Internship Report

A Gesture-based Tool for Sterile Browsing of Radiology Images CNN and OpenCV

Dorsaf Tlili

4 erp-BI 6

TheSmartBridge summer internship:

15th june 2021 - 15th september 2021

Mentored by: Pradeepthi Duggaraju

pradeepthi

Table of content

General introduction……..…….…4

I- project context……..…...5

1-overview……...….5

2- purpose……..…….5

II- Literature survey……...…….....7

1-existing problem…….....7

2- proposed solution……8

III- theoretical analysis……..9

1- block diagram……....….9

2- software designing……....10

IV- Experimental investigation……...11

V- Flowchart……...13

1-WorkFlow…….…13

2- Sequence diagram……14

VI- Result…….….16

VII- Advantages and disadvantages……...….21

VIII- Application……22

General conclusion…….....23

Acronyms

AI: Artificial intelligence
CNN: Convolutional Neural Network
DL: Deep Learning

List of figures

Figure 1: Doctor holding MRI image 7

Figure 2: block diagram.....9

Figure 3: software design 10

Figure 4: mediapipe hand points.....11

Figure 5: workflow.....13

Figure 6: Sequence diagram.....14

Figure 7: Browse gesture.....17

Figure 8: Zoom in gesture.....18

Figure 9: Zoom out gesture.....18

Figure 10: Rotate clockwise gesture.....19

Figure 11: Rotate ccw gesture.....19

Figure 12: empty gesture…....20

General introduction:

Artificial Intelligence is the number one problem solving solution in the world. With the current advancement in technology, it is so much easier now to solve almost any problem with AI. From a simple algorithm to predict a variable from a dataset to analysing and predicting the economical state of a whole country to analyzing a person's body language in real time, with Artificial Intelligence you can do anything.

Deep Learning is a subdivision of Artificial Intelligence, mainly it imitates the works of the human brain in decision making based on the data processed. In this project we will be using Deep Learning to implement our solution for a gesture based tool to browse through MRI images and zoom or/and rotate them.

Using the power of deep learning and artificial intelligence we were able to implement a project that works with opencv and cnn to capture and analyze hand gestures in real time and perform actions on the MRI images depending on the gesture prediction.

The purpose of this project is to create a system which will allow the surgeons to browse and navigate through updated MRI images of the patient and decide about the next steps to follow in the surgery. Hand gestures are nonverbal and seem suitable for the precise control of the image navigation system.

Five gestures were used in this project to help perform 5 actions: scroll to browse through MRI images, zoomIn to zoom in on a specific image, zoomOut to zoom out on an image, rotateCW to rotate clockwise on an image and rotateCCW to rotate counterclockwise on an image.

In this report I will walk you through the problem, the solution, the purpose of this project as well as the technologies used and of course the future scope for this project.

I- Project context:

Introduction:

In this section of the report we will have a general overview of the project as well as the main purpose of it, providing the reader with a summary of what this project is all about.

1- Overview

This project was proposed by TheSmartBridge which is the company I did the summer internship for. TheSmartBridge is an edTech organization with a vision to bridge the gap between academia & industry. Their outcome-based experiential learning programs on emerging technologies (Internet of Things, Machine Learning, Data Science, Artificial Intelligence, Robotics) are building skilled entry - level engineers, for the corporate world. SmartBridge is on a mission to build technology communities in academia to encourage students towards innovation & entrepreneurship.

The main aim of this project is to provide a solution to browse and manipulate MRI images safely during surgery with no risk of infection.

This project is 100% written in python. The whole code was written, compiled and executed in PyCharm. There was a lot of use of the library cv2 (also known as OpenCV-python) which is a library of Python bindings designed to solve computer vision problems. There was also a lot of use of tensorflow and keras which are used for fast computing and model training.

The first step in this project was to collect the training images and for that I used a python script that uses video capture to capture a lot of images in matters of seconds.

After that the images were used to train the model and test it.

After having the model tested multiple times, the main function was written. The file main.py contains the code for real time gesture prediction and the actions for each gesture.

2- Purpose

The purpose of this project is to maintain the surgery environment sterile while the doctor browses through radiology images. This tool will allow the doctor to zoom, rotate or scroll

through MRI images without having to touch them, eventually reducing the risk of infection.

Hence it will provide a much safer environment for both the doctor and the patient.

A hand gesture system for medical images manipulation can be used during any surgery but we can consider a brain biopsy. This type of gesture based interface will avoid surgeon's focus shift and change of location while achieving, rapid intuitive reaction, and natural interaction. The major advantages of this approach is naturalness (surgeon use his hands), and rapid reaction. The build system responds to the surgeon's gesture commands in real-time.

The human–machine interface is of supreme importance because it is the means by which the surgeon controls medical information while staying sterile. We have to design a novel human–machine interface, for manipulating medical imaging through the use of hand gestures. The proposed human interface is an image recognition based system which tracks the surgeon's hand motions robustly in real time and prevents contact with external devices, so the surgeon remains sterile. The surgeon can easily browse, zoom or rotate MRI or X-rays by simply making the appropriate hand gesture.

Conclusion:

Now that we have a general overview of the project let's look at the problem and solution that was discussed when this project was proposed.

II- Literature survey:

Introduction

In this section we will have a greater understanding of the existing problem that caught our attention and the proposed solution for it

1- Existing problem

The current problem is that if the surgent, during surgery, needs to double check the MRI images to make a possible life changing decision, they risk infection because they have to touch the MRI image to thoroughly look at it.

Figure 1: Doctor holding MRI image

2- proposed solution

The solution was to use the power of AI to make the operation of investigating MRI images more safe.

The proposed solution is to create a tool that will capture the surgent’s hand gesture in real time and according to that gesture perform an action on the MRI images.

A vision-based system will capture the surgeon's hand gesture and predict it using a deep learning model, then perform an action on the MRI image to either browse, rotate clockwise, rotate counterclockwise, zoom in or zoom out.

Conclusion

This part of the report covered the details of the existing problem and the proposed solution for it

In the next section we will present the theoretical analysis containing the block diagram and the softwares needed.

III- Theoretical analysis:

Introduction

In this part, we will view and analyze the block diagram as well as discuss the softwares used.

1- Block diagram

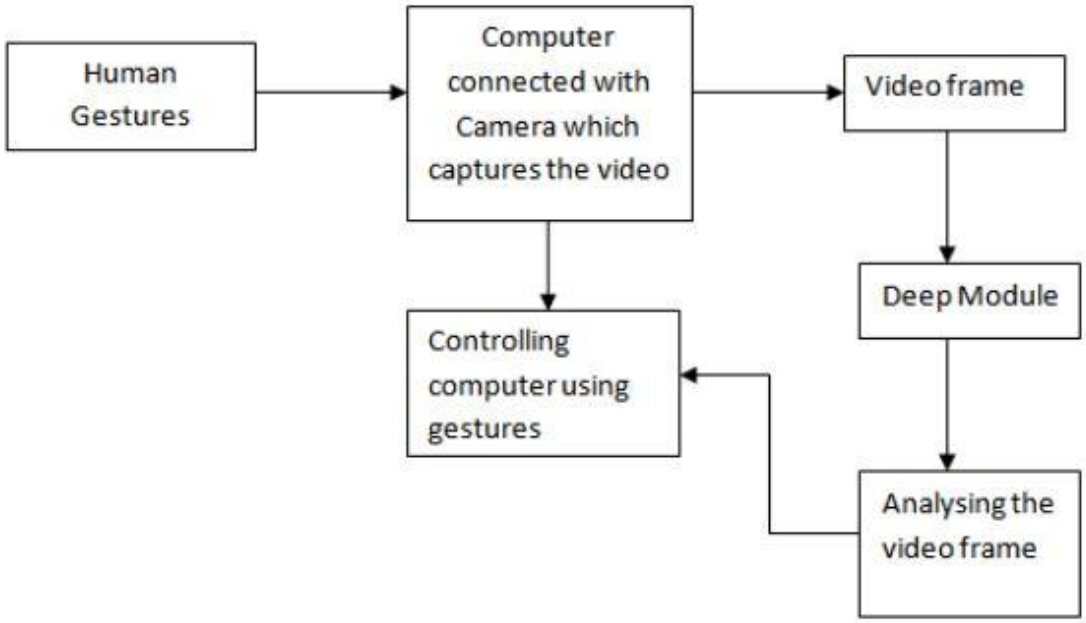


Figure 2: block diagram

2- Software designing

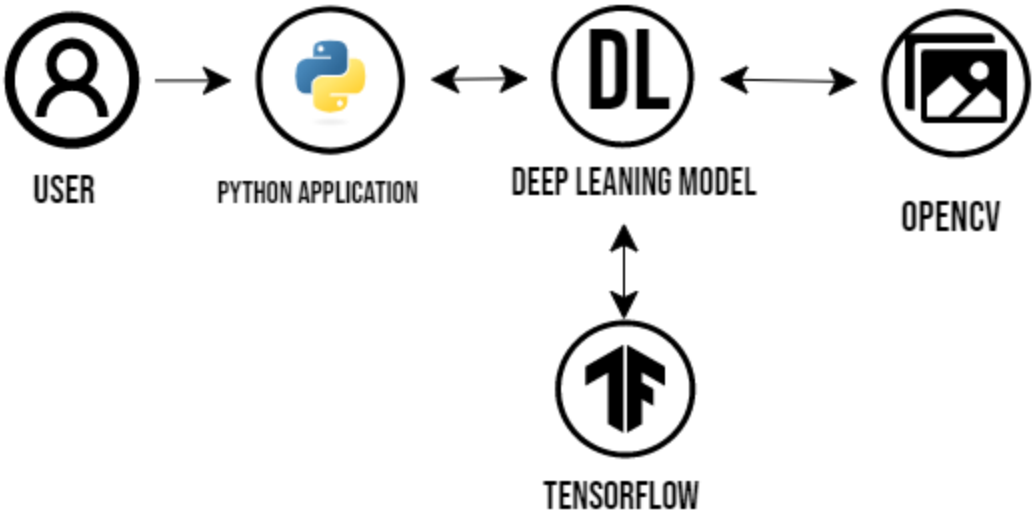


Figure 3: software design

There were so many technologies used to create the hand gesture tool.

Python: is an interpreted high-level general-purpose programming language.

Keras: is an open-source software library that provides a Python interface for artificial neural networks.

OpenCV: OpenCV is a library of programming functions mainly aimed at real-time computer vision.

Tensorflow: is a free and open-source software library for machine learning and artificial intelligence

Conclusion

This part of the report covered the theoretical analysis so in the next part we will talk about the experimental investigation done during the creation of this project.

IV- Experimental investigation:

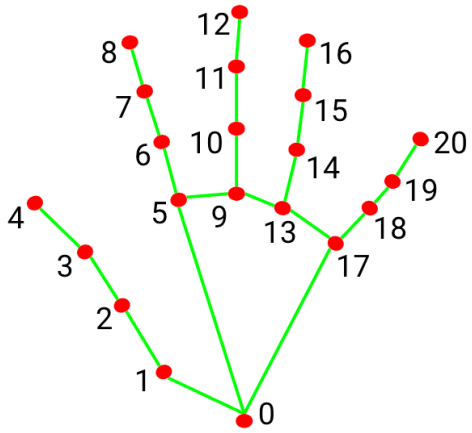
Introduction

In this section we will have a greater understanding of the research done during the realisation of this project.

The research:

The first step in working on a project is to do some research, then suggest some possible ways to establish the problem and finally decide on the best way that will meet what the proposed solution suggests.

During my research I found that the proposed solution can be solved using a hand tracking library called “mediapipe” created by google. mediapipe is a real time hand tracking tool that tracks 20 points on your hand, so with the use of this library we can detect the gestures based on the hand position using the points.



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

Figure 4: mediapipe hand points

Why didn’t I use this solution?

Using a pre-prepared library by Google is an easy way out. If I used this solution I wouldn’t need to create a deep learning model and that would make this project less beneficial for my learning experience.

Conclusion

In this part of the report we talked about all the research done during the realization of the project, specially the secondary solution that could be implemented to create the tool. In the next section we will view the flowchart of this project.

V- Flowchart:

Introduction

In this section we will view the WorkFlow of our solution as well as the Sequence Diagram. Reports.

1- WorkFlow

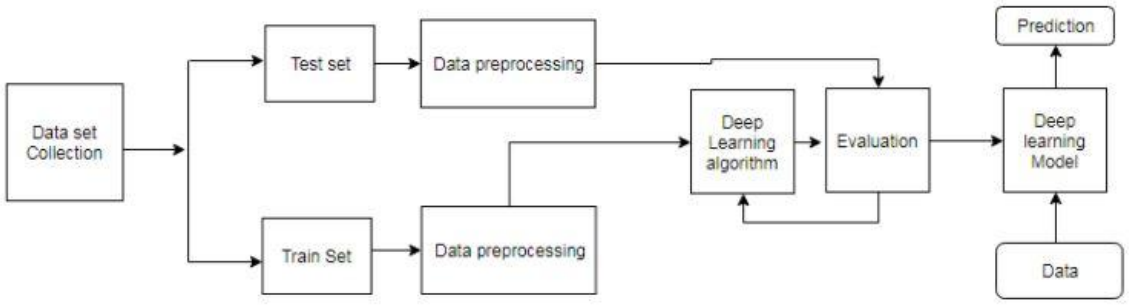


Figure 5: workflow

2- Sequence diagram

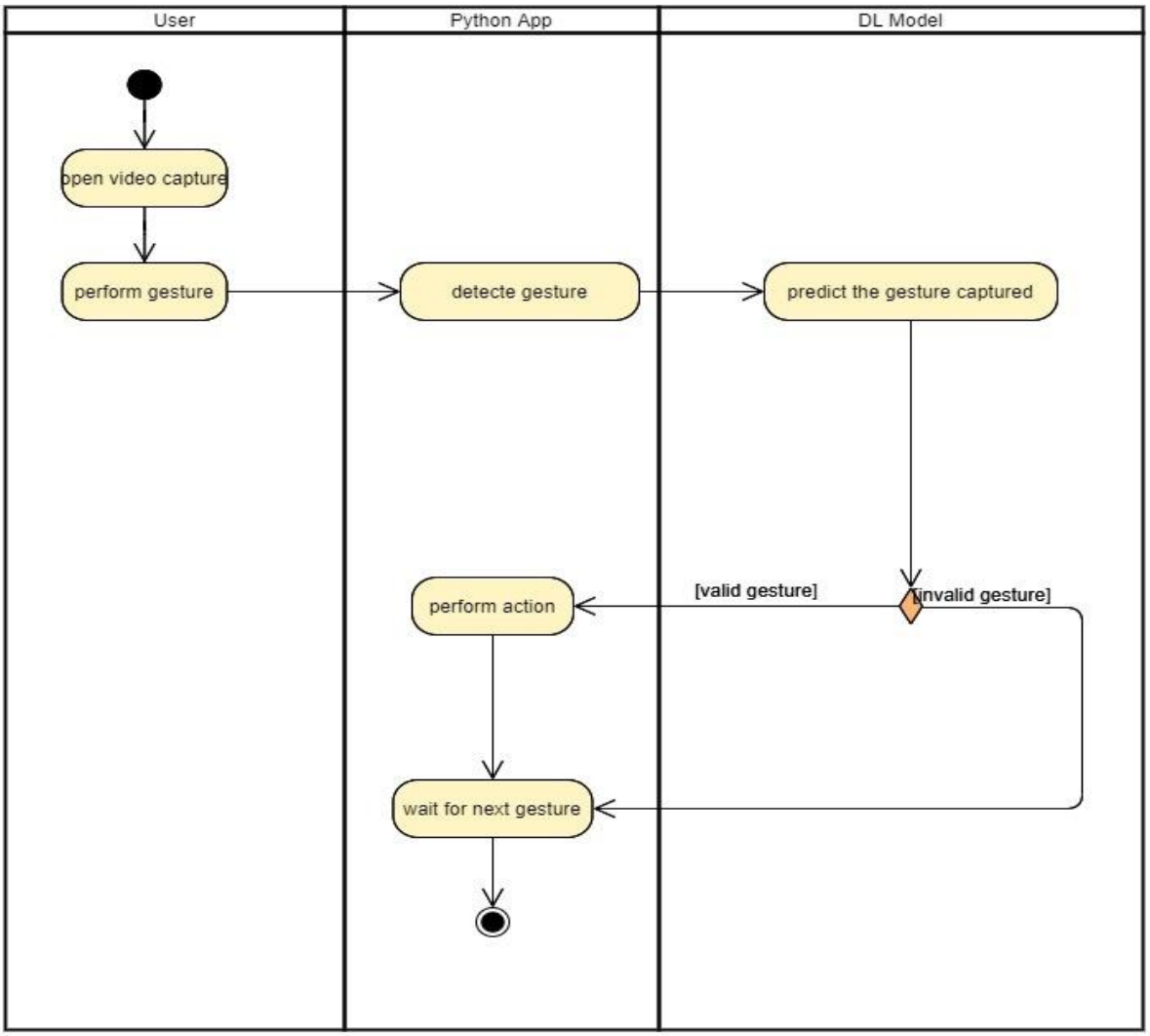


Figure 6: Sequence diagram

Conclusion

This part of the report covered the details of the current existing solution and why it is not enough to meet the needs of the company.

In the next section we will present our solution and specify the requirement and the objectives needed.

VI- Result:

Introduction

In this section we will discuss the result established by our proposed solution and dig deep into details of how it was created.

The project and the result:

This project was done step by step so let's see what each step was.

STEP 1: Collect the images:

I used a python script to collect images faster and for each and every gesture I collected 1300 images to make the prediction more accurate. I moved the position of my hand, without changing the gesture, around the image border a lot so that the prediction won't require that the gesture will be in the middle of the screen for it to be accurate. Also, I took the images with both my left and right hand and some of them while wearing a glove since the doctors wear gloves during surgery.

Here are the gestures:

Scroll: to browse through MRI images.

Figure 7: browse gesture

Zoom In: to zoom in on a specific image:

Figure 8: zoom in gesture

Zoom out: to zoom out of an image:

Figure 9: zoom out gesture

Rotate clockwise: to rotate clockwise on an image:

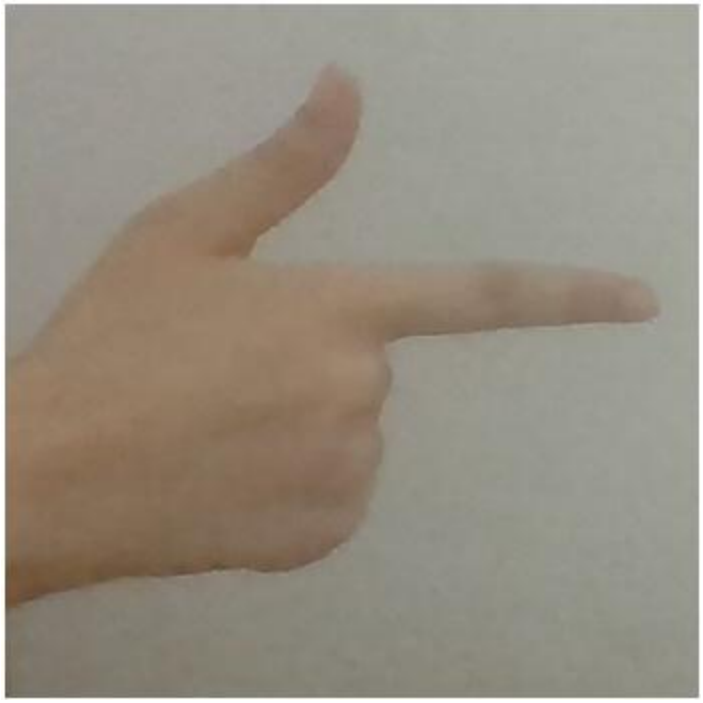


Figure 10: rotate clockwise gesture

Rotate counterclockwise: ro rotate on an image counterclockwise:

Figure 11: rotate ccw gesture

And i had to add an empty image set so it wont predict anything while there is no gesture:

Figure 12: empty gesture

STEP 2: Train the model:

After collecting the images we used them to train the model.

