JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

DIGITAL AIR-CHARACTER AND DIGIT RECOGNITION

[1]Mr. M. P. Gangawane, [2]Mr. Pratik Rao, [3]Ms. Sanika Saurkar, [4]Mr. Samarth Saurkar, [5]Mr. Shreyash Padhar

[1]Assistant Professor, Guru Gobind Singh College of engineering and research center , Nashik-422009 , India
[2][3][4][5]Student, Guru Gobind Singh College of engineering and research center, Nashik-422009, India

1. ***Abstract :*** In the realm of human-computer interaction, our research introduces an innovative system for Fingertip Detection and Tracking to recognize air-written gestures. This project leverages computer vision, machine learning, and multi-core processing to enable users, including those with disabilities, to interact with digital devices through intuitive air-written commands. The system offers customization features, allowing users to define and personalize gestures. Beyond accessibility, this technology finds applications in education, healthcare, and industrial automation. The project's outcomes showcase advancements in gesture recognition, opening new avenues for inclusive and interactive computing experiences.

Keywords **- Fingertip Detection, Air-writing Recognition, Gesture Recognition, Computer Vision, Accessibility.**

2. INTRODUCTION

In the ever-evolving landscape of human-computer interaction, our project introduces an innovative system that not only recognizes air-written gestures but also understands strings conveyed through these dynamic movements. This research delves into the realms of computer vision, machine learning, and multi-core processing to create a seamless interface where users can communicate and input text through intuitive air-written gestures.

Traditional methods of text input often present challenges for individuals with diverse abilities. Our approach aims to redefine this interaction paradigm by enabling users to express themselves naturally in the air, translating their movements into meaningful strings. This goes beyond conventional gesture recognition, as our system incorporates advanced algorithms to decipher the intricate patterns and nuances inherent in air-written text.

The versatility of this project extends beyond recognizing predefined gestures; it includes the ability to interpret and understand strings of characters, opening up possibilities for touchless, expressive communication. Users can not only command their devices through intuitive gestures but also input text in a manner that mirrors the fluidity of their natural movements.

This project holds particular significance for individuals with disabilities, offering an alternative and inclusive means of text input. Moreover, the customization features empower users to define their own air-written strings, enhancing personalization and adaptability.

3. LITERATURE SURVEY

Sohom and Arif Ahmed [1] have proposed a new framework for the recognition of mid-air finger writing the use of internet-cam video as input. They supplied a new writing hand pose detection set of rules for the initialization of air-writing. Further, we used a novel signature function known as distance-weighted curvature entropy for fingertip detection and monitoring.

Adil Rahman [2] have proposed a robust and cheaper machine for the popularity of multi-digit numerals traced in an air-writing environment which uses most effective a prevalent tool camera for input. RNN-LSTM networks are used for noise elimination and digit recognition.

In [3] they have got developed an set of rules that solves the problem of detecting air-writing from well known finger movements on a motion-primarily based consumer interface. Use a controller-loose tracking system consisting of LEAP to song the finger motion.

Xin Zhang [4], they gives a finger-writing device that recognizes characters written within the air without the need for an extra hand held device. A depth-skin-history mixture model (DSB-MM) is proposed for hand segmentation to solve traditional troubles related to Kinect such as illumination variant, hand-face overlapping issues, and coloration-intensity mismatch.

In [5] The RGB digicam detects the fingertip and tracks its motion through out the screen. Whenever the hand is available in front of the digicam, the initial component to do is come across the fingertip, The RGB digital camera detects the fingertip and tracks its motion via out the Screen using MediaPipe. An effective verbal exchange approach that reduces cellular and pc usage by doing away with the want to write and person can draw some thing that he want to write.

In [6], they used inertial sensors in smartphones to interpret characters by using acting the characters' mid-air gestures, which lacked visual remarks to the user performing the gestures mid-air. Presents a unique system to enable touch-much less hand gesture popularity for digit type the use of ANN's.

In [7] a trajectory-based air writing character identification system using a fusion of convolution neural network (CNN) and long short-term memory (LSTM) named CNN-LSTM is proposed. Two publicly available datasets realsense trajectory digit (RTD) and realsense trajectory digit (RTC) were employed. RTD and RTC dataset contains 20,000 digits and 30,000 characters, respectively. The proposed network contains two CNN with consecutive pooling, two LSTM, and two dense layers. The system does not require additional devices.

