International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 3, pp 242-246, March-2024

https://doi.org/10.47001/IRJIET/2024.803035

Cursor Movement by Hand Gestures

[1]Umesh Bhosale, [2]Abhishek Badade, [3]Rushikesh Bhalerao, [4]Samiksha Chorge, [5]Prof. M.P. Desai

[1,2,3,4]Student, Information Technology, Smt. Kashibai. Navale College of Engineering, Pune, Maharashtra, India

[5]Professor, Information Technology, Smt. Kashibai. Navale College of Engineering, Pune, Maharashtra, India

***Abstract* - Our project develops a system that can recognize the air-written words in 3D space, and then classify the recognized character into one of the possible classes. Air-writing is the new way of writing the linguistic characters or words in free area using hand or finger movements. Writing within the air may be a method to jot down one thing in an exceedingly 3D house with our finger-tip. This paper presents a simple yet effective air-writing recognition approach based on deep convolutional neural networks (CNNs). A robust and efficient hand tracking algorithm is proposed to extract air-writing trajectories collected by a single web camera. The algorithm addresses the push-to-write problem and avoids restrictions on the users' writing without using a delimiter and an imaginary box. A novel preprocessing scheme is also presented to convert the writing trajectory into appropriate forms of data, making the CNNs trained with these forms of data simpler and more effective. This project could be a combination of computer vision object chase and handwriting recognition machine learning. The air writing recognition system uses the digital camera of a pc to trace character, digits written within the air by the user and then uses a convolutional neural network to classify the character and digits into one of the possible classes. Several current systems use advanced and high-priced chase setups to realize gesture recognition, however, we tend to get to form a system that may attain similar work with a far cheaper setup.**

Keywords: Handwriting Recognition, Handwriting Classification, Computer Vision, Machine Learning, and Convolutional Neural Networks (CNN), etc.

I. INTRODUCTION

Several new interfaces provide possibilities for the realization of human-computer interaction including the controlling systems based on human movement. The human - computer interaction is able to detect the cognitive actions, the human and computer can have much more than an interaction (HCI). On the basis of the results of the mouse cursor movement analysis differences were shown in the course of controlling performed by a conventional mouse, and hand gesture detection.

This project focuses on enabling cursor movement on a computer screen through hand gestures, enhancing accessibility and user experience. It involves the development of a real-time hand gesture recognition system for precise cursor control, with potential applications in accessibility and interactive interfaces.

Handwriting analysis could be a classic, well-explored downside within the realm of introductory machine learning that encompasses several of the necessary topics of neural networks. Air-writing is particularly helpful for user interfaces that don't permit the user to kind on a key- board or pen a trackpad/touchscreen, or for text input for sensible system management, among several applications. Air-writing differs from typical hand writing; the one that performs air writing will solely use associate degree unreal coordinate to guide the writing motion. The variability of motion knowledge represents a letter is so significantly broader in air-writing than in paper writing. Our project aims to use a mixture of computer vision and handwriting recognition to form a system that acts as a virtual whiteboard. Model users would be ready to write "air words" facing a web camera either real-time or in advance and have those gestures translated into letters or digits. The aim of this project is to further explore the task of classifying handwritten text and to convert handwritten text into a digital format and achieving a virtual whiteboard system at a cost that is accessible to the average user. We want to introduce alternative interfaces for communication that have high affordability, usability, and accessibility.

II. LITERATURE SURVEY

An early notable attempt in the area of character recognition research is by Grimsdale in 1959. The origin of a great deal of research work in the early sixties was based on an approach known as the analysis-by-synthesis method suggested by Eden in 1968. The great importance of Eden's work was that he formally proved that all handwritten characters are formed by a finite number of schematic features, a point that was implicitly included in previous works. This notion was later used in all methods in syntactic (structural) approaches of character recognition.

K. Gaurav, Bhatia P. K. Et al,[6] this paper deals with the various pre-processing techniques involved in character recognition with different kinds of images ranges from simple

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 3, pp 242-246, March-2024

https://doi.org/10.47001/IRJIET/2024.803035

handwritten form-based documents and documents containing colored and complex backgrounds and varying intensities. In this, different pre-processing techniques like skew detection and correction, image enhancement techniques of contrast stretching, binarization, noise removal techniques, normalization and segmentation, morphological processing techniques are discussed. It was concluded that using a single technique for pre-processing, we can't completely process the image. However, even after applying all the said techniques might not possible to achieve full accuracy in a pre-processing system.

In [7], Renata F. P. Neves have proposed SVM based offline handwritten digit recognition. Authors claim that SVM outperforms the Multilayer perceptron classifier. Experiment is carried out on NIST SD19 standard dataset. Advantage of MLP is that it is able to segment non-linearly separable classes. However, MLP can easily fall into a region of local minimum, where the training will stop assuming it has achieved an optimal point in the error surface. Another hindrance is defining the best network architecture to solve the problem, considering the number of layers and the number of perceptron in each hidden layer. Because of these disadvantages, a digit recognizer using the MLP structure may not produce the desired low error rate.

In [12] the system proposed used the depth and color information from Kinect sensor to detect the hand shape. As for the gesture recognition, even with the Kinect sensor. It is still a very challenging problem. The resolution of this Kinect sensor is only 640×480. It works well to track a large object. e, g. the human body. But tracking a very small like finger is complex

Existing System:

Multimedia content analysis is applied in different real-world computer vision applications, and digital images constitute a major part of multimedia data. In last few years, the complexity of multimedia contents, especially the images, has grown exponentially, and on daily basis, more than millions of images are uploaded at different archives such as Twitter, Facebook, and Instagram. To search for a relevant image from an archive is a challenging research problem for computer vision research community. Most of the search engines retrieve images on the basis of traditional text-based approaches that rely on captions and metadata. In the last two decades, extensive research is reported for content-based image retrieval (CBIR), image classification, and analysis. In CBIR and image classification-based models, high-level image visuals are represented in the form of feature vectors that consists of numerical values. The research shows that there is a significant gap between image feature representation and human visual understanding. Due to this reason, the research presented in this area is focused to reduce the semantic gap between the image feature representation and human visual understanding. In this paper, we aim to present a comprehensive review of the recent development in the area of CBIR and image representation using Optical Character Recognition (OCR). We analyzed the main aspects of various image retrieval and image representation models from low level feature extraction to recent semantic deep-learning approaches. The important concepts and major research studies based on CBIR and image representation are discussed in detail, and future research directions are concluded to inspire further research in this area.

III. PROBLEM STATEMENT

Nowadays human computer interaction Handwriting analysis is a classic, well-explored problem in the introductory machine learning that catches many of the important topics of conventional neural networks. When we discussed handwriting analysis during project work, we realized that handwriting analysis is a project that could be extended with other machine learning concepts for interesting combination applications. In this proposed system, presents a simple yet effective Cursor Movement with air-movement recognition approach based on deep convolutional neural networks (CNNs). A robust and efficient hand tracking algorithm is proposed to extract air-writing trajectories collected by a single web camera.

IV. MOTIVATION

The world is changing incredibly fast, and we are not all aware of it. Proposed technology and Machine Learning are irreversible advancements that is disrupting established industries and the ways in which we interact financially. While a large body of research exits on the topic of learning technologies from a variety of disciplines, there is still lack of classrooms based evidence which describes and evaluates their use. With implementing this concept in classrooms we take a step towards growing digital world. Now is the era of digital world, in future teachers will use this method to make their teaching easy and understandable.

This concept motivates us to make this program more advance by implementing further features for eg. Digitally drawing, sketching, etc.

V. METHODOLOGY

Object Detection:

In our model, object tracking is realized with the detection of the target object through image analysis of each

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 3, pp 242-246, March-2024

https://doi.org/10.47001/IRJIET/2024.803035

video frame OpenCV -an open-source computer vision and ma- chine learning library provides enough functions and facilities so that we can process our image with various computer vision algorithms hence the OpenCV library is used to detect, track, and save the trajectory of the target object as its position shifts throughout the video. The digital representation of each frame is pre-processed. The digital image is nothing more than a matrix of scalar or vector values (depending on whether the image is monochromatic or colored) and it can be processed in numerous ways. That is where the OpenCV library may be used since it is optimized for such operations. We use optical flow function. Motion of image objects between two consecutive frames which is caused by the motion of cameras that creates a pattern and it is called as Optical flow [14]. This is 2D vector field where each vector is a displacement vector showing the movement of points from first frame to second.

Data Acquisition:

The creation of a mask or subtraction of background depends on several circumstances. Whether there is a separate image of background available, whether there is a lot of noise in the image, variability in lights, and others. Not all objects have strong features which can be immediately recognized. There was the conversion of color space performed before actual sample selection. The reason for conversion is that RGB color space (default color space for OpenCV) defines color using 3 values, where the HSV (or HSB) color space defines color with only one value. We pre-process each frame by resizing and reducing Gaussian noise. Then, we construct a binary mask around the object and perform morphological transformations to clean up. We find the contours using an OpenCV algorithm which calculates the hierarchy of contours in the image and compresses it.

Normalization:

Image normalization is a typical process in image processing that changes the range of pixel intensity values. Its normal purpose is to convert an input image into a range of pixel values that are more familiar or normal to the senses, we normalize strokes and remove variations that would otherwise complicate recognition and reduce the recognition rate. The video frames captured from the camera may have noise especially when the ambient light around the sensor is low. The Gaussian blurring process reduces the sensor noise and reduces the effects of low light condition on input frame to some extent, resulting in less number of false contours and regions in our subsequent operation. Blurring of image is done by convolving the image with a low-pass filter kernel. The kernel here is a simple Gaussian kernel matrix. Each pixel in the input frame matrix gets multiplied by the Gaussian kernel. The output values of these multiplications are summed up and that result is used to set value at the destination pixel.

Classification of Letters:

After obtaining the pre-processed, 28x28 sized image of the written character, our next step is to run it through a trained convolutional. While training with pre-processed images, the input to CNN is a fixed-size 28x28 grayscale image. T The batch size was set to 256 and the number of epoch was 25. The image is passed through a stack of convolutional layers, each of which has 3x3 filters. When an input image is presented to the system, its features are extracted and given as input to the trained classifier like an artificial neural network or support vector machine. Classifiers com- pare the input feature with the stored pattern and find out the best matching class for input. Two activation functions are used for training the CNN are ReLU and Softmax. Relu (Rectified Linear Unit) Activation function has output 0 if the input is less than 0, and raw output otherwise f(x) = max(x,0) It results in much faster training for large net- work. The Softmax Activation function's input is a set of values and it operates on it to make their sum equal to one, which will give probabilities for each value, i.e 0¡softmax¡1.

Gesture Classification with A CNN:

We used a CNN to classify the image patterns. CNNs are extensively used as a deep learning technique that mimics the human vision system. A CNN consists of convolutional, pooling, and fully connected layers. In convolutional layers, the key features of the input image are extracted by a convolutional filter.

VI. PROPOSED SYSTEM

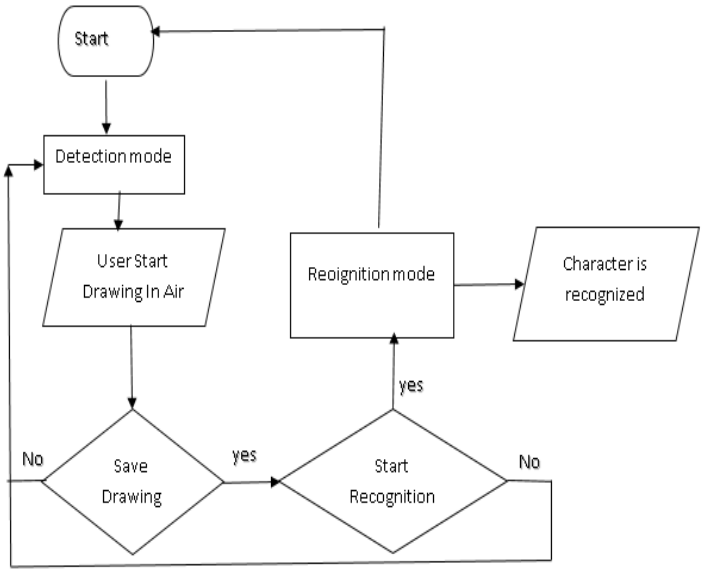

Figure 1: System Architecture

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 3, pp 242-246, March-2024

https://doi.org/10.47001/IRJIET/2024.803035

The proposed system uses the camera for Hand gesture recognition. In our prototype, two basic techniques are used that includes processing of a video input frame and then pattern recognition of processed frame using CNN.

Figure 1 depicts the architecture design of our system which has four modules that helps us to detect alphabets and numbers drawn by motion of any object in air. The basic working is that a live input from camera will be captured and pre-processed using background subtraction, Gaussian blurring and thresholding i.e. CNN algorithms to separate foreground objects of user from background and correctly detect it in any background condition. Object motion of user is then traced and the frame with tracked path is pre-processed and then given to the trained convolutional neural network which consists of large enough training dataset and our trained CNN model performs recognition of written text

VII. RESULTS AND DISCUSSIONS

The results showed that the Cursor Movement Based on Hand Gesture's validation test was deemed valid. The result screenshots of the system as follows:

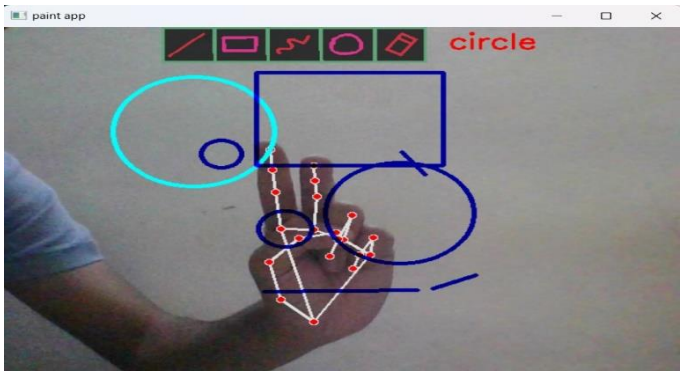

Figure 2: shape

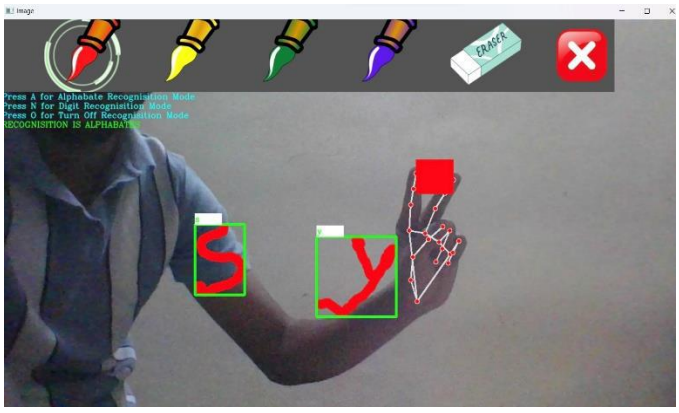

Figure 3: Character recognition

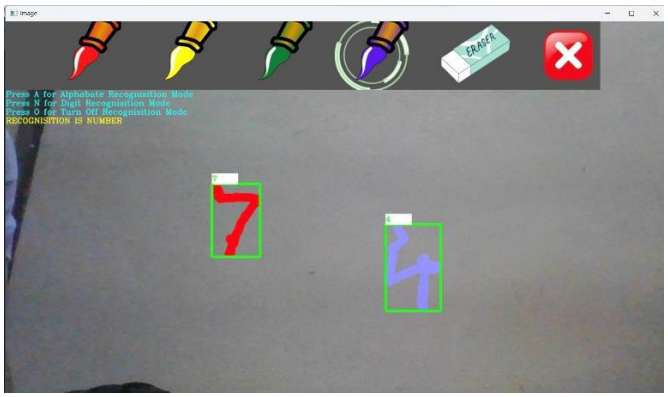

Figure 4: Number recognition

VIII. CONCLUSION

We have built the suggested system in this proposed system. Maximum work is done manually in the current system, which is prone to errors and requires time to update. In order to control a laptop device, a strong CNN model is provided in this study for air written digit and air movement recognition. To avoid the issues associated with human skin division, a uniformly shaded marker is used. A CNN model that has been trained beforehand and adjustments made to the air-writing data are used to achieve recognition. In this trial, the suggested solution supported all laptop devices and achieved a 93% recognition rate for English characters and digits. This method is totally independent of depth and makes use of a camera.

REFERENCES

[1] Badi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M. and Ghemawat, S., 2015. TensorFlow: Large-scale machine learning on heterogeneous systems.

[2] Bradski, G., 2000. The openCV library. Dr. Dobb's Journal: Software Tools for the Pro- fessional Programmer, 25(11), pp.120-123.

[3] Ganger, M., Duryea, E. and Hu, W., 2016. Double Sarsa and double expected Sarsa with shallow and deep learning. Journal of Data Analysis and Information Processing, 4(4), pp.159-176.

[4] Gregory Cohen, Saeed Afshar, Jonathan Tapson, and Andre´ van Schaik. 2017. EMNIST: an extension of MNIST to handwritten letters. arXiv preprint arXiv:1702.05373 (2017).

[5] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 12 (2011), 2825–2830.

ISSN (online): 2581-3048
Volume 8, Issue 3, pp 242-246, March-2024
https://doi.org/10.47001/IRJIET/2024.803035

[6] K. Gaurav and Bhatia P. K., "Analytical Review of Preprocessing Techniques for Offline Handwritten Character Recognition", 2nd International Conference on Emerging Trends in Engineering Management, ICETEM, 2013.
[7] Renata F. P. Neves, Alberto N. G. Lopes Filho, Carlos A.B.Mello, Cleber Zanchettin, "A SVM Based Off-Line Handwritten Digit Recognizer", International conference on Sys- tems, Man and Cybernetics, IEEE Xplore, pp. 510-515, 9-12 Oct, 2011, Brazil.
[8] Air-writing Recognition, Part 1: Modeling and Recognition of Characters, Words and Connecting Motions Mingyu Chen, Ghassan AlRegib, Senior Member, IEEE, and Biing- Hwang Juang, Fellow, IEEE.
[9] C.-M. Karat, C. Halverson, D. Horn, and J. Karat, "Patterns of entry and correction in large vocabulary continuous speech recognition systems," in Proc. of the SIGCHI Conference on Human Factors in Computing Systems, 1999, pp. 568–575.
[10] Yuan-Hsiang Chang, Chen-Ming Chang, "Automatic Hand-Pose Trajectory Tracking Sys- tem Using Video Sequences", INTECH, pp. 132- 152, Croatia, 2010.

AUTHORS BIOGRAPHY

Umesh Bhosale,
Student, Information Technology, Smt. Kashibai. Navale College of Engineering, Pune, Maharashtra, India.

Abhishek Badade,
Student, Information Technology, Smt. Kashibai. Navale College of Engineering, Pune, Maharashtra, India.

Rushikesh Bhalerao,
Student, Information Technology, Smt. Kashibai. Navale College of Engineering, Pune, Maharashtra, India.

Samiksha Chorge,
Student, Information Technology, Smt. Kashibai. Navale College of Engineering, Pune, Maharashtra, India.

Prof. M.P. Desai,
Professor, Information Technology, Smt. Kashibai. Navale College of Engineering, Pune, Maharashtra, India

Citation of this Article:

Umesh Bhosale, Abhishek Badade, Rushikesh Bhalerao, Samiksha Chorge, Prof. M.P. Desai, "Cursor Movement by Hand Gestures", Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 8, Issue 3, pp 242-246, March 2024. Article DOI https://doi.org/10.47001/IRJIET/2024.803035

****

