**Air Canvas-Drawing in Air using AI**

**ABSTRACT:**

In the study of visual pattern recognition, Abstract-Drawing in Air has been one of the most fascinating and challenging research fields. In this context, visual pattern recognition refers to the ability to spot finger-tip movement. In a variety of applications, it enhances human-computer interaction. With the help of this concept, human-computer interaction will become more naturally seamless (HCI). The proposed approach performs two primary functions: first, it tracks the fingertip; second, it shows the fingertip's coordinates on the screen in any specified color. Instead of a camera, there is no need for a keyboard, a pen, or a glove. The standard flat-dimensional, rectangular, and empty (white) canvas used in traditional artworks is something that this concept of Air Canvas goes beyond. To create this project, we're using OpenCV and computer vision techniques. This project uses fingertip tracking and detection procedure to accomplish its purpose. It is advised especially during COVID-19 situation or any critical weather conditions to draw virtually on an air canvas using hand gestures without touching anything. For the deaf, especially the able, the elderly, and youngsters, this initiative will be a potent means of communication for educational reasons.

**INTRODUCTION:**

Modern technology has advanced to such a degree that "HUMAN COMPUTER INTERACTION" has come to play an increasingly significant role in our daily lives. The traditional form of writing is being overtaken by digital art in the age of the internet. Digital art refers to methods of artistic expression and dissemination using digital media. One of the defining features of digital manifestation is its reliance on modern science and technology. Traditional art refers to the kind of art that was developed prior to digital art. It can be easily broken down into visual art, audio art, audio-visual art, and audio-visual imaginary art, which encompasses literature, painting, sculpture, architecture, music, dance, theatre, and other works of art, depending on who is doing the analysis. Traditional and digital art are interdependent and related to one another. Although social progress is not a result of popular demand, basic human necessities nonetheless serve as the primary inspiration. In art, the same thing takes place. In the circumstances at hand, we need to thoroughly grasp the fundamental differences in form between digital art and conventional art because they are both inclusive of the symbiotic condition. The writing methods that are traditionally used include chalk and a board and a pen. The primary goal of digital art is to develop a system for hand motion recognition so that digital writing may be done. There are numerous different ways to write in digital art, including utilizing a keyboard, touch-screen surface, digital pen, stylus, electronic hand gloves, etc. However, in this system, hand gesture detection is used in conjunction with a machine learning algorithm and python programming to produce a natural interface between the human and the machine. To develop this AI based Project, we will be using our trending techniques namely OpenCV and Python. Open cv is mainly known as an open-source computer vision and machine learning software.

**PROBLEM STATEMENT:**

Hand gesture recognition is an incredibly challenging task in the computer vision field. In this field you required detection and explanation of certain movements and poses of the hands. The goal of this research is to develop a healthy and accurate system for hand gesture recognition using computer vision techniques. In this research, the system should be able to detect and classify accurately for different hand movements in real phase of time. Also, the system should be performed in various conditions and in different viewpoints. In this research we aim to improve the performance for hand recognition systems and try to make them more universally applicable in fields such as gaming and sign language reorganization.

**PROJECT VISION:**

The project focuses on creating a motion-to-text converter that may one day act as software for intelligent wearables that enable writing from the air and can be used to good effect. This endeavor serves as a reporter of infrequent gestures. The finger's route will be traced using computer vision. Messages, emails, and other types of communication can all be sent using the created text. The deaf will be able to effectively communicate thanks to it. It is an efficient means of communication that decreases the use of mobile devices and laptops by doing away with the need to write.

In the era of digital world, traditional art of writing is being replaced by digital art. Digital art refers to forms of expression and transmission of art form with digital form. Relying on modern science and technology is the distinctive characteristics of digital manifestation. Traditional art refers to the art form which was created before digital art. From the recipient to analyze, it can simply be divided into visual art, audio art, audio-visual art, and audio-visual imaginary art, which includes literature, painting, sculpture, architecture, music, dance, drama and other works of art. Digital art and traditional art are interrelated and interdependent. Social development is not a people’s will, but the needs of human life are the main driving force anyway. The same situation happens in art. In the present circumstances, digital art and traditional art are inclusive of the symbiotic state, so we need to systematically understand the basic knowledge of the form between digital art and traditional art. The traditional way includes pen and paper, chalk and board method of writing. The essential aim of digital art is to build hand gesture recognition systems to write digitally. Digital art includes many ways of writing like by using keyboard, touch-screen surface, digital pen, stylus, using electronic hand gloves, etc. But in this system, we are using hand gesture recognition with the use of machine learning algorithm by using python programming, which creates natural interaction between man and machine. With the advancement in technology, the need for development of natural ‘human – computer interaction (HCI)’ [10] systems to replace traditional systems is increasing rapidly.

**METHODOLOGY:**

Writing in the air has become a fascinating and challenging field of study in image processing and pattern recognition in recent years. It has the potential to greatly advance automation techniques and improve how people and technology interact in a variety of applications.

The development of new processes and techniques that can reduce processing time and increase recognition precision has been the focus of numerous research studies. Due to the availability of inexpensive, high-quality video cameras, faster computers, and the rising need for automated video analysis, object tracking has become a crucial job in the field of computer vision. Object tracking is the process of spotting and following an object's movement in a video sequence. It has many uses, including surveillance, navigation, and activity recognition. To complete this job, you must overcome a number of obstacles, including handling occlusions, changes in scale and orientation, and motion blur, among others. However, the improvements in computer vision hardware and algorithms have significantly advanced object tracking methods and allowed for the development of real-world applications.

In general, object detection, object tracking, and object behavior analysis are the three major steps in video analysis. The right choice of object representation, feature selection for tracking, object recognition, and object tracking are the four main problems that need to be solved in object tracking. There are many practical uses for object tracking algorithms, such as automatic surveillance, video indexing, and car navigation.

Due to a variety of variables, including the complexity of the scene, changes in illumination, occlusions, and variations in scale and viewpoint, object tracking is a difficult job. In order for the object to be described and recognized in consecutive frames, object representation is essential for object tracking. Another crucial issue is featuring selection for tracking, which entails choosing pertinent visual characteristics that can accurately represent the object's appearance and distinguish it from other objects in the scene.

In order to fill this gap, the project is concentrating on creating a motion-to-text converter that could function as software for wearable technology that allows for in-air writing. This endeavor serves as a report of infrequent movements. The finger's route will be traced using computer vision. Additionally, the generated text can be used for a number of things, including sending emails and texts. The mute will be able to communicate effectively thanks to it. It is an efficient method of communication that decreases the use of mobile devices and laptops by doing away with the need to type.

**5 c’s:**

**CONSENT**:

Before collecting, processing, or sharing a user's personal data, specific consent from stakeholders or users is required. Consent is a crucial factor in the AI canvas sketching project because the system will be gathering and analyzing hand gestures and translating them into text. Thus, it's critical to make sure users are aware of the system's goals and provide clear authorization for the usage of their personal information. The system should give users the ability to opt out or withdraw their consent at any time, and it should clearly describe the data collecting and use procedure to ensure that consent is received in a clear and transparent manner.

**CLARITY**:

Clarity describes the necessity of openness and unambiguous communication in all facets of the endeavor. Clarity is fundamental in the case of the AI canvas-drawing project in terms of the system's functionality, goal, and data utilization. As a result, the system must be created in a fashion that is simple for users to comprehend and that gives precise information about how data is collected, used, and any potential hazards or system restrictions. Clarity also entails making sure users can make knowledgeable choices about their involvement in the system and are aware of any potential drawbacks.

**CONTENT**:

The ethical issues surrounding the data that is gathered and used by the system are referred to as content. In our AI canvas-drawing project, contents might include private information like hand gestures, SMS messages, and other forms of communication. As a result, it's critical to guarantee that the data is gathered and handled in an ethical and responsible manner. This could entail putting in place data security measures to stop unwanted access to or use of the data, or it could entail performing routine evaluations to find and rectify any potential biases or ethical issues relating to the data.

**CONSISTENCY:**

Consistency refers to the requirement that the system be consistent throughout time with moral precepts and values. This can entail making sure that the AI canvas-drawing project is conceived and developed in a manner that is consistent with moral concepts like openness, justice, and respect for user privacy. To make sure that the system remains compliant with moral standards and to spot any areas in need of adjustment or improvement, it must also be continuously monitored and reviewed. To assist uncover any potential ethical concerns or issues, this may entail carrying out routine audits or evaluations or putting in place a system for user feedback and input.

**CONSEQUENCES:**

Consequences refer to the system's and its use's possible ethical repercussions. Consequences could include privacy problems, biases in the data, or unanticipated uses of the system in the case of the AI canvas-drawing project. As a result, it's crucial to thoroughly analyze the system's possible effects and to take action to reduce any risks or moral dilemmas. This could entail putting in place privacy and security measures to safeguard user data, carrying out routine audits to spot and correct any biases in the data, or creating explicit rules for system usage to ward off any unauthorized or unethical applications.

