International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 12 Dec. 2022, pp: 339-346 www.ijaem.net ISSN: 2395-5252

A Survey on Virtual Air Sketching Using OpenCV

Ranjitha N, Rakshitha V, Prakshi Singh, Aditi B Prahalad, Smt.Akshitha Katkeri

Dept of CSE,BNM Institute of Technology

Date of Submission: 08-12-2022 Date of Acceptance: 16-12-2022

ABSTRACT

In today's world due to the pandemic situation online learning gained immense importance and is widely in use. But there are many barriers for Online learning as well. One of the most important barriers is concentrating on the line that is taught. It is also necessary for the Professors to mark the important words or lines. Using Virtual writing and controlling system contributes extremely to the advancement of an automation process and can improve the interface between man and machine in numerous applications. Using OpenCV and Python we can solve this problem but implementing Virtual Air Sketching Where it allows the user to draw by moving the finger, which makes it both awesome and interesting to draw simple thing.

I. INTRODUCTION:

Writing in air has been one of the most fascinating and enthralling exploration areas in field of image processing and pattern recognition in the recent times. It contributes immensely to the advancement of an automation process and can ameliorate the interface between man and machine in multitudinous operations. Several exploration works have been fastening on new ways and styles that would reduce the processing time while furnishing advanced recognition accuracy.

Numerous studies have concentrated on new strategies and approaches that would speed up recognition while lowering processing time. This jamboard-like Python application uses the OpenCV library, which is based on real-time webcam data, to track an object-of-interest (in this case, a human palm or finger), and it enables the user to draw by moving the finger, which makes it awesome and interesting to draw simple things.

The task of object tracking is regarded as crucial in the field of computer vision. The development of faster computers, the accessibility of inexpensive, high-quality videotapes, and the requirements for automated video analysis have all increased the demand for object tracking methods. The three main steps of a video analysis procedure are typically the detection of the object, tracking its movement from frame to frame, and preliminary analysis of the object's nature.

Virtual Air Sketching is a hands-free digital drawing canvas that utilizes a Raspberry Pi, a PiCamera, and OpenCV to recognize and map hand gestures onto a screen. The user's "brush" can be modified in color and size by using built-in buttons. The direction of the brush is controlled completely using OpenCV software and modified to map the pointer finger onto the screen using Pygame following a calibration screen to measure and record the color of the user's hand. The idea for Virtual was a result of our interest in digital drawing.

Four different types of issues are considered when tracking an object: picking an appropriate method of object representation, choosing the features needed for tracking, finding the object, and tracking the object. The algorithms related to object tracking are an essential component of many applications today, including automatic surveillance, vehicle navigation, and video indexing. The survey identifies this gap and is focused on creating a motion-to-text converter application that may one day be used as computer software for devices for air sketching. To track the finger's movement, it uses computer vision. It might be a way for deaf people to communicate. It is a potent communication technique that does away with paper-based writing.

II. MOTIVATION:

The initial driving force was the requirement for a student study environment free of dust in the classroom.Thereare numerous options, including touch screens, but what about the schools that cannot afford to purchase such enormous large screens and use them for classroom instruction. We, therefore, pondered why a finger couldn't be tracked, albeit superficially and without deep learning. Despite the use of tactile graphics and audio guides, blind and visually impaired people still face challenges to experience and understand visual artworks independently at art exhibitions Art museums and other art places are increasingly exploring the use of interactive guides to make their collections more accessible.As a result, OpenCV saved the day for these computer vision projects.

III. PROBLEM DEFINITION:

The current system only functions with your fingers, not crayons or paints. We concentrate on the difficult task of identifying and separating a finger from an RGB image without a depth sensor. The lack of a top and movement under the pen are additional issues. One RGB camera is used by the system, which you can replace. It is impossible to discover the bottom up, and a pen cannot be followed up. The result is an abstract, model-unseen image because every finger path has been drawn. It takes a lot of code care to change the position of the process from one region to another using real-time hand touch. To properly control his plan, the user should also be familiar with numerous movements.

The survey focuses on finding solutions to some of the most pressing social problems. The problems that hearing-impaired people face in daily life are numerous. While listening is something that most people take for granted, sign language is not always used when communicating with someone who is disabled. Second, Paper waste is not unusual. Wasted paper includes paper used for writing, drawing, etc. Paper makes up 25% of landfills, 50% of commercial waste, 93% of sources, and the list goes on. These issues can easily be resolved through on-air writing.It will help the hearing impaired communicate. Your online text may be spoken back to you or displayed in augmented reality.On the air, one can write quickly and work.

IV. LITERATURE SURVEY:

Detailsabouttheliteraturereviewthat weobserved on the topic Virtual Air Sketching are viewed in this section.Thepapers'contentsare listed below.

1.1 An Economical Air Writing System Converting Finger Movements To Text Using Web Camera: [1]

The system is being developed using fingertip detection and finger movement techniques. Fingertip is first detected using Python, OpenCV, and CNN techniques, and then its trajectory is tracked and shown on the screen. The tracking of the hands and the tips of the fingers is done using the MediaPipe package. The movements of the LED-fitted fingers are recorded using a web camera, and the patterns are recognized using characters from the database. The characters are printed onto the screen after identification. As it is simpler to track down the red color, a red-colored LED pointed light source is attached to the user's finger to speed up finger movement tracking. The precise character is found and displayed on the screen using the optical character recognition (OCR) technique. This OCR method makes use of a pre-built database that is filled with the entire English alphabet, from A to Z. Thisdatabase is used to identify the English alphabet and compare it to the cropped black-and-white image. A text editor, such as Notepad, displays the English character. The user's characters are processed one at a time by this system, and the process is looped until the user has finished typing. Software called MATLAB is used to program all intended operations and produce the desired result. By dividing the problem solution into various modules, the aforementioned methodology is put into practice. The method for automatic video indexing andvideo search in large lecture video archives ispresented in this research. They have used key-framedetectionandautomaticvideosegmentation to provide avisual roadmapfornavigatingthevideomaterial.Videoisconvertedinto textual data using OCR (Optical CharacterRecognition) on different frames of video andASR (Automatic Speech Recognition) on audiotracks.

Automated Lecture video indexing: In order toextract each individual slide frame with its owntemporalscopeandberegardedavideosegment,they firstdetectthe slide transitions from thevisual screen. Then textual metadata is extractedfromslide frames usingvideoOCRanalysis.

The first step in the implementation phase is to capture a video and then separate it into sequences. Red-colored objects are tracked from this series of images, which includes 100 images. In this instance, it is assumed that the only red-colored objects in the environment are those illuminated by the tracking LED light mounted on the finger. The images are converted into binary images after the red-colored objects have been located, where the red color is replaced by white and the rest of the environment is turned black. After creating an image database to compare with the input image, the matched alphabet is then displayed on the screen.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 12 Dec. 2022, pp: 339-346 www.ijaem.net ISSN: 2395-5252

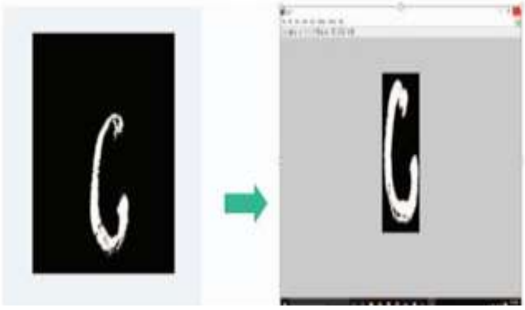

Figure 4.1.1 : Cropping of the Region of Interest [1]

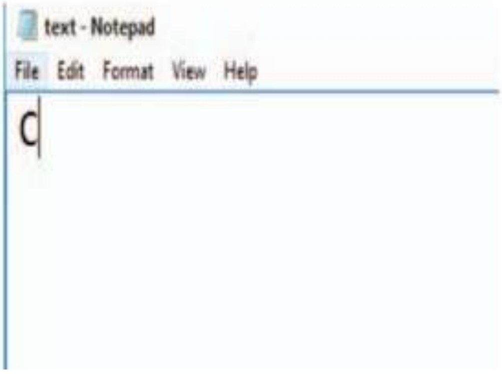

Figure 4.1.2 : Displaying of the identified character in Text Editor [1]

4.2 Air Canvas Application Using OpenCV and Numpy python : [2]

This system needs a dataset for the fingertip detection model as it is the primary purpose to record the motion i.e. the air character.

A. Fingertip detection model: Air writing can be merely achieved using a stylus or air pens that have a unique color. The system, though, makes use of the fingertip. We think it should be possible for people to write in the air without having to carry around a stylus. They have used Deep Learning algorithms to detect fingertips in every frame, generating a list of coordinates.

B. Techniques of Fingertip Recognition Dataset Creation:

a. Video to Images: In the first method, two-second videos of a person's hand motion in various settings were recorded. Then, 30 different images were created from these videos. Then a dataset of 2000 images was extracted. This dataset was labelled manually using Labelling. The best model trained on this dataset yielded an accuracy of 99%. However, since the generated 30 images were from the same video and the same environment, the dataset was monotonous. As a result, the model performed poorly for discrete backgrounds like those in the dataset.

b. Take Photos with Different Backgrounds: To address the issue with the previous This time, we were aware that we needed some gestures to control the system. So, we collected the four distinct hand poses. The goal was to efficiently enable the model to identify the tips of all four fingers. This would enable the user to operate the system by displaying a certain number of fingers. He or she could promptly write by showing one index finger, convert this writing motion to e-text by offering two fingers, add space by showing three fingers, hit backspace by showing five fingers, etc. This dataset consisted of 1800 images. The previously trained model was made to automatically label this dataset using a script. After that, we changed the labels on the images and added a new model. A 94% accuracy was achieved. This model worked well in different backgrounds.

C. Fingertip Recognition Model Training: The dataset is split into train and dev sets (85%–15%) after it has been prepared and labelled. We used Single Shot Detector (SSD) and Faster RCNN pre-trained models to train, but Faster RCNN was much better in terms of accuracy as compared to SSD. SSDs combine two standard object detection modules – one which proposes regions and the other that classifies them. This speeds up the performance as objects are detected in a single shot. It is commonly used for real-time object detection. Faster RCNN uses an output feature map from Fast RCNN to compute region proposals. They are submitted to a Region of Interest pooling layer after being evaluated by a Region Proposal Network. Finally, two fully connected layers for classification and bounding box regression are given in the result. To identify the fingertip in the image, we tuned the Faster RCNN's final fully connected layer.

Figure 4.2.1 :Word written in air traced on a black image [2]

4.3 Virtual Air Canvas application using opencv and numpy in python : [3]

Various models were looked into to determine which one to be used for detection. NN multilayer (Neural Perceptron) a network with a line divided into sections a real-time video capture function processing. This activation function reduces means NN complex time. They use NN in two stages:

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 12 Dec. 2022, pp: 339-346 www.ijaem.net ISSN: 2395-5252

color recognition and steps to go check hand shape. Their results show that the method proposed well works with the acquisition rate of 99.2% .

Jordan's Recurrent Neural Network (JRNN) can be used for hand touch recognition. Their program compared 5 and 9 different hand positions through repetitive sequence still images.

He then took the recording asre-installation begins to separate the shape of the hands. JRNN gets input touch after a temporary behaviour of sequence of positions has been detected. In addition, they created a new way of training, the proposed method shows a 99.0% accuracy.

The main objective is to create a space where user can draw freely in air by moving their hands in front of the camera screen.

Fingertip detection:

The focus is on detecting the whole hand first then the segmentation of the region is done. Region segmentation is a two step approach that has skin segmentation and background subtraction. Faster R-CNN is used for background subtraction. Centre if gravity is determined as it helps in detecting proper hand gesture. The proposed system aims to use two algorithms for centroid calculation and then take the average value of both as the final result. Distance transformation is the algorithm used and the pixel with the highest intensity is the centre of gravity.

Fingertip Tracking:

After detection of hand gesture and centre of gravity, we detect fingertip movement. Faster R-CNN handheld detector is intensive and the frames produced are below real-time performance. Thus we are aiming to use KCF tracing algorithm. This algorithm converts the detected fingertip into HSV colour space. After detecting the mask, the system does morphological operations to remove impurity from masked image. After contours are detected the line is drawn. Then python deque is used to memorize the positions of the outline in each subsequent frames, and we will use these accumulated points to create a line using OpenCV's drawing capabilities.

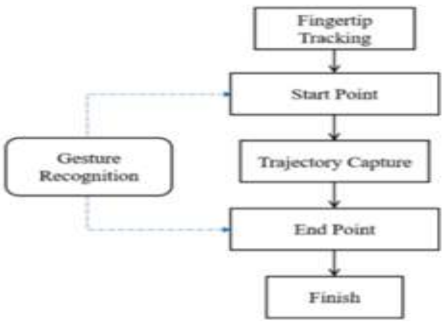

Figure 4.3.1 : Fingertip Tracking Model [3]

4.4 Air Canvas Using OpenCV,Mediapipe:[3]

To overcome the limitations of using data with different background differences or the skin color difference, they have made use of mediapipe. In this hand tracking is done using mediapipe that first detects hand landmarks and then obtains positions relative to it.

Using opencv, mediapipe the efficiency of existing system are improved. The stages of proposed system are:

Run or execute the code:

After all libraries are installed execute the code, that turns the camera on automatically and opencv frame with various buttons displaying various shapes, colors, size, save, clear, erase etc. is visible.

2. Webcam Starts:

Webcam starts recording and converts the video each frame and sends the frame to hand tracker class to track or detect the positions of finger.

3. Detects Hand Landmarks

Each frame received is compared with mediapipe hand landmarks i.e, the positions of finger are found using getPositions() and which finger is opened using getUpFingers() functions of handtracker class. These 2 functions are part of the Handtracker class.

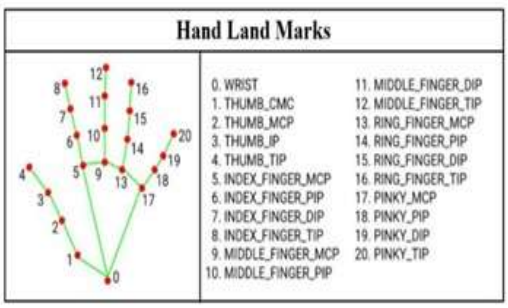

Figure 4.4.1: Mediapipe Hand Landmarks [4]

4. Perform actions according to button

Different buttons are chosen with the help of index finger hovering over those buttons. Same way different shapes are chosen. Each button has its own functionality to perform. If shapes or colors or size are chosen then respective functionalities are performed on the canvas . The shapes are resized according to the distance between the thumb and the index finger. If pdf is chosen then it gets opened in ms edge and editing is done.

5. Depict on canvas

Shapes like rectangle, circle, ellipse, arrow head line, line, free style are depicted on canvas. The content on canvas can be edited, as well as saving of work. Shapes are drawn by considering the distance between the thumb and index finger as the diagonal length for rectangle, radius for circle, length for line and arrowhead line

6. Annotate PDF

If pdf option is selected then a Tkinter window opens asking to choose the pdf needed to open. After opening pdf, select draw and use finger to edit or draw over the pdf at required areas.

7. Exit

By hovering index finger over exit closes the frame.

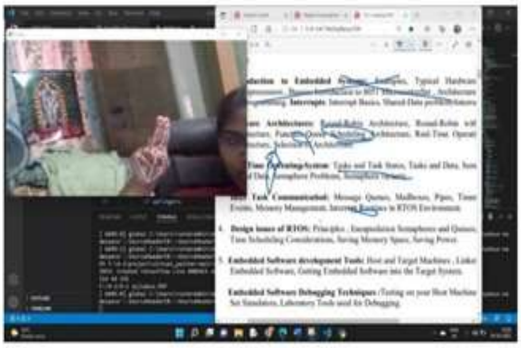

Figure 4.4.2: Annotated PDF [4]

4.5 Building a Air Canvas using Numpy and Opencv in Python [5]

The proposed system needs a datatset for fingertip detection model. The main purpose of this model is to record motion.

A.Fingertip detection model:

For writing in air we need stylus or one of many colorairpens in the market. But this system relies on finger for it to operate. Our main focus is to write in air without having to carry out a pen around with them. Every frame has been analysed using Deep Learning techniques to generate a list of coordinates for each fingertip

B.Techniques of Fingertip Recognition Dataset Creation:

1. Video to Images: Hand motion films were recorded for two seconds and then transferred to a variety of diverse contexts. This footage was sliced up into 30 individual pictures, so that more than 2000 photos were taken. Data was manually annotated in this dataset. The most accurate model developed using this dataset has a precision of 99.5 percent. Dataset monotony resulted from using a same surroundings to produce 30 different photos. It follows that for non-continuous environments like those in the sample, the model failed miserably.

2. Take Pictures in Distinct Backgrounds: We built a new dataset to solve the problem of the prior method's lack of variety. This time, we were well aware of the fact that gestures were required in order to operate the system. To that end, we gathered the four hand gestures. The goal was to create a model that could recognize the tips of all four fingers with equal efficiency.

One index finger is all that's required to start writing, and the user can now - quickly write by showing one index finger,convert this writing motion to e-text using two fingers, etc.

The photos in this collection totalled 1800. An auto-labelled dataset was created using a software that has previously been trained on this dataset Once the photos had been properly renamed, a new model was introduced. At least 94% of the time, it worked. Different backdrops proved to be suitable for the model.

C. Fingertip Recognition Model Training:

Divided into train and dev sets after completion of the dataset preparation and labeling process (85% -15%). Our dataset was trained using the Single Shot Detector (SSD) and the Faster RCNN pre-trained models. In comparison to SSD, the accuracy of the faster RCNN model was much higher. Object detection in SSDs consists of two basic modules: one that suggests areas and another that classifies them. Detecting items in a single shot reduces processing time. It is often used for real-time object detections. To calculate region suggestions, Faster RCNN utilizes the output feature map from Fast RCNN. They are reviewed by a Region Proposal Network and forwarded to a Region of Interest pooling layer. The categorization and bounding box regression results are eventually shown in two completely linked layers . We

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 12 Dec. 2022, pp: 339-346 www.ijaem.net ISSN: 2395-5252

modified the final completely linked layer of Faster RCNN to detect the fingertip in the picture.

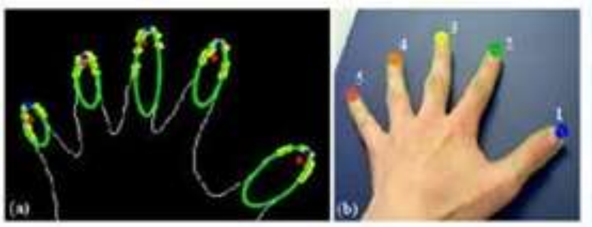

Figure 4.5.1: Fingertip Recognition Model [5]

Figure 4.5.2:Final Output [5]

V. COMPARATIVE ANALYSIS OF PAPERS

Title	Authors	Methodology	Limitations	Publication,Year
An economical air writing system converting finger movements to text using web camera	PavithraRamasamy ,Prabhu. G, Dr. R. Srinivasan	Python, OpenCV, CNN techniques,MediaPipe ,Red LED light source,OCR, MATLAB	• Single color detection model where it is assumed that there is no red colour in the background.	Fifth International Conference On Recent Trends In Information Technology, 2016
Air Canvas Application Using OpenCV and Numpy python	Prof. S.U. Saoji ,NishthaDua, Akash Kumar Choudhary, Bharat Phogat	Faster RCNN,SSD, Deep Learning algorithms to detect fingertip,dataset for training	• Detection model is trained using a dataset with similar background. • Hence it does not work for diverse background.	International Research Journal of Engineering and Technology,2021
Virtual Air Canvas application using opencv and numpy in python	AsstProf.Jahnavi S,K Sai Sumanth Reddy, Abhishek R, AbhinandanHeggde, Lakshmi Prashanth Reddy	KCF tracting algorithm, MSER, faster RCNN	• Faster RCNN and KCF tracting algorithm do not provide efficient accuracy and speed,	International Journal of Creative Research Thoughts,2022

Building a Air Canvas using Numpy and Opencv in Python	D. Vijendra Kumar, G.Vijaya Raj Siddarth, R. Venkata Satya Sravani, I.VishnuVardhan Reddy, Y. Lalitha Sri Naga DurgaVyshnavi	Deep Learning Algorithm,FasterRCNN,SSD,Manually annotated dataset	• This system is color sensitive cuch that presence of any red color in the background may lead to wrong interpretation.	International Journal for Modern Trends in Science and Technology,2022
Air Canvas Using OpenCV,Mediapipe	Dr. B. Esther Sunanda, M. Bhargavi, M. TulasiSree, M.R.S. Ananya, N. Kavya	Mediapipe,Tkinter, Hand Tracking, Pdf Annotator	• Lack of Documentation of Mediapipe poses a disadvantage for complete functional usage	International Journal of Research of Modernization in Engineering Technology and Science,2022

VI. CONCLUSION

As we develop this system, we offer a very straightforward and cost-effective solution for one of the burgeoning areas of research in the field of gesture recognition: translating finger movements to text. This system is designed in such a way that it can recognize the English alphabet drawn in the air and translate it into text using just a basic Webcam. This system's implementation is carried out using Matlab software, which enables straightforward, understandable, and effective coding structures. Any computer that has a Web Camera and the necessary configurations can have the system installed. This system can undoubtedly be seen as the starting point for the more recent invention of translating finger movement to text and can act as a benchmark for future improvements in the same field.

Future work could focus on increasing the system's precision in identifying the alphabets drawn, particularly those with intricate geometrical structures. To make the system offer better real-time Human-Computer interaction, the speed of alphabet recognition can also be increased.

REFERENCES

[1] PavithraRamasamy, Prabhu. G, Dr. R. Srinivasan, " An Economical Air Writing System Converting Finger Movements To Text Using Web Camera" in Fifth International Conference on Recent Trends in Information Technology, 978-1-4673-9802-2, 2016.

[2] Prof. S.U. Saoji, NishthaDua, Akash Kumar Choudhary, Bharat Phogat, " Air canvas application using OpenCV and numpy in python" in IRJET, (Deemed to be University) College of Engineering, Pune, Volume: 08 Issue: 08, e-ISSN: 2395-0056, p-ISSN: 2395-0072, Aug 2021.

[3] AsstProf.Jahnavi S , K Sai Sumanth Reddy ,Abhishek R , AbhinandanHeggde , Lakshmi Prashanth Reddy, "Virtual Air Canvas Application using OpenCV and Numpy in Python" in IJARIIE, Dayananda Sagar Academy of Technology and Management,Bangalore,Karnataka,INDIA, Vol-8 Issue-4 , ISSN(O)-2395-4396, 2022.

[4] Dr. B. Esther Sunanda, M. Bhargavi, M. TulasiSree, M.R.S. Ananya, N. Kavya, " Air canvas using OpenCV, Mediapipe " in

IRJETS, College of Engineering for Women, Andhra Pradesh, India, Volume:04/Issue:05, e-ISSN: 2582-5208, May-2022.

[5] D. Vijendra Kumar, G.Vijaya Raj Siddarth, R. Venkata Satya Sravani, I.VishnuVardhan Reddy, Y. Lalitha Sri Naga DurgaVyshnavi, “Building a Air Canvas using Numpy and Opencv in Python” in IJMTST, Godavari Institute of Engineering and Technology(A), JNTUK, Kakinada, ISSN: 2455-3778, DOI: https://doi.org/10.46501/IJMTST08S0528, 2022

[6] SumedhBansode , SaurabhVarkhad , Santosh Dhaigude , ShivamWaghmare , Prof. ShubhangiSuryawanshi, “ Computer Vision Based Virtual Sketch Using Detection” in IJRASET, DR D Y Patil institute of Technology, Pimpri, Pune, Volume 10 Issue V, ISSN: 2321-9653, IC Value: 45.98, SJ Impact Factor: 7.538, May 2022.

[7] Sayli More, PrachiMhatre, ShrutiPakhare, SurekhaKhot, “Air Canvas: Draw in Air” in IRJET, A.C Patil College of Engineering, Navi Mumbai, India, Volume: 09 Issue: 02, e-ISSN: 2395-0056, p-ISSN: 2395-0072, Feb 2022.

[8] K. Kim et al., “Real Time Foreground-Background Segmentation Using Code Book Model,” Real-Time Imaging, vol. 11, no.3, 2005, pp. 172–185.

[9] Y. Huang, X. Liu, X. Zhang, and L. Jin, "A Pointing Gesture Based Egocentric Interaction System: Dataset, Approach, and Application," 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Las Vegas, NV, pp. 370-377, 2016

[10] HemalathaVadlamudi ,Evaluation of Object Tracking System using Open-CV In Python International Journal of Engineering Research& Technology (IJERT) Vol. 9 Issue 09, September-2020.