width: 300px) {
  span.psw {
     display: block;
     float: none;
  }
  .cancelbtn {
     width: 100%;
  }
}

.home{
        margin:80px;

  width: 84%;
  height: 500px;
  padding-top:10px;
  padding-left: 30px;

}
.login{
        margin:80px;
        box-sizing: content-box;
  width: 84%;
  height: 420px;
  padding: 30px;
```

```
  border: 10px solid rgb(13, 53, 68);
}
.left,.right{
 box-sizing: content-box;
 height: 400px;
 margin:20px;
 border: 10px solid rgb(13, 53, 68);
}

.mySlides {display: none;}
img {vertical-align: middle;}

/* Slideshow container */
.slideshow-container {
  max-width: 1000px;
  position: relative;
  margin: auto;
}

/* Caption text */
.text {
  color: #9ac0c0;
  font-size: 15px;
  padding: 8px 12px;
  position: absolute;
  bottom: 8px;
  width: 100%;
  text-align: center;
}
/* The dots/bullets/indicators */
.dot {
  height: 15px;
  width: 15px;
  margin: 0 2px;
  background-color: #bbb;
  border-radius: 50%;
  display: inline-block;
  transition: background-color 0.6s ease;
}

.active {
  color: rgb(145, 216, 221);
}

/* Fading animation */
.fade {
  -webkit-animation-name: fade;
  -webkit-animation-duration: 1.5s;
  animation-name: fade;
  animation-duration: 1.5s;
}

@-webkit-keyframes fade {
  from {opacity: .4}
  to {opacity: 1}
}

@keyframes fade {
  from {opacity: .4}
  to {opacity: 1}
}
```

```
/* On smaller screens, decrease text size */
@media only screen and (max-width: 300px) {
  .text {font-size: 11px}
}



@import url("https://fonts.googleapis.com/css?family=Luckiest+Guy");
/* BODY */
body {
  position: absolute;
  top: 0;
  left: 0;
  right: 0;
  bottom: 0;
  width: 100%;
  height: 100%;
  overflow: hidden;
  font-family: "Arial", cursive;
  -webkit-font-smoothing: antialiased;
}

::selection {
  background: transparent;
}
/* CLOUDS */
body:before {
  content: "";
  position: absolute;
  top: 0;
  left: 0;
  right: 0;
  width: 0;
  height: 0;
  margin: auto;
  border-radius: 100%;
  background: transparent;
  display: block;
  box-shadow: 0 0 150px 100px rgba(255, 255, 255, 0.6),
    200px 0 200px 150px rgba(255, 255, 255, 0.6),
    -250px 0 300px 150px rgba(255, 255, 255, 0.6),
    550px 0 300px 200px rgba(255, 255, 255, 0.6),
    -550px 0 300px 200px rgba(255, 255, 255, 0.6);
}
/* JUMP */
h1 {
  cursor: default;
  position: absolute;
  top: 0;
  left: 0;
  right: 0;
  bottom: 0;
  width: 100%;
  height: 100px;
  margin: 70px;
  display: block;
  text-align: center;
}
```

```
h1 span {
  position: relative;
  top: 5px;
  display: inline-block;
  font-size: 25px;
  color: #061a1f;
  text-shadow: 0 1px 0 rgb(151, 201, 197), 0 2px 0 rgb(151, 201, 197), 0
3px 0 rgb(151, 201, 197), 0 4px 0 rgb(151, 201, 197),
    0 5px 0 rgb(151, 201, 197), 0 6px 0 transparent, 0 7px 0 transparent, 0
8px 0 transparent,
    0 9px 0 transparent, 0 10px 10px rgba(58, 159, 167, 0.4);
}

h1 span:nth-child(2) {
  -webkit-animation-delay: 0.1s;
}

h1 span:nth-child(3) {
  -webkit-animation-delay: 0.2s;
}

h1 span:nth-child(4) {
  -webkit-animation-delay: 0.3s;
}

h1 span:nth-child(5) {
  -webkit-animation-delay: 0.4s;
}

h1 span:nth-child(6) {
  -webkit-animation-delay: 0.5s;
}

h1 span:nth-child(7) {
  -webkit-animation-delay: 0.6s;
}

h1 span:nth-child(8) {
  -webkit-animation-delay: 0.2s;
}

h1 span:nth-child(9) {
  -webkit-animation-delay: 0.3s;
}

h1 span:nth-child(10) {
  -webkit-animation-delay: 0.4s;
}

h1 span:nth-child(11) {
  -webkit-animation-delay: 0.5s;
}

h1 span:nth-child(12) {
  -webkit-animation-delay: 0.6s;
}

h1 span:nth-child(13) {
  -webkit-animation-delay: 0.7s;
}
```

```
h1 span:nth-child(14) {
  -webkit-animation-delay: 0.8s;
}

/* ANIMATION */
@-webkit-keyframes bounce {
  100% {
    top: -20px;
    text-shadow: 0 1px 0 #ccc, 0 2px 0 #ccc, 0 3px 0 #ccc, 0 4px 0 #ccc,
      0 5px 0 #ccc, 0 6px 0 #ccc, 0 7px 0 #ccc, 0 8px 0 #ccc, 0 9px 0 #ccc,
      0 50px 25px rgba(0, 0, 0, 0.2);
  }
}

</style>

<body>
<h1 style="color: rgb(193, 207, 207);">
  <table style="width:100%">
    <tr>
      <th><img
src="https://media.tenor.com/images/30169e4a670daf12443df7d2dd140176/tenor.
gif" alt="NOT AVAILABLE" width="300" height="300"></th>
      <th>
        <!-- <iframe width="780" height="440"
src="https://www.youtube.com/embed/nD621G8u6oc?start=3&loop=1&autoplay=1&mu
te=1&controls=0">
      </iframe><br> -->
    </th>
      <th></th>
    </tr>
  </table>
    <br>
    <span>HAND   GESTURE   RECOGNITION</span>
  </br>
    <span>OF</span>
  </br>
    <span>RADIOLOGY   IMAGES</span>
  </br>
    <span>THROUGH</span>
  </br>
    <span>STERILE   BROWSING</span>

</h1>
<div class="header">
<div style="width:50%;float:left;font-size:2vw;text-
align:left;color:#c4dfd7; padding-top:1%;padding-left:5%;"><b>HAND GESTURE
RECOGNITION</b></div>
  <div class="topnav-right"style="padding-top:0.5%;color:white">
    <a  class="active" href="{{ url_for('home')}}"><u>Home</u></a>
    <a  class="active" href="{{ url_for('intro')}}">Introduction</a>
    <a  class="active" href="{{ url_for('image1')}}">Launch</a>
  </div>
</div>

</body>

</html>
```

Intro.html:

```
<html>
<script>


</script>


<style>
.header {       position: relative;
                        top:0;
                        margin:0px;
                        z-index: 1;
                        left: 0px;
                        right: 0px;
                        position: fixed;
                        background-color: rgb(10, 102, 109) ;
                        color: white;
                        box-shadow: 0px 8px 4px grey;
                        overflow: hidden;
                        padding-left:20px;
                        font-family: 'Josefin Sans';
                        font-size: 2vw;
                        width: 100%;
                        height:8%;
                        text-align: center;
                }
                .topnav {
  overflow: hidden;
  background-color: #FCAD98;
}

.topnav-right a {
  float: left;
  color: black;
  text-align: center;
  padding: 14px 16px;
  text-decoration: none;
  font-size: 18px;
}

.topnav-right a.active {
  background-color: #07201e;
  color: rgb(238, 226, 234);
}

.topnav-right a:hover {
  background-color: rgb(181, 228, 236);
  color: rgb(6, 27, 36);
}

.topnav-right {
  float: right;
  padding-right:100px;
}

body {

  background-color: ;
  background-repeat: no-repeat;
```

```
  background-size:cover;
  background-image:
url("https://i.pinimg.com/originals/b2/1d/c6/b21dc69346915015bc4e19bd502f40
1b.gif");
    background-size: cover;
  background-position: 0px 0px;
  }
  .button {
  background-color: #091425;
  border: none;
  color: white;
  padding: 15px 32px;
  text-align: center;
  text-decoration: none;
  display: inline-block;
  font-size: 12px;
  border-radius: 16px;
}
.button:hover {
  box-shadow: 0 12px 16px 0 rgba(0,0,0,0.24), 0 17px 50px 0
rgba(0,0,0,0.19);
}
form {border: 3px solid #f1f1f1; margin-left:400px;margin-right:400px;}