**LIBRARIES**:

**Computer Vision and OpenCV:**

Computer vision is a branch of science that focuses on teaching computers to understand and analyze visual data from their surroundings. It involves the application of various techniques and algorithms to extract useful information from images or videos. Object recognition, motion analysis, and image enhancement are just a few of the uses for computer vision.

In this project, computer vision is used to allow the motion-to-text converter to recognize and interpret the user's gestures. The OpenCV library, which offers a set of pre-built functions for different image and video processing tasks, is a widely used tool in computer vision research. Because of its efficient and optimized implementation, OpenCV is especially well-suited for real-time applications.

In this project, the use of OpenCV allows for the creation of robust and efficient algorithms for detecting and tracking the user's finger movement in real-time. The library includes a variety of image processing functions, such as edge detection, filtering, and feature detection, that can be used to extract the required information from the video stream. Accurate and robust object tracking across frames is a crucial challenge in computer vision. Several object tracking algorithms, such as mean-shift and particle filter-based trackers, are provided by OpenCV and can be used to monitor the user's finger movement. These algorithms can be optimized to operate in real-time on devices with limited resources, such as wearable devices.

To summarize, the use of computer vision and OpenCV in this project allows for the creation of a motion-to-text converter that can identify and track the user's finger movement in real-time. The library contains many pre-built functions for different image processing tasks and object tracking algorithms that can be used to create efficient and robust algorithms for the project.



**Media Pipe:**

MediaPipe is a cross-platform open-source framework for creating real-time video processing pipelines developed by Google. It offers a variety of pre-built modules for different multimedia processing tasks such as object detection, face detection, and hand tracking. MediaPipe can be used for hand tracking in the context of this project, which is a crucial component of the motion-to-text converter. Deep learning techniques are used by Media Pipe’s hand tracking module to identify and track hand movements in real time. The module can precisely monitor the position and orientation of the hand, as well as the movement of individual fingers.

In addition, MediaPipe includes a collection of pre-built modules for image and video processing tasks such as filtering, feature detection, and optical flow estimation. The OpenCV Python library is a computer vision library that is extensively used for image analysis, image processing, detection, and recognition, among other things. The use of multi-threading allows Media Pipe to accomplish its speed. Such development techniques are usually difficult, but Media Pipe handles them for you as long as you follow good graph-making practices.

In overall, this project's use of MediaPipe makes it possible to create a precise and effective hand tracking module that can operate in real-time on devices with limited resources. A reliable and effective motion-to-text converter can be created using the framework's extensive library of pre-built components for handling multimedia processing tasks.

**Tensor Flow:**

TensorFlow is an open-source machine learning framework developed by Google. It has a wide variety of uses in computer vision, natural language processing, and other fields, and is frequently used to create deep learning models, including neural networks.

TensorFlow can be used in this project to learn and improve the hand tracking model. The high-level TensorFlow API, such as Keras or TensorFlow Lite, which offers a user-friendly UI for creating deep learning models, can be used to create the hand tracking module. TensorFlow also offers pre-built models that can be adjusted for hand tracking for a variety of computer vision tasks, including object detection and picture classification.

Quantization and pruning are two of the many methods for model optimization provided by TensorFlow, which can be used to improve the hand tracking model for effective deployment on devices with limited resources. Additionally, the framework offers APIs for deploying models on different operating systems, such as iOS and Android.



**Hand Tracking:**

The process of hand tracking involves a computer identifying a hand from an input picture and maintaining attention on the hand's movement and orientation. We can create a variety of programmes that use hand movement and direction as their input thanks to hand tracking. The hand tracking module is an important component of the overall system, as it provides the input for the motion-to-text converter.

To implement hand tracking as part of our program, we frequently reuse the same code across various tasks. This issue is resolved by building a hand monitoring module because we only need to write the code once. Next, we create a module out of this bit of code. This tool performs hand tracking and can be imported into any Python project that we are working on.

* **Palm Detection:**

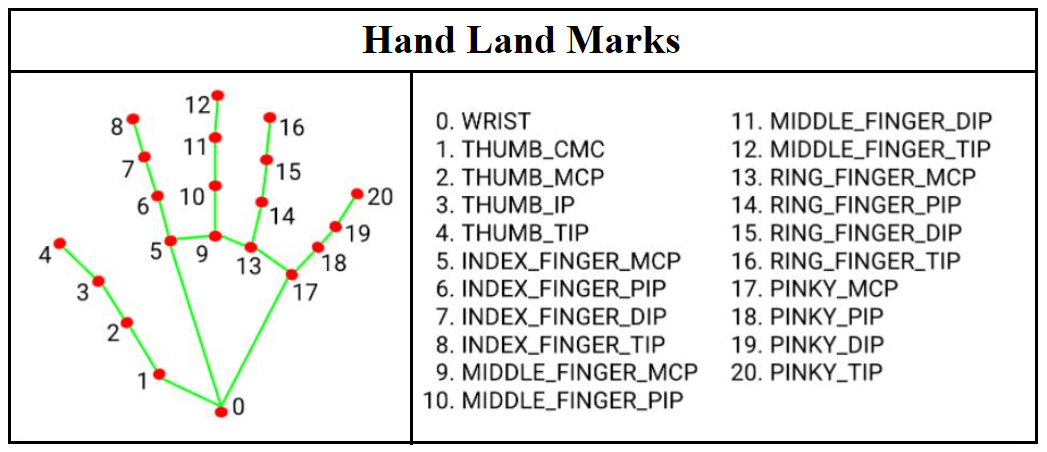
Palm detection is a crucial part of the hand monitoring module in the Air Canvas project. The process of palm recognition entails detecting the hand's palm from an input image or video stream. The palm detection algorithm looks at the input picture or video frame to determine the hand's palm and locates the palm within the hand area after first identifying the hand region in the image. It is developed using the OpenCV library, which offers a number of tools for processing images and videos. Implementing the palm recognition algorithm can be done easily through the OpenCV library's built-in functions for contour analysis, convex hull computation, and palm detection.

Once the palm has been identified, it is used as a point of reference to follow the hand's motion and position. This makes it possible for us to translate hand gestures quickly and precisely into digital writing.

* **Hand Landmark Identification:**

The hand landmark recognition algorithm looks at the input image or video frame to identify the hand's important points. This is accomplished by using a sizable dataset of hand image data to build a deep learning model. Convolutional neural networks (CNNs) are used in the deep learning model to learn features unique to hands, allowing it to recognize hand landmarks with accuracy. Each received frame is compared to MediaPipe hand landmarks, and finger positions are determined using getPositions() and which finger is opened using getUpFingers() methods of the hand tracker class.

The Hand Tracker class includes these two methods.



**FUNCTIONALITIES**:

**Hand Gesture Detection:**

Hand gesture detection is a critical component of the Air Canvas project because it allows users to engage with the system using hand movements and gestures. Computer vision techniques are used to track and identify the hand's position and motion in real-time to spot hand gestures. This is accomplished by analyzing video footage from a camera aimed at the user's hand.

As mentioned before, the first step in hand gesture detection is hand detection and landmark identification. After detecting the hand and identifying landmarks, the hand's motion is tracked over time using methods such as optical flow or Kalman filtering. Following that, hand gestures have been identified by analyzing hand motion over time. Machine learning methods such as support vector machines (SVMs), random forests, and deep neural networks can be used to accomplish this.

Machine learning algorithms are taught on a dataset of hand gestures, with each gesture labelled with the class to which it belongs. The training data is used to teach the algorithm how to recognize various hand motions.

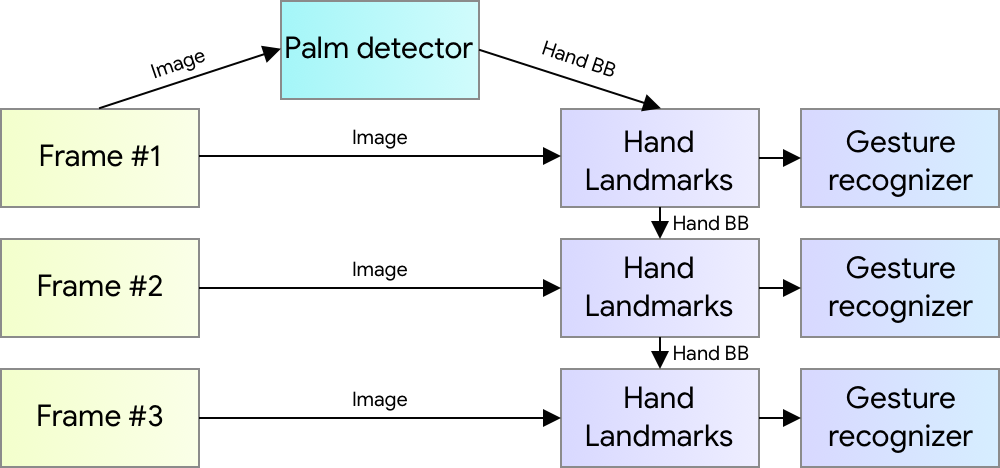
During real-time operation, the algorithm receives motion data from the hand and predicts the associated hand gesture class based on the learned patterns in the training data. As a result, the system can recognize and react to various hand gestures in real time.