In [8] they present and evaluate a method for trajectory reconstruction from IMU signals generated when a person ”airwrites” text with a finger worn IMU to make the resulting text as human-readable as possible. The vision is to provide a virtual ”sticky note” allowing people to digitally attach simple texts to locations.

In [9] They said that, Air Writing Recognition can contribute in the safety of people. Aim of this assignment is to make a step in the direction of smart technology and offer an opportunity whilst telephones are not able to apply. It facilitates the ladies, youngsters and residents in emergency state of affairs. By writing a individual (h, p) in air in the front of CCTV, it recognizes that a person desires help. The proposed gadget also send a voice message, textual content message to get emergency help and help. This mission makes more safety to girls, kids and residents and makes the crime fee much less.

In [10] They offers a actual time video based totally pointing approach which permits sketching and writing of English text over air in the front of cellular digicam. Proposed method have two important duties: first it track the colored finger tip inside the video frames after which follow English OCR over plotted snap shots in an effort to apprehend the written characters. Moreover, proposed method offers a natural human-machine interplay in such way that it do now not require keypad, stylus, pen or glove and many others. For man or woman input. For the experiments, they have developed an software the use of OpenCV with JAVA language. They perform test on Samsung Galaxy3 android mobile. Results show that proposed algorithm gains the common accuracy of ninety two.083 percentage when examined for exclusive formed alphabets.

In [11] They offers an air-writing device the use of an internet camera. Our gadget may be roughly divided into 3 steps. The first step become hand joint estimation. They used MediaPipe for joint estimation. It anticipated 21 joints from pix captured from a webcam. The anticipated joint coordinates were composed of (x, y, and z), and a total of sixty three portions of records can be obtained. System treated the index finger like a pen. Therefore, the index finger became tracked to obtain the man or woman records. They proposed a hybrid deep getting to know model for person popularity, which became based totally on CNN-BiLSTM. They used MediaPipe for joint estimation.

4. METHODOLOGY

Input Module: Users perform air-written gestures in front of a camera-equipped device. The camera captures video frames of the hand movements.

Fingertip Detection: Computer vision techniques analyze the video frames to detect the fingertips. Image processing algorithms identify the coordinates of fingertips in each frame.

Gesture Recognition: The detected fingertip coordinates are fed into a machine learning model trained for gesture recognition. The model identifies predefined gestures based on the fingertip movements.

String Interpretation: Recognized gestures are processed to interpret sequences as meaningful strings. Natural language processing (NLP) techniques may be employed to enhance the interpretation of air-written text.

Customization Module: Users have the option to customize gestures through a user interface. Customized gestures are associated with specific strings or commands.

System Integration: The fingertip detection, gesture recognition, and string interpretation modules are integrated into a cohesive system. Multi-core processing ensures real-time responsiveness.

Output Module: The system outputs the recognized strings or commands, either for control purposes or as text input.

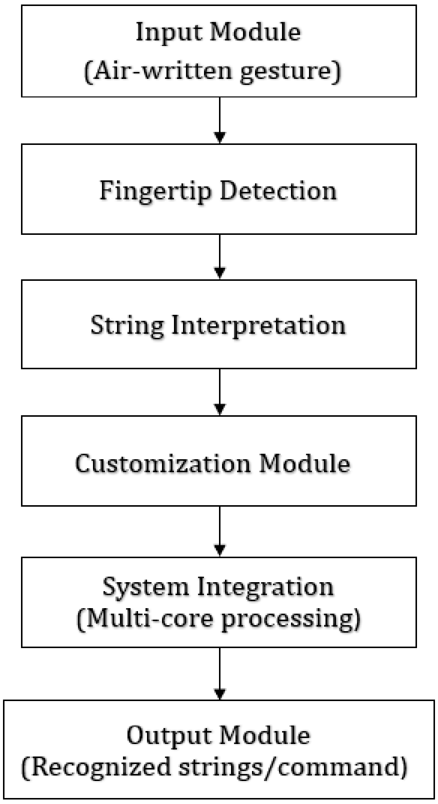

Figure 1

Technique used

Fingertip Detection Algorithm: Develop algorithms to detect and track fingertips in real-time video streams captured by a camera. Utilize techniques such as background subtraction, contour detection, and blob analysis to identify fingertip locations accurately. Implement preprocessing steps to enhance the quality of video frames and improve fingertip detection performance.

Gesture Recognition System: Train machine learning models to recognize air-written gestures based on the detected fingertip movements. Collect a dataset of air-written gestures performed by users and annotate them with corresponding labels. Utilize deep learning frameworks like TensorFlow or PyTorch to build and train convