iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** XII **Month of publication:** December 2023

DOI: https://doi.org/10.22214/ijraset.2023.57456

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XII Dec 2023- Available at www.ijraset.com

AirPen: Real-Time Gesture-based Interactive Drawing using Computer Vision

Vikram Kulkarni[1], Shubham Kulkarni[2], Kashish Bhatia[3]
Assistant Professor, Dept. of Information Technology, MPSTME, Maharashtra, India
B.Tech in Information Technology, MPSTME, Mumbai

Abstract: The project provides someone who uses a system that allows him to draw on video without switching to a separate whiteboard program while he or she is talking. The sphere of education and training has advanced greatly throughout this epidemic era therefore, this project would be quite valuable for speakers who wish to explain their subject by writing or drawing while speaking. The target is either his fingers or a pen, and its movement is traced to create the desired image.

I. INTRODUCTION

AirPen aims to provide a platform for the person to sketch while conversing on video without transferring to an external white board software which is the primary goal. - The field of learning and instruction has evolved significantly as we are very much aware of during this whole epidemic period, and thus this project which we as a team of 3 would prove extremely valuable to the speakers in expressing his subject by writing or drawing while speaking. - As a target which is to obtain the necessity, the user's fingers or a pen/scale are selected, and their activity is traced along to draw as needed.

Most Joysticks, keyboards, trackballs, and light pens are all examples of conventional input ads we are very much aware of, and they are just not very good at simulating natural hand gestures like drawing and sketching. Conventional input methods fall short particularly in comparing to the smoothness of the human hand when drawing or sketching which is an integral part of the exercise. Designing a very simple user experience which is unique and that allows users to submit the drawing that defines the product's appearance is a tough design problem to solve but somehow we will tackle it. Due in large part to the fingertip's relatively small size which makes it difficult and also accurate and strong detection needed for further appearance of the target object and also tracking those has proven very difficult despite recent advancements in object detection and tracking. Furthermore, the lack of a widely accepted limiting criterion makes it difficult to start or finish midair finger writing.

Using the Faster R-CNN framework for precise hand detection, hand segmentation, and other important properties which counts the number of lifted fingers based on geometrical characteristics of the hand, we as a group of three offer a new writing hand pose detection technique for the initiation of air-writing which will be very unique. We as a group came with an idea of a novel signature function and also other attributes which also include distance-weighted curvature entropy, as a means of reliable fingertip detection which is unique. We have come up with a cutoff criterion based on the velocity of the writer's fingertips to indicate when the air-writing motion is complete.

II. METHODOLOGY

We offer an alternative method which can be very unique for deciphering air-written content in movies captured with a regular laptop camera or a web-cam to address these concerns. Hand shape distortion, fingertip motion blur, a cluttered background, and varying illumination all contribute to making this a challenging task. Our results can contribute the following to solving those issues: Using the Faster R-CNN framework for reliable hand detection, hand segmentation, and eventually counting the number of uplifted fingers which is based on geometrical characteristics of the hand, we have come up with a new writing hand pose detection approach for the initiation of air-writing which is intrusive. We propose a novel signature function, distance-weighted curvature entropy, as a manner of effective fingertip detection in an effective manner. We offer a cutoff all the points that leads towards the end of the air-writing motion based on the velocity of the reporter's fingers.

A. Proposed System

By employing the device's built-in webcams, Media Pipe gives machine learning-based solutions for augmented reality apps. Specifically, it provides various kinds of (face, hand, body, etc.) detection and tracking algorithms which allow programmers to generate stunning visuals.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XII Dec 2023- Available at www.ijraset.com

The best is that many of these tools are supported in JavaScript, which means you can add superb features to your web apps and pages. The x, y, and z coordinates of 21 nodes get presented for each hand that it tool recognizes and maintains in a video feed (the equivalent of 4 joints per finger plus 1 palm). Large hand and finger tracking is possible using Media Pipe Hands, which uses machine learning (ML) to deduce 21 3D landmarks of a hand from a single frame.

It employs machine learning (ML) to infer 21 3D landmarks of a hand from just a single frame. As we know, Media Pipe Hands also utilizes an ML pipeline consisting of multiple models working together: A palm detection model which can be very effective that operates on the full image and returns an oriented hand bounding box. A hand landmark model that operates on the cropped image region is defined by the palm detector and returns high-fidelity 3D hand key points. Whereas the existing state-of-the-art techniques also increase in production desktop settings for inference, our method remains very simple but magnificent which runs at real-time speeds on a mobile device and can even handle users with multiple fingers.

By making this hand perception capability available we can also do a broader research and development community, we want to inspire the development of all the novel use cases and practices and lines of inquiry. An image-wide palm identification model as we are very much aware of can produce an aligned bounding box of the hand. High-fidelity 3D landmarks of the hand are returned by this model which can be further saved, which is applied to the clipped image region set by the palm detector.

We know that our Media Pipe method also makes use of a face detector and a representation of facial landmarks, thus this approach is conceptually comparable. By presenting the hand landmark model with a perfectly cropped image of a hand, the need for data augmentation which is a unique and effective method (such as rotations, translations, and scaling) is much reduced, freeing up the network's resources to focus on precision in its coordinate prediction.

Furthermore, we are aware that in our pipeline crops can also be created based on the 13 hand landmarks determined in the prior frame, and palm detection is only triggered to relocalize the hand when the landmark model can no longer identify hand presence. The pipeline is a Media Pipe graph that renders using a dedicated hand renderer subgraph and tracks hand landmarks with the hand landmark module.

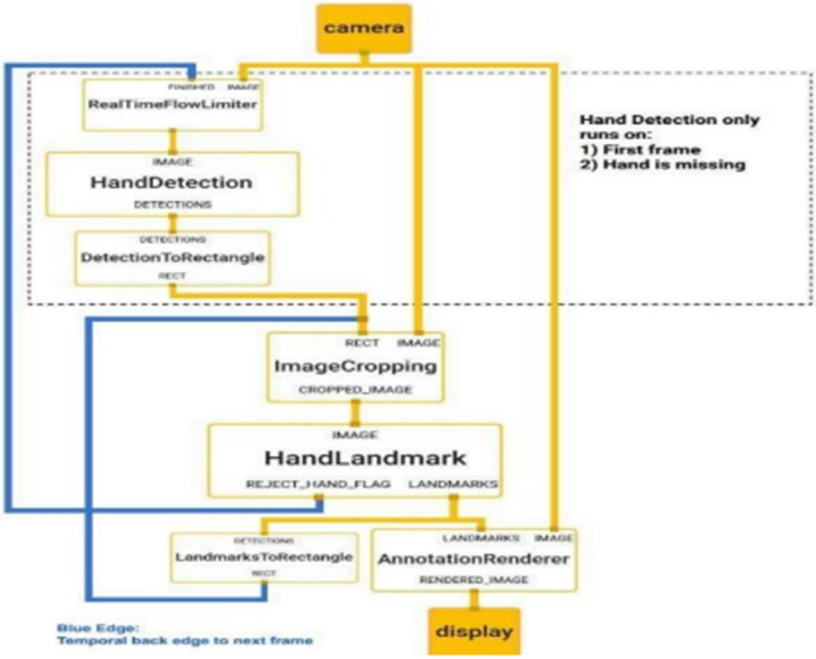


Figure 1: Flowchart of the Proposed System

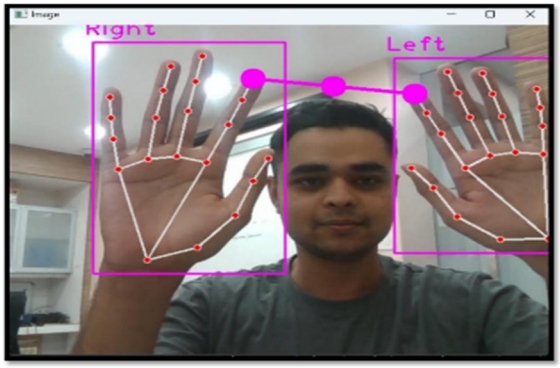
III. MODELING AND ANALYSIS

A. Palm Detection

Guided by Media Pipe's face detection model, we have developed a single-shot detector which is unique and that detects the initial hand locations, making it suitable for mobile real-time applications. Hands recognition as we know, presents a challenging problem because our lite model and complete model need to be able to detect hands that are occluded, and that span a wide range of sizes (20x) in proportion to the image frame.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

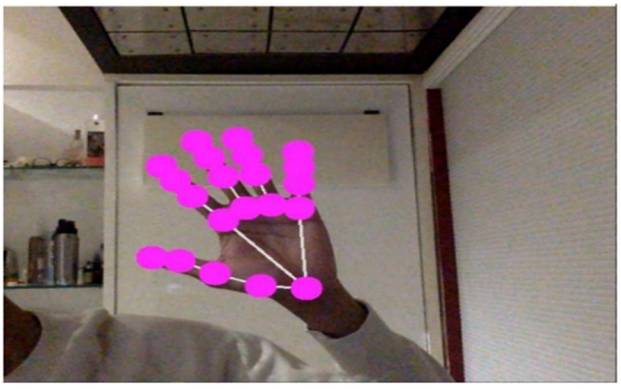
Detecting hands consistently based on visual cues can alone be more challenging than with faces because hands lack high contrast patterns found in the eye and mouth region, for example However, precise hand identification is aided by supplying extra information, such as arm, body, or human traits. Our approach takes a multi-pronged effort which makes it very easy to overcome the difficulties. To begin, instead of training a hand detector, we train a palm detector which is the hidden secret as it's considerably simpler to compute the bounding boxes of hard objects like palms and fists than it is to recognise hands with articulated fingers. The non-maximum suppression procedure is also good for two-hand self-occlusion scenarios, such as handshakes, because palms are relatively small objects. In addition, the number of anchors can be reduced by a factor of 3-5 by modelling palms using squared bounding boxes (anchors in Dl language) and ignoring all the other aspect ratios. Secondly, larger scene context awareness is provided, even for very little objects, by employing an encoder-decoder classifier (similar to the RetinaNet approach). Finally, because to the significant scale variance, we limit the focusing loss during training so as to sustain a large number of anchors which makes it unique.



5.2 Hand Detection Module

B. Hand Landmark Recovery

After a full-frame hand detection is completed, the hand landmark model uses regression models to just provide consistent key-point localization of 3D hand coordinates within the detected hand regions. This is achieved through the use of 21 hand-knuckle landmarks. Below are a few of them:



6.1 Landmark precision close to 91%

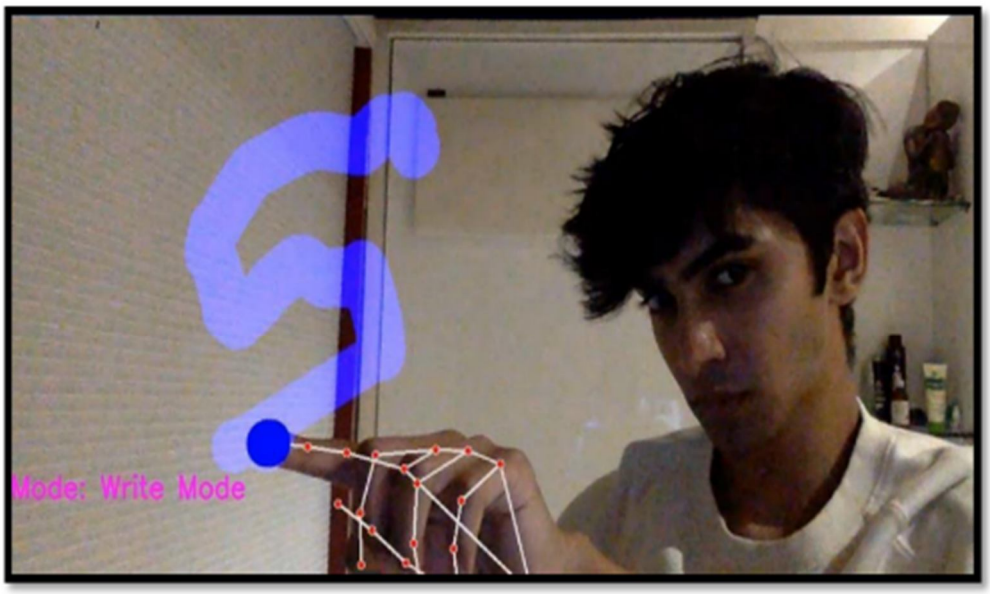


6.2 Fingers and Landmark positions

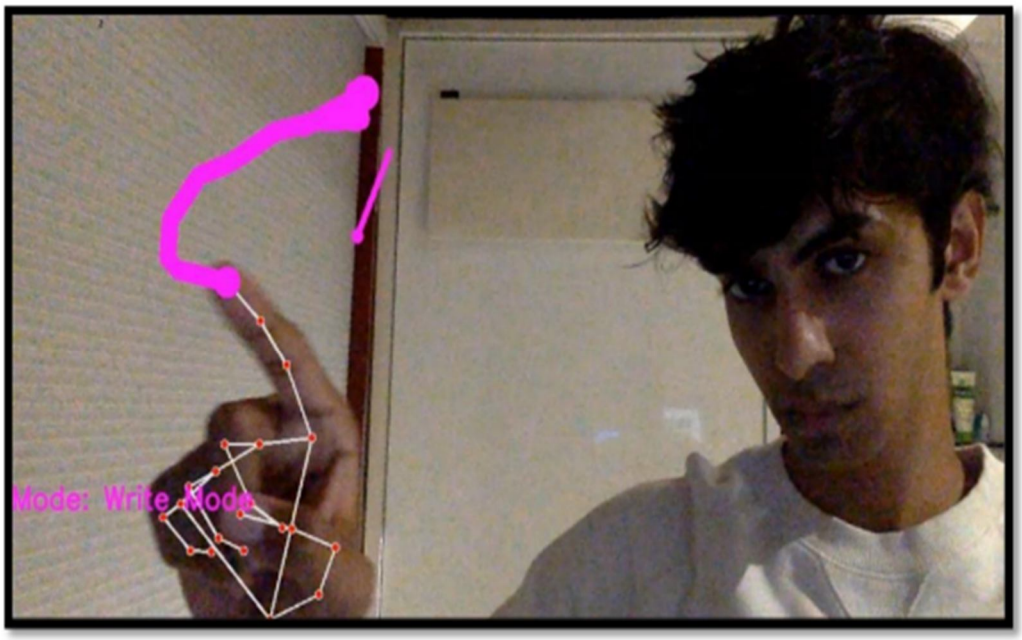
©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XII Dec 2023- Available at www.ijraset.com

IV. RESULTS AND DISCUSSION



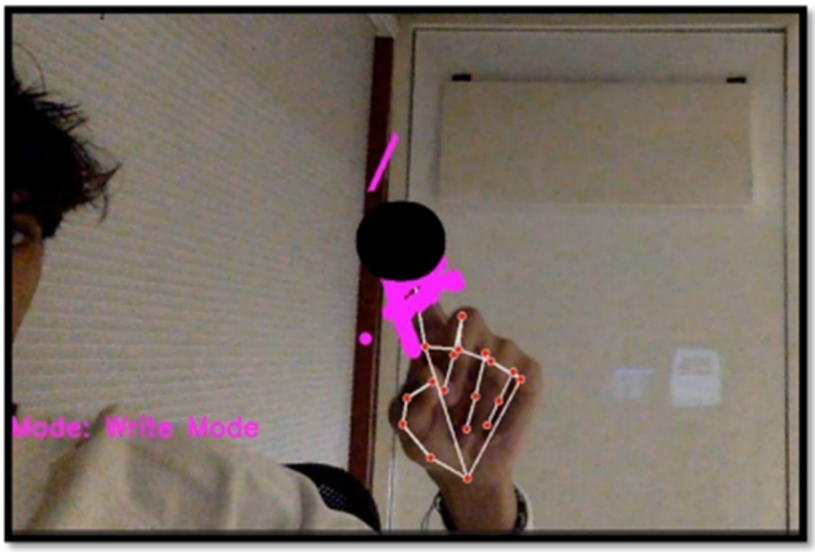
7.1 Writing with Highlighter



7.2 Writing with Pen

In its entirety, the Media Pipe Hand Tracking Flowchart as we know provides the hand landmarks as the final product. As was demonstrated above, once the last part was installed, the system was ready for testing in its process. Then, we had to put through its boosts in a variety of test cases, finding that it exceeded all expectations in each one, including hand detection and section switching based on the classification of the user's hand postures. So, the infrastructure is suitable for use as a digital and comprehensive medium, or AirPen.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



7.4 Selecting Eraser

V. CONCLUSION

In this paper, we presented the system architecture of Virtual Hand, a virtual reality interface designed for interactive sketching. The system combines hand gesture recognition and sketching techniques to allow users to sketch in mid-air using their fingers. We introduced a new framework that utilizes webcam video input to recognize mid-air finger writing. Our framework includes a new writing hand pose detection algorithm for air-writing initialization, a distance-weighted curvature entropy signature function for fingertip detection and tracking, and a fingertip velocity based termination criterion to mark the completion of air-writing gestures. Our experiments on our air-writing dataset demonstrate the superior performance of our fingertip detection and tracking approach compared to state-of-the-art trackers, with real-time performance in terms of frame rate. In the future, we plan to extend our framework to recognize words and signatures, which could have potential applications in touchless and marker-less biometric authentication systems. We also intend to explore more robust fingertip detection techniques to enhance the overall performance of the air-writing recognition system.

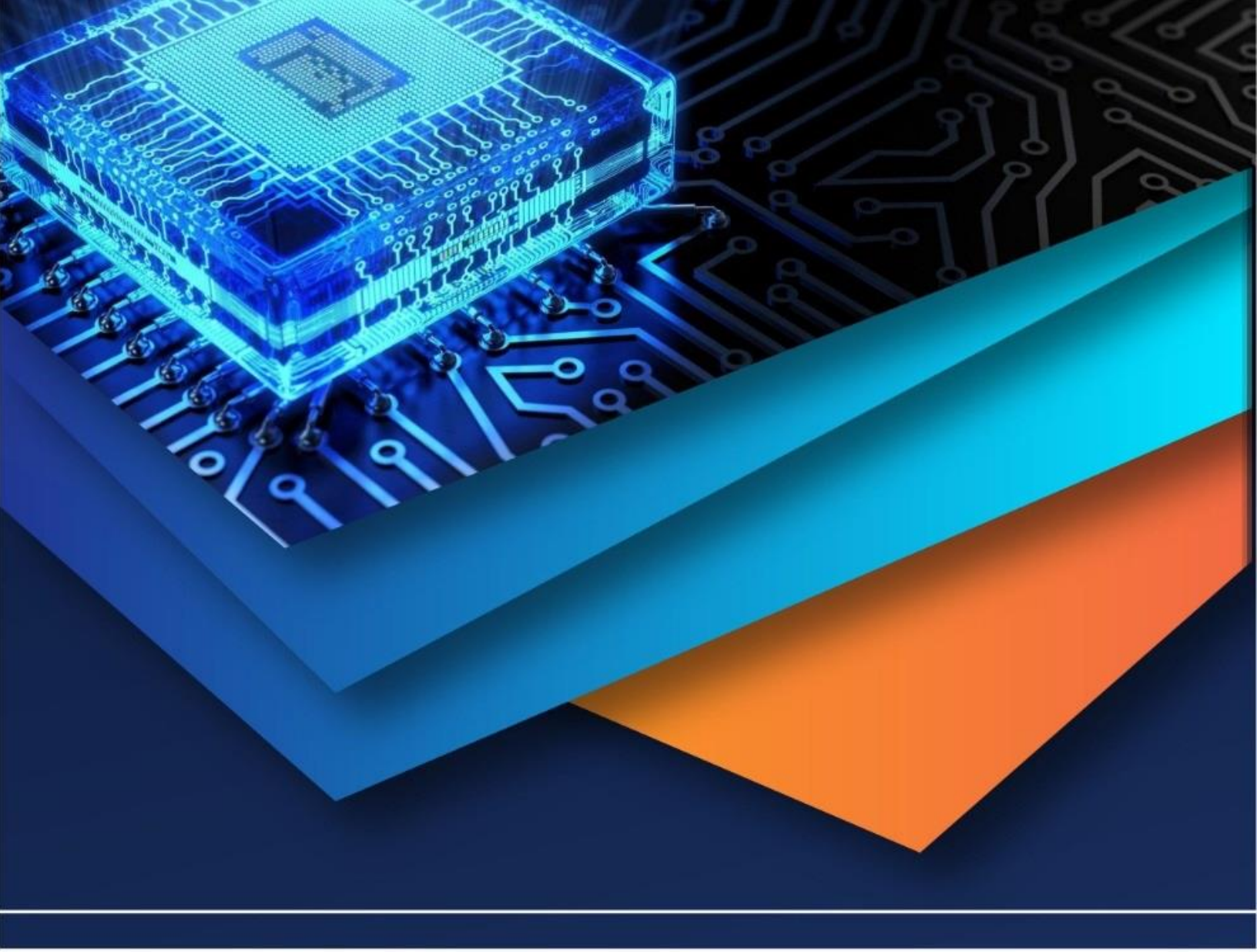
VI. ACKNOWLEDGEMENTS

We have invested effort in this endeavor. But, it would not have been possible without the generous assistance and assistance of numerous persons and mentors. We would want to express our gratitude to each of them. We owe Prof. Vikram Kulkarni a great debt of gratitude for his sincere advice and frequent supervision, as well as for supplying the essential information about the project and for his assistance in completing the project. Finally, we would want to convey our gratitude and appreciation to both our mentor and our examiners for their time and consideration. Our gratitude also extends to our coworkers and friends who assisted us in developing the project, as well as those who volunteered their assistance.

REFERENCES

[1] Mukherjeea, S., Ahmedb, S. A., Dograc, D. P., Kard, S., & Roye, P. P. (2018). Fingertip detection and tracking for recognizing air-writing in videos. Journal of Expert Systems with Applications.
[2] Kavakli, M., & Jayarathna, D. (2005). Virtual Hand: An interface for interactive sketching in virtual reality. In IEEE International Conference on Computational Intelligence for Modelling, Control and Automation, and International Conference on Intelligent Agents, Web Technologies and Internet Commerce.
[3] Chang, Y. H., & Chang, C. M. (2010). Automatic hand-pose trajectory tracking system using video sequences. INTECH
[4] Sudderth, E. B., Mandel, M. I., Freeman, W. T., & Willsky, A. S. (2004). Visual hand tracking using nonparametric belief propagation. MIT Laboratory for Information & Decision Systems Technical Report P-2603, presented at IEEE CVPR Workshop on Generative Model Based Vision.
[5] Wang, R. Y., & Popovi'c, J. (2008). Real-time hand-tracking with a color glove
[6] Bragatto, T. A. C., Ruas, G. I. S., & Lamar, M. V. (2006). Real-time video-based finger spelling recognition system using low computational complexity artificial neural networks. IEEE ITS.
[7] Araga, Y., Shirabayashi, M., Kaida, K., & Hikawa, H. (2012). Real-time gesture recognition system using posture classifier and Jordan recurrent neural network. IEEE World Congress on Computational Intelligence, Brisbane, Australia.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)