**Gesture Recognition:**

After hand movements are identified, Machine learning algorithms are used to map hand gestures to their associated text labels to perform gesture recognition. This is accomplished by using a dataset of hand gestures and their associated text labels to train the algorithm. Gathering a dataset of hand motions and the text labels that go with them is the first stage in the gesture recognition process. Users can accomplish this by making various hand motions while concurrently recording the corresponding text characters.

The relevant traits are then extracted from the hand gesture data using feature extraction techniques. These characteristics can include the hand's location and orientation, its movement direction and speed, and its shape.

Overall, collecting data, extracting features, and machine learning are all important steps in the complicated process of gesture detection. Users will be able to interact with the Air Canvas system using their own natural hand movements if the proper methods and tools are used to correctly recognize and map hand gestures to their corresponding text labels.



**Text Generation:**

Text generation, which allows the system to generate digital text from the hand gestures identified by the gesture recognition module, is an essential part of the Air Canvas project. The system must integrate a text generation module with a gesture detection module in order to produce text from hand gestures. Several natural language processing (NLP) methods, such as language models and sequence-to-sequence models, can be used to accomplish this.

A particular kind of neural network called sequence-to-sequence models produces a sequence of output data (In this case, digital text) from a sequence of input data (in this instance, data on hand gestures). These models usually use a combination of encoder and decoder modules and are trained on large datasets of input-output pairs to map input sequences to output sequences.

**User Interface:**

A GUI is likely the best choice in the case of Air Canvas because it enables users to view their hand motions and the resulting digital text in real-time. The user will need to receive clear feedback on their hand movements and the output text from the GUI in order for it to be intuitive and simple to use. A video feed of the user's hand motions, a list of the recognized hand gestures, and a preview of the generated digital text are just a few features that the GUI might offer. Additionally, it might have buttons or other interactive components that let users change the system's settings, like the text's font or size.

The system might include a voice-based interface in addition to the GUI to allow for hands-free operation. This might entail using voice commands to start and stop the system, choose various operating modes, or manage other UI elements.

**Reduction of Device Usage:**

One of the main benefits of the Air Canvas system is the decrease in device utilization. The system can do away with physical input devices like keyboards and touchscreens by enabling users to generate digital text using only hand gestures. This is especially helpful when using a physical device would be difficult or impractical, such as when the user's hands are full or when one is not immediately accessible.

Additionally, the fact that the Air Canvas system can run without extra hardware reduces the system's cost and complexity, opening up a broader range of users to it. Additionally, the system removes the need for users to transport and maintain extra hardware, which can be a significant burden, particularly for people who are constantly on the move.

Additionally, since fewer physical devices need to be produced and disposed of, the Air Canvas system's decreased device utilization may have positive environmental effects. The system might assist in reducing some of the harmful effects that electronic devices have on the environment by cutting down on the quantity of e-waste produced.

**Performance Optimization:**

The project will optimize the performance of the system to ensure that it can recognize gestures in real-time and generate text accurately. This may involve optimizing the computer vision algorithms or using specialized hardware, such as GPUs.

**Accessibility:**

This project's ability to increase accessibility for those with hearing or speech impairments is one of its main advantages. These people can interact effectively thanks to the motion-to-text translation, which eliminates the need for laptops or mobile devices.

**LITERATURE SURVEY:**

[1] [**https://www.irjet.net/archives/V9/i2/IRJET-V9I2183.pdf**](https://www.irjet.net/archives/V9/i2/IRJET-V9I2183.pdf)

**Summary:** This article is from the International Research Journal of Engineering and Technology (IRJET). It explained that the proposed profound outfit community is ready disconnected using records stuck through an outer tracker (Opti track V120) and the define primarily based totally approach. They use the brush tip function to take the brush deal with acts like a record, allowing us to make use of actual cloth with an actual brush. During the checking out system, the framework works continuously, considering that round then, it tracks the brush deal with present (function and direction) and the proposed profound troupe community takes this brush deal with act like information and predicts the brush tip function in genuine time. For information assortment, played out various strokes for 60 seconds on a superficial level inside the following area.

**[2]**[**https://ijariie.com/AdminUploadPdf/Air\_Canvas\_Application\_using\_OpenCV\_and\_Numpy\_in\_Python\_ijariie17660.pdf**](https://ijariie.com/AdminUploadPdf/Air_Canvas_Application_using_OpenCV_and_Numpy_in_Python_ijariie17660.pdf)

**Summary:** This survey is from Dayananda Sagar Academy of Technology and Management,Bangalore,Karnataka,INDIA. It explained that hand gesture recognition has multiple techniques. It might be difficult to identify the hand because of the different skin tone and hand from each person. The technology uses two neural networks to identify the outcomes. The SCD (Scalable color descriptor) neural network is the first algorithm. The picture pixels are fed into the SCD neural network, which determines whether they are skin pixels. The second is HGR (Hand gesture recognition) neural network in which the extracted features will be added. The features are to be extracted by two distinct algorithms namely Finding the fingertip and Pixel segmentation algorithm.

**CONCLUSION:**

In conclusion, by utilizing computer vision and machine learning techniques to give users the ability to sketch and write in the air using hand gestures recorded by a camera, the Air Canvas project has the potential to revolutionize conventional drawing and writing methods. This project functions as a cutting-edge illustration of how users can engage with the digital world in a distinctive and immersive manner.

This project removes the need for physical devices like laptops or mobile devices, making it a convenient and natural way to create digital content. It does this by utilizing hand tracking and gesture recognition technologies. Users can easily express their creativity in the air using the visual-based pointing technique, which gives text and drawings a completely new life.

Overall, the Air Canvas project shows great promise and has the potential to influence digital art and how people engage with computers in the future. To add depth and credibility to the conclusion, references to pertinent technologies and methods used in the project, such as computer vision, machine learning, and graphical user interface (GUI), can also be made. The conclusion can be strengthened by providing examples of real-world applications and the possible effect of Air Canvas in pertinent industries or domains. The main conclusions and project consequences can be concisely summarized up in a conclusion that is thorough and well-supported, leaving readers with a strong impression.

We have been successful in creating an AI algorithm that can detect and understand 10 different hand gestures and convert those gestures into text output with an accuracy score of 86%. We have also been documenting all the progress using different documents. We have used MNIST dataset that has handwritten numbers and is used to train our model to understand the inputs written by the user and understand handwriting and convert it to text outputs.

For future work we will be using CV2 libraries of python to develop a user interactable and computer vision enabled environment that can capture user inputs without using any external device to make art and interaction with technology more immersive and user friendly. We aim to revolutionize how humans interact with technology the same way when touch devices changed the interaction with technology.

**REFERENCES:**

[1] L. A. Gatys, A. S. Ecker, and M. Bethge. “Image Style Transfer Using Convolutional Neural Networks”. June 2016.

[2] Gangadhara Rao Kommu, Assistant ProfessorDepartment of Information technology, Chaitanya Bharati Institute of Technology, Hyderabad, India “AN EFFICIENT TOOL FORONLINE TEACHING USING OPENCV”.

[3] PranaviSrungavarapu, Eswar Pavan Maganti, SrilekkhaSakhamuri, Sai Pavan Kalyan Veerada, Anuradha Chinta “Virtual Sketch using Open CV” International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075 (Online), Volume-10 Issue-8, June 2021.

[4] P. Ramasamy, G. Prabhu and R. Srinivasan, "An economical air writing system converting finger movements to text using web camera," 2016 International Conference on Recent Trends in Information Technology (ICRTIT), 2016, pp. 1-6, doi: 10.1109/ICRTIT.2016.7569563.

[5] Prof. S.U. SaojiNishthaDua Department of Computer Engineering Department of Computer Engineering Bharati Vidyapeeth Bharati Vidyapeeth (Deemed to beUniversity) (Deemed to be University) College of Engineering, Pune College of Engineering, Pune Akash Kumar Choudhary Bharat Phogat Department of Computer Engineering Department of Computer Engineering Bharati Vidyapeeth Bharati Vidyapeeth (Deemed to be University) (Deemed to be University) College of Engineering, Pune College of Engineering, Pune “ AIR CANVAS APPLICATION USING OPENCV AND NUMPY IN PYTHON”International Research Journal of Engineering and Technology (IRJET) Volume: 08 Issue: 08 | Aug 2021.

[6] 6. SairaBeg, M. Fahad Khan, FaisalBaig “Text Writing in Air” Journal of Information Display Volume 14, Issue 4, 2013.