input[type=text], input[type=password] {
  width: 100%;
  padding: 12px 20px;
  display: inline-block;
  margin-bottom:18px;
  border: 1px solid #ccc;
  box-sizing: border-box;
}

button {
  background-color: #091425;
  color: white;
  padding: 14px 20px;
  margin-bottom:10px;
  border: none;
  cursor: pointer;
  width: 17%;
  border-radius:4px;
  font-family:Montserrat;
}

button:hover {
  opacity: 0.8;
}

.cancelbtn {
  width: auto;
  padding: 10px 18px;
  background-color: #f44336;
}

.imgcontainer {
  text-align: center;
  margin: 24px 0 12px 0;
}

img.avatar {
```

```
  width: 30%;
  border-radius: 50%;
}

.container {
  padding: 16px;
}

span.psw {
  float: right;
  padding-top: 16px;
}

/* Change styles for span and cancel button on extra small screens */
@media screen and (max-width: 300px) {
  span.psw {
     display: block;
     float: none;
  }
  .cancelbtn {
     width: 100%;
  }
}

.home{
        margin:80px;

  width: 84%;
  height: 500px;
  padding-top:10px;
  padding-left: 30px;

}
.login{
        margin:80px;
        box-sizing: content-box;
  width: 84%;
  height: 420px;
  padding: 30px;
  border: 10px solid blue;
}
.left,.right{
 box-sizing: content-box;
 height: 400px;
 margin:20px;
 border: 10px solid blue;
}

.mySlides {display: none;}
img {vertical-align: middle;}

/* Slideshow container */
.slideshow-container {
  max-width: 1000px;
  position: relative;
  margin: auto;
}

/* Caption text */
.text {
  color: #f2f2f2;
```

```
  font-size: 15px;
  padding: 8px 12px;
  position: absolute;
  bottom: 8px;
  width: 100%;
  text-align: center;
}
/* The dots/bullets/indicators */
.dot {
  height: 15px;
  width: 15px;
  margin: 0 2px;
  background-color: #bbb;
  border-radius: 50%;
  display: inline-block;
  transition: background-color 0.6s ease;
}

.active {
  background-color: #FCAD98;
}

/* Fading animation */
.fade {
  -webkit-animation-name: fade;
  -webkit-animation-duration: 1.5s;
  animation-name: fade;
  animation-duration: 1.5s;
}

@-webkit-keyframes fade {
  from {opacity: .4}
  to {opacity: 1}
}

@keyframes fade {
  from {opacity: .4}
  to {opacity: 1}
}

/* On smaller screens, decrease text size */
@media only screen and (max-width: 300px) {
  .text {font-size: 11px}
}



@import
url("https://fonts.googleapis.com/css?family=Montserrat&display=swap");

* {
  padding: 0;
  margin: 0;
}

body {
  height: 100vh;
  display: flex;
  flex-direction: column;
  justify-content: center;
  align-items: center;
```

```
}

h1 {
  font-family: "Montserrat Medium";
  max-width: 90ch;
  text-align: center;
  transform: scale(0.94);
  animation: scale 3s forwards cubic-bezier(0.5, 1, 0.89, 1);
}
@keyframes scale {
  100% {
    transform: scale(1);
  }
}

span {
  display: inline-block;
  opacity: 0;
  filter: blur(4px);
}

span:nth-child(1) {
  animation: fade-in 1s 0.1s forwards cubic-bezier(0.11, 0, 0.5, 0);
}

span:nth-child(2) {
  animation: fade-in 0.8s 0.2s forwards cubic-bezier(0.11, 0, 0.5, 0);
}

span:nth-child(3) {
  animation: fade-in 0.8s 0.3s forwards cubic-bezier(0.11, 0, 0.5, 0);
}

span:nth-child(4) {
  animation: fade-in 0.8s 0.4s forwards cubic-bezier(0.11, 0, 0.5, 0);
}

span:nth-child(5) {
  animation: fade-in 0.8s 0.5s forwards cubic-bezier(0.11, 0, 0.5, 0);
}

span:nth-child(6) {
  animation: fade-in 0.8s 0.6s forwards cubic-bezier(0.11, 0, 0.5, 0);
}

span:nth-child(7) {
  animation: fade-in 0.8s 0.7s forwards cubic-bezier(0.11, 0, 0.5, 0);
}

span:nth-child(8) {
  animation: fade-in 0.8s 0.8s forwards cubic-bezier(0.11, 0, 0.5, 0);
}

span:nth-child(9) {
  animation: fade-in 0.8s 0.9s forwards cubic-bezier(0.11, 0, 0.5, 0);
}

span:nth-child(10) {
  animation: fade-in 0.8s 1s forwards cubic-bezier(0.11, 0, 0.5, 0);
}
```

```
span:nth-child(11) {
  animation: fade-in 0.8s 1.1s forwards cubic-bezier(0.11, 0, 0.5, 0);
}

span:nth-child(12) {
  animation: fade-in 0.8s 1.2s forwards cubic-bezier(0.11, 0, 0.5, 0);
}

span:nth-child(13) {
  animation: fade-in 0.8s 1.3s forwards cubic-bezier(0.11, 0, 0.5, 0);
}

span:nth-child(14) {
  animation: fade-in 0.8s 1.4s forwards cubic-bezier(0.11, 0, 0.5, 0);
}

span:nth-child(15) {
  animation: fade-in 0.8s 1.5s forwards cubic-bezier(0.11, 0, 0.5, 0);
}

span:nth-child(16) {
  animation: fade-in 0.8s 1.6s forwards cubic-bezier(0.11, 0, 0.5, 0);
}

span:nth-child(17) {
  animation: fade-in 0.8s 1.7s forwards cubic-bezier(0.11, 0, 0.5, 0);
}

span:nth-child(18) {
  animation: fade-in 0.8s 1.8s forwards cubic-bezier(0.11, 0, 0.5, 0);
}
span:nth-child(19) {
  animation: fade-in 0.8s 1.9s forwards cubic-bezier(0.11, 0, 0.5, 0);
}
span:nth-child(20) {
  animation: fade-in 0.8s 2.0s forwards cubic-bezier(0.11, 0, 0.5, 0);
}
span:nth-child(21) {
  animation: fade-in 0.8s 2.1s forwards cubic-bezier(0.11, 0, 0.5, 0);
}
span:nth-child(22) {
  animation: fade-in 0.8s 2.2s forwards cubic-bezier(0.11, 0, 0.5, 0);
}
span:nth-child(23) {
  animation: fade-in 0.8s 2.3s forwards cubic-bezier(0.11, 0, 0.5, 0);
}span:nth-child(24) {
  animation: fade-in 0.8s 2.4s forwards cubic-bezier(0.11, 0, 0.5, 0);
}span:nth-child(25) {
  animation: fade-in 0.8s 2.5s forwards cubic-bezier(0.11, 0, 0.5, 0);
}span:nth-child(26) {
  animation: fade-in 0.8s 2.6s forwards cubic-bezier(0.11, 0, 0.5, 0);
}span:nth-child(27) {
  animation: fade-in 0.8s 2.7s forwards cubic-bezier(0.11, 0, 0.5, 0);
}span:nth-child(28) {
  animation: fade-in 0.8s 2.8s forwards cubic-bezier(0.11, 0, 0.5, 0);
}
@keyframes fade-in {
  100% {
    opacity: 1;
    filter: blur(0);
  }
```

```
}

</style>

<body>
  <h1>INTRODUCTION</h1>
<h1>
  <span> Hand Gesture </span> <span> recognition system </span> <span>
provides us </span > <span> an innovative,</span> <span>natural,</span>
<span>  user friendly </span> <span> way of interaction </span > <span>
with the computer</span>
  <span> which is more </span> <span> familiar to the </span> <span>human
beings. </span > <span> In our project, </span> <span> the hand region
</span> <span> is extracted from </span> <span> the background  </span >
<span> by using</span>
<span> Region of intrest. </span> <span> Then, </span> <span> we will be
</span > <span>predicting the labels </span> <span> based on the </span>
<span> CNN trained model weights  </span> <span> of hand gestures </span >
<span> using that predicted labels</span>
<span> we apply if conditions </span> <span> to control some of the actions
</span> <span>like </span > <span>reshaping , blur, flip of the given
image.</span>

</h1>
<!--Brian Tracy-->

<div class="header">
<div style="width:50%;float:left;font-size:2vw;text-
align:left;color:#c1e2d9; padding-top:1%;padding-left:5%;">Hand Gesture
System</div>
  <div class="topnav-right"style="padding-top:0.5%;">

    <a class="active" href="{{ url_for('home')}}">Home</a>
    <a class="active" href="{{ url_for('intro')}}"><u>Introduction</u></a>
    <a class="active" href="{{ url_for('image1')}}">Launch</a>
  </div>
</div>
</body>

</html>
```

Launch.html:

```
<html lang="en">

  <head>
  <meta charset="utf-8">
      <meta http-equiv="X-UA-Compatible" content="IE=edge">
      <meta name="viewport" content="width=device-width, initial-
scale=0.6">
      <script
src="https://ajax.googleapis.com/ajax/libs/jquery/3.3.1/jquery.min.js"></sc
ript>

      <link href="https://fonts.googleapis.com/icon?family=Material+Icons"
rel="stylesheet">
      <meta charset="UTF-8">
      <title>Predict</title>
```

```
      <link
href="https://cdn.bootcss.com/bootstrap/4.0.0/css/bootstrap.min.css"
rel="stylesheet">

      <script
src="https://cdn.bootcss.com/popper.js/1.12.9/umd/popper.min.js"></script>
      <script
src="https://cdn.bootcss.com/jquery/3.3.1/jquery.min.js"></script>
      <script
src="https://cdn.bootcss.com/bootstrap/4.0.0/js/bootstrap.min.js"></script>
      <link href="{{ url_for('static', filename='css/main.css') }}"
rel="stylesheet">
  <style>
  .bar
  {
  margin: 0px;
  padding:20px;
  background-color:black;
  opacity:0.6;
  color:black;
  font-family:'Roboto',sans-serif;
  font-style: italic;
  border-radius:20px;
  font-size:25px;
  }
  a
  {
  color:black;
  float:right;
  text-decoration:none;
  font-style:normal;
  padding-right:20px;
  }
  a:hover{
  background-color:black;
  color:black;
  font-size:30px;
  padding-left:10px;
  }

  div1{
    text-align: center;
    width: 650spx;
    height: 800px;
    padding: 190px;
    margin: 10px;
    position: absolute;
  }



  body
  {
      background-image: url("./img.png");
      background-size: cover;
  }

  .header {    position: relative;
        top:0;
        margin:0px;
        z-index: 1;
```

```
        left: 0px;
        right: 0px;
        position: fixed;
        background-color: rgb(10, 102, 109) ;
        color: black;
        box-shadow: 0px 8px 4px grey;
        overflow: hidden;
        padding-left:20px;
        font-family: 'Josefin Sans';
        font-size: 2vw;
        width: 100%;
        height:8%;
        text-align: center;
      }
      .topnav {
    overflow: hidden;
    background-color: #056959;
  }

  .topnav-right a {
    float: left;
    color: black;
    text-align: center;
    padding: 14px 16px;
    text-decoration: none;
    font-size: 18px;
  }

  .topnav-right a.active {
    background-color: #07201e;
    color: rgb(238, 226, 234);
  }

  .topnav-right a:hover {
    background-color: rgb(181, 228, 236);
    color: rgb(6, 27, 36);
  }



  .topnav-right {
    float: right;
    padding-right:100px;
  }


  .button {
  background-color: #091425;
  border: none;
  color: black;
  padding: 15px 32px;
  text-align: center;
  text-decoration: none;
  display: inline-block;
  font-size: 12px;
  border-radius: 16px;
  }
  .button:hover {
    box-shadow: 0 12px 16px 0 rgba(0,0,0,0.24), 0 17px 50px 0
rgba(0,0,0,0.19);
  }
```

```
form {border: 2px solid black; margin-left:400px;margin-right:400px;}

input[type=text], input[type=password] {
  width: 100%;
  padding: 12px 20px;
  display: inline-block;
  margin-bottom:18px;
  border: 1px solid #ccc;
  box-sizing: border-box;
}

button {
  background-color: #091425;
  color: black;
  padding: 14px 20px;
  margin-bottom:10px;
  border: none;
  cursor: pointer;
  width: 17%;
  border-radius:4px;
  font-family:Montserrat;
}

button:hover {
  opacity: 0.8;
}

.cancelbtn {
  width: auto;
  padding: 10px 18px;
  background-color: #f44336;
}

.imgcontainer {
  text-align: center;
  margin: 24px 0 12px 0;
}

img.avatar {
  width: 30%;
  border-radius: 50%;
}

.container {
  padding: 16px;
}

span.psw {
  float: right;
  padding-top: 16px;
}

/* Change styles for span and cancel button on extra small screens */
@media screen and (max-width: 300px) {
  span.psw {
     display: block;
     float: none;
  }
  .cancelbtn {
     width: 100%;
  }
```

```
}

.home{
  margin:80px;

  width: 84%;
  height: 500px;
  padding-top:10px;
  padding-left: 30px;

}
.login{
  margin:80px;
  box-sizing: content-box;
  width: 84%;
  height: 420px;
  padding: 30px;
  border: 10px solid rgb(12, 91, 94);
}
.left,.right{
 box-sizing: content-box;
 height: 400px;
 margin:20px;
 border: 10px solid rgb(12, 91, 94);
}

.mySlides {display: none;}
img {vertical-align: middle;}

/* Slideshow container */
.slideshow-container {
  max-width: 1000px;
  position: relative;
  margin: auto;
}

/* Caption text */
.text {
  color: #f2f2f2;
  font-size: 15px;
  padding: 8px 12px;
  position: absolute;
  bottom: 8px;
  width: 100%;
  text-align: center;
}
/* The dots/bullets/indicators */
.dot {
  height: 15px;
  width: 15px;
  margin: 0 2px;
  background-color: #bbb;
  border-radius: 50%;
  display: inline-block;
  transition: background-color 0.6s ease;
}

.active {
  background-color: #267481;
}
```

```
  /* Fading animation */
  .fade {
    -webkit-animation-name: fade;
    -webkit-animation-duration: 1.5s;
    animation-name: fade;
    animation-duration: 1.5s;
  }

  @-webkit-keyframes fade {
    from {opacity: .4}
    to {opacity: 1}
  }

  @keyframes fade {
    from {opacity: .4}
    to {opacity: 1}
  }

  /* On smaller screens, decrease text size */
  @media only screen and (max-width: 300px) {
    .text {font-size: 11px}
  }

  </style>
  </head>

  <body>
  <div class="header">
  <div style="width:50%;float:left;font-size:2vw;text-
align:left;color:#c1e2d9; padding-top:1%;padding-left:5%;">Hand Gesture
System</div>
    <div class="topnav-right"style="padding-top:0.5%;">

      <a class="active" href="{{ url_for('home')}}">Home</a>
      <a class="active" href="{{ url_for('intro')}}">Introduction</a>
      <a class="active" href="{{ url_for('image1')}}"><u>Launch</u></a>
    </div>
  </div>
  <br>

    <div1 style=""><h1><font color="Black" size="6" font-
family="Roboto">Hand Gesture Recognition</h1><br>
    <p><i><font color="Black" size="4" fonr-family="sans-serif"></i>Provide
an image for which you want to perform various operations</p>
    <br>
          <div>
            <h4>Upload Image Here</h4>
        <form action = "http://localhost:5000/" id="upload-file"
method="post" enctype="multipart/form-data">
          <label for="imageUpload" class="upload-label">
            Choose...
          </label>
          <input type="file" name="image" id="imageUpload" accept=".png,
.jpg, .jpeg,.pdf">
        </form>
             <center>
        <div class="image-section" style="display:none;">
          <div class="img-preview">
            <div id="imagePreview">
            </div>
          </div>
```

```
          <div>
            <button type="button" class="btn btn-info btn-lg " id="btn-
predict">Predict!</button>
          </div>
        </div>
        <div class="loader" style="display:none;"></div>
             </center>
      </div>
      </div1>

          <footer>
      <script src="{{ url_for('static', filename='js/main.js') }}"
type="text/javascript"></script>
  </footer>