First we had to set the model parameters.

Then we compiled the model.

And lastly, fit the model and save it.

STEP 3: Test the model:

I tested the model using images I saved on my computer and unless the background is too messy the prediction is accurate each and every time.

STEP 4: predict hand gestures in real time:

With the use of the model and OpenCV we were able to capture real time gestures and predict them real fast.

STEP 5: perform the actions:

Depending on the gestures, actions were performed on the MRI images.

Conclusion

This part of the report covered the details of the solution created. In the next section we will present the advantages and disadvantages of our solution.

VII- Advantages and disadvantages:

Introduction

In this section we will discuss the advantages and disadvantages of our solution.

The advantages:

- Real time:

This project provides a real time tool to analyse hand gestures.

- Gesture based:

This tool is gesture based so it’s perfect for the existing problem.

- Fast prediction:

The prediction of the gesture is immediate.

The disadvantages:

- Video lag:

While performing the action the video frame lags a lot maybe due to the fact that my computer is not very well performing.

Conclusion

In this section we talked about the advantages and disadvantages, next part we will discuss the possible applications for this project.

VIII- Application:

A hand gesture system for medical images manipulation can be used during any surgery but we can consider a brain biopsy. This type of gesture based interface will avoid surgeon's focus shift and change of location while achieving, rapid intuitive reaction, and natural interaction. The major advantages of this approach is naturalness (surgeon uses his hands), and rapid reaction. The build system responds to the surgeon's gesture commands in real-time.

The human–machine interface is of supreme importance because it is the means by which the surgeon controls medical information while staying sterile. We have to design a novel human–machine interface, for manipulating medical imaging through the use of hand gestures.

The proposed human interface is an image recognition based system which tracks the surgeon's hand motions robustly in real time and prevents contact with external devices, so the surgeon remains sterile. The surgeon can easily browse, zoom or rotate MRI or X-rays by simply making the appropriate hand gesture.

Like it was mentioned in the "problem/solution" section, this project will be used for surgeons to browse and manipulate through MRI images safely with no risk of infecting the environment.

So the medical field is the number one beneficials of this gesture-based tool.

We can also expand the fields to regular users by making this tool an app that will help everyday people browse through images using hand gestures.

It can also be implemented in the educational field allowing the teachers to browse and manipulate images without having to touch a mouse or a screen.

IX- General conclusion:

The aim of this AI project is to solve the problem of safely browsing and manipulating MRI images during surgery. We used the power of deep learning and computer vision to create a tool that can predict hand gestures in real time and depending on the gesture perform an action.

The surgeon usually needs to avoid the rotation of the focus of attention and the change of his position during the operation because this causes additional mental effort and delay in the surgery. The purpose of this project is to create a system which will allow the surgeons to browse and navigate through updated MRI images of the patient and decide about the next steps to follow in the surgery.

Hand gestures are nonverbal and seem suitable for the precise control of the image navigation system. Furthermore, gestures do not require the use of pedals, and so they are convenient for the operating surgeon. vision-based gesture capture system interprets user's gestures in real-time to navigate through and manipulate an image and data visualization environment.

Dynamic navigation gestures are translated to commands based on their relative positions on the screen. The gesture system relies on tracking of the user's hand. we can implement the following states

1) BROWSE state: Navigation through MRI images while maintaining the image of interest in the centre of the video frame. (we can implement opening an particular image from desktop)
2) ZOOM state: Zooming an selected image
3) ROTATION state: Rotating the image of interest CW or CCW according to the rotation of the palm of the hand.

All the visualization work needed for the surgeons can be summarized into any of these states.

In this report, we covered every aspect of the project, from the initial problem to the result created.the realisation of this tool took effort and research as well as time. We hope that this project will get future development and enhancements.