Gesture recognition - Automating surgery process

-Vaidehi Som

Supervisor: Dr Vijay Kumar Pal

Overview

- Aim
- Challenges
- Project steps
- Literature survey
- Algorithm
- Data Sets Used and Methodology
- Output
- Other applications

Aim

Controlling computer screen using gestures.

Hand gesture recognition systems **in** operating rooms (ORs) are crucial **for** browsing and controlling computer-aided devices, which have been developed to decrease the risk of contamination **during surgical** procedures because doctors have to control screen during the surgery. Doctors need to use screen to check MRI scans, etc.

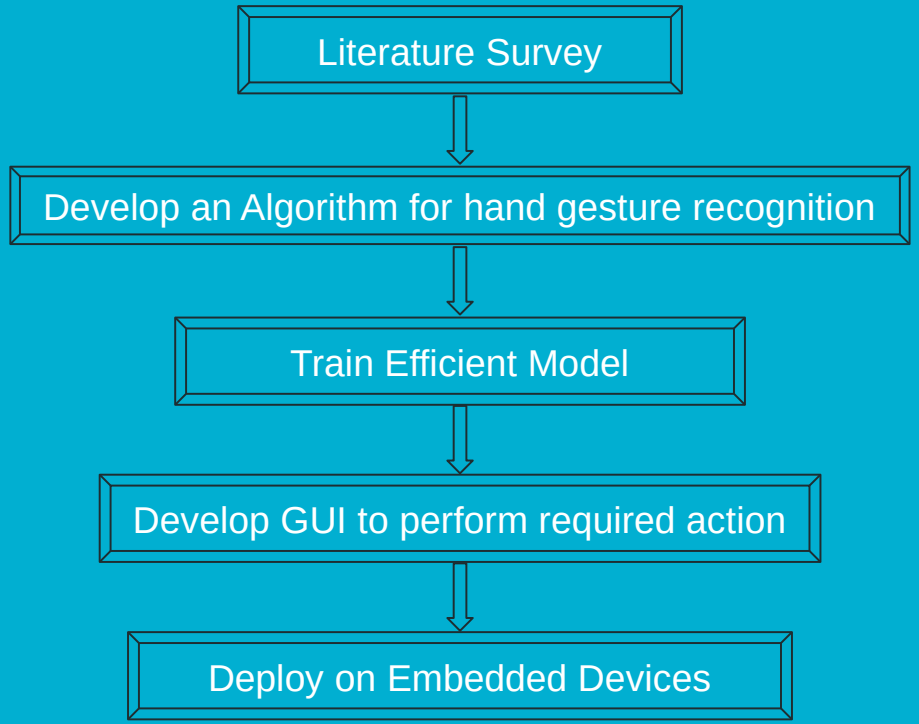
- Either take help from assistant sometimes communication problem and it take some time to adjust the screen according to the doctor.
- Touch the screen, then wash hands to prevent spread of infections: 10 min wastage per cycle

Efficient model that is generic and computationally less expensive.

Challenges

- Environment: Background illumination, color differentiation between hands and background
- Robustness
- Real time recognition
- User independent
- Fake signal
- High accuracy

Project Steps



Literature Survey

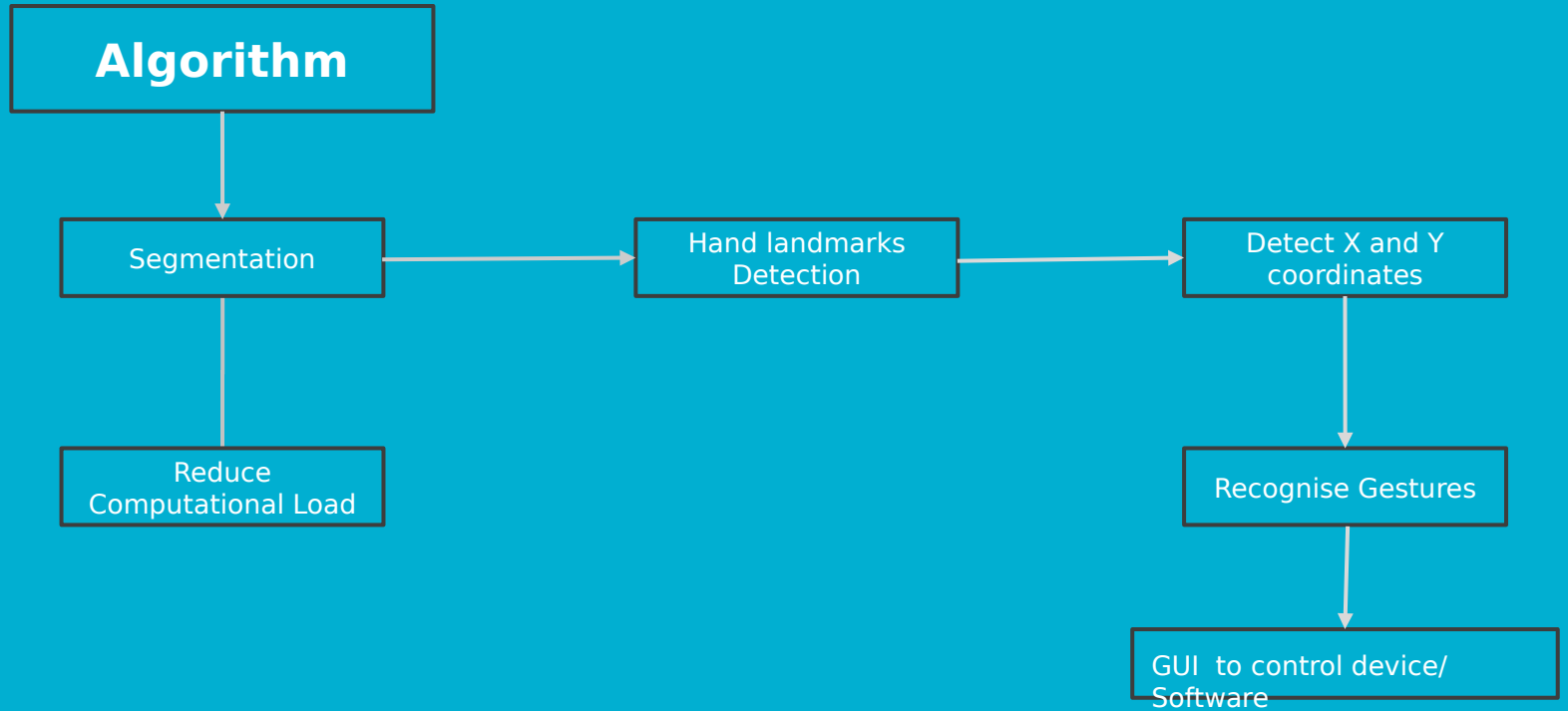
Available methods

- Skeleton based : Detecting points on hands. Defining gesture by angle parameters
- Vision based : Using camera, stereo cameras, etc.
- Sensor based : Using external sensors (LMC, Kinect, IMU, ultrasonic) to detect hands. Have to add on an external sensor.

Skeleton based gesture recognition stages

Data Acquisition	Image Preprocessing	Segmentation	Deep Learning Models	Identification
Standard camera, stereo camera, thermal camera, webcam	Remove noise (Gaussian filter, median filter) Downsize image to reduce processing time Histogram equalizer to enhance	Region of interest is segmented from rest of the image. Palm is segmented. Easier without fingers	VGG, ResNet, C3D, FasterRCNN Hand landmarks detection	Gesture recognition

Algorithm


Algorithm
Segmentation
Reduce Computational Load
Hand landmarks Detection
Detect X and Y coordinates
Recognise Gestures
GUI to control device/ Software

Dataset used

Large-Scale Multiview 3D Hand Pose Dataset by Robotics and Tridimensional Vision Research Group, 20500 images with anno

Methodology - Two stage architecture

1. Train palm detector instead of hand detector. Estimate bounding boxes using Non-maximum suppression algorithm and anchors
2. Cross entropy loss : precision 86%

Two ML models: Palm detector + hand landmarks

Palm detector returns Region of Interest (ROI)

Landmark models predict key points from ROI. These are to use ROI from previous frame

Our Contribution

1. Segmentation + keypoint detection
2. Gesture detection using these key points
3. Using these gestures to control computer screen, mouse and keyboard.

Gestures we can detect:

4. Zoom in
5. Zoom out
6. Control Mouse
7. Double click
8. Single click

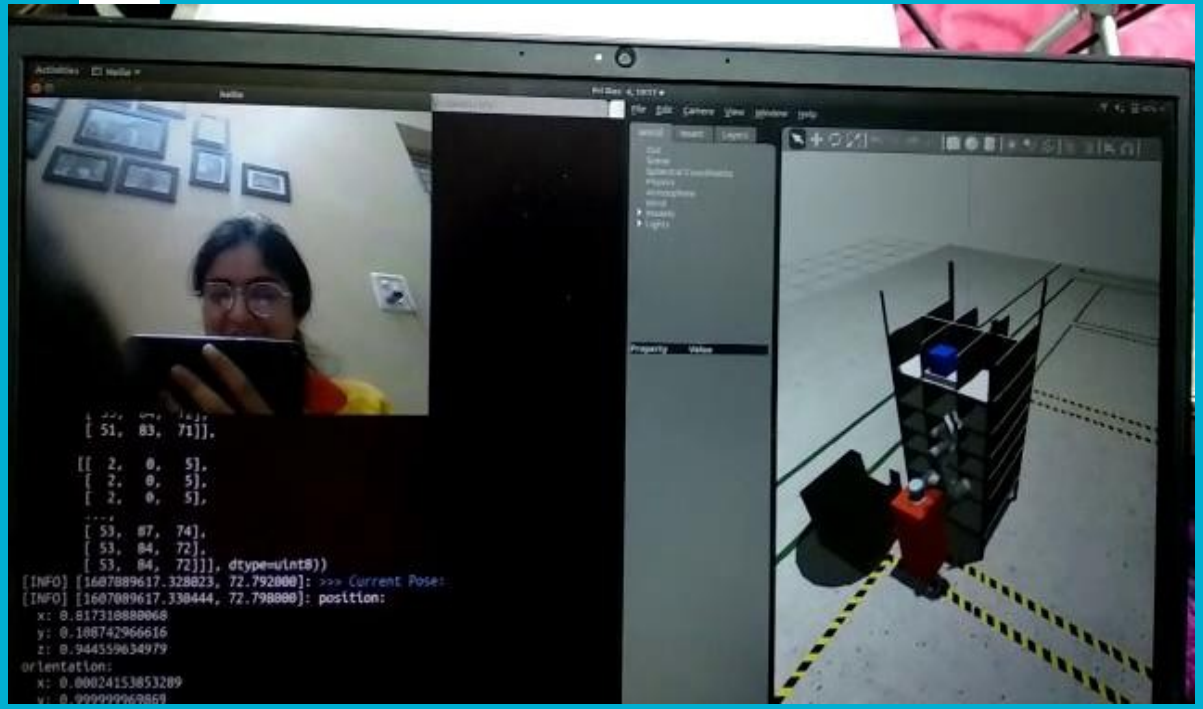
Example Output

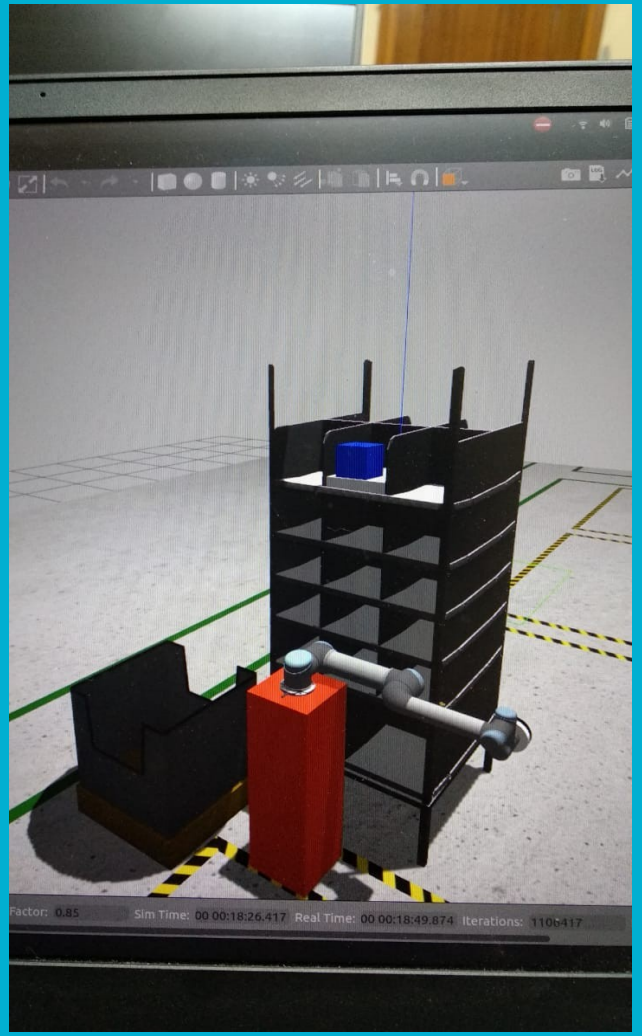
These are the types of outputs we are aiming to obtain:

- http://tedcas.es/tedcube/
- https://youtu.be/Blrdttwc8js
- https://www.youtube.com/watch?v=-wypzKpWeFk

Our output

Our output


[51, 83, 71]],
[[2, 0, 5],
[2, 0, 5],
[2, 0, 5],
...,
[53, 87, 74],
[53, 84, 72],
[53, 84, 72]]], dtype=uint8))
[INFO] [1607089617.328823, 72.792000]: >>> Current Pose:
[INFO] [1607089617.330444, 72.798000]: position:
x: 0.817318880068
y: 0.188742966616
z: 0.944559634979
orientation:
x: 0.00024153853289


Factor: 0.85
Sim Time: 00 00:18:26.417
Real Time: 00 00:18:49.874
Iterations: 110417

Our output

Other Applications

- Controlling robots during surgery
- Household robots
- Wheelchair control
- Workspace robotics control - Controlling industrial robots
- Augmented reality
- Gaming
- Sign language interpretation
- Similar other touchless technologies

Future Scope

- Incorporation of more gestures
- Using better algorithm to increase robustness
- Decrease sensitivity of gestures detected and operations performed
- Incorporate for scenarios involving more than one hand