  </html>
```

App.py:

```
from flask import Flask,render_template,request
# Flask-It is our framework which we are going to use to run/serve our
application.
#request-for accessing file which was uploaded by the user on our
application.
import operator
import cv2 # opencv library
import matplotlib.pyplot as plt
import matplotlib.image as mpimg
import numpy as np

from tensorflow.keras.models import load_model#to load our trained model
import os
from werkzeug.utils import secure_filename

app = Flask(__name__,template_folder="templates") # initializing a flask
app
# Loading the model
model=load_model('gesture.h5')
print("Loaded model from disk")


@app.route('/')# route to display the home page
def home():
    return render_template('home.html')#rendering the home page


@app.route('/intro') # routes to the intro page
def intro():
    return render_template('intro.html')#rendering the intro page

@app.route('/image1',methods=['GET','POST'])# routes to the index html
def image1():
    return render_template("launch.html")


@app.route('/predict',methods=['GET', 'POST'])# route to show the
predictions in a web UI
def launch():
```

```
    if request.method == 'POST':
        print("inside image")
        f = request.files['image']

        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)

            # Displaying the predictions
            cv2.putText(frame, prediction[0][0], (10, 120),
cv2.FONT_HERSHEY_PLAIN, 1, (0,255,255), 1)
            cv2.imshow("Frame", frame)

            #loading an image
            image1=cv2.imread(file_path)
            if prediction[0][0]=='ONE':

                resized = cv2.resize(image1, (200, 200))
                cv2.imshow("Fixed Resizing", resized)
                key=cv2.waitKey(3000)

                if (key & 0xFF) == ord("1"):
                    cv2.destroyWindow("Fixed Resizing")
```

```
            elif prediction[0][0]=='ZERO':

                cv2.rectangle(image1, (480, 170), (650, 420), (0, 0, 255),
2)
                cv2.imshow("Rectangle", image1)
                cv2.waitKey(0)
                key=cv2.waitKey(3000)
                if (key & 0xFF) == ord("0"):
                    cv2.destroyWindow("Rectangle")

            elif prediction[0][0]=='TWO':
                (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")

            elif prediction[0][0]=='THREE':
                blurred = cv2.GaussianBlur(image1, (21, 21), 0)
                cv2.imshow("Blurred", blurred)
                key=cv2.waitKey(3000)
                if (key & 0xFF) == ord("3"):
                    cv2.destroyWindow("Blurred")

            elif prediction[0][0]=='FOUR':

                resized = cv2.resize(image1, (400, 400))
                cv2.imshow("Fixed Resizing", resized)
                key=cv2.waitKey(3000)
                if (key & 0xFF) == ord("4"):
                    cv2.destroyWindow("Fixed Resizing")

            elif prediction[0][0]=='FIVE':
                '''(h, w, d) = image1.shape
                center = (w // 2, h // 2)
                M = cv2.getRotationMatrix2D(center, 45, 1.0)
                rotated = cv2.warpAffine(image1, M, (w, h))'''
                gray = cv2.cvtColor(image1, cv2.COLOR_RGB2GRAY)
                cv2.imshow("OpenCV Gray Scale", gray)
                key=cv2.waitKey(3000)
                if (key & 0xFF) == ord("5"):
                    cv2.destroyWindow("OpenCV Gray Scale")

            else:
                continue


            interrupt = cv2.waitKey(10)
            if interrupt & 0xFF == 27: # esc key
                break


        cap.release()
        cv2.destroyAllWindows()
    return render_template("home.html")

if __name__ == "__main__":
   # running the app
```

```
app.run(debug=False)
```

Project Links:

1. Data Set link:

https://drive.google.com/drive/u/0/folders/1CBmtxPdpcHLYnGesQcUidz27DpuefuMI

2. Training and testing models link: https://drive.google.com/drive/u/0/folders/15ccLs7XeKw5jpI3kMJhXY8wUS6IyGUbM

3. Flask link:

   https://drive.google.com/drive/u/0/my-drive

4. Project demo link: https://drive.google.com/drive/u/0/folders/1u7R3LgLxfybmQgWXW2qDp0nt3iAX66sk

5. GitHub link: https://github.com/IBM-EPBL/IBM-Project-5849-1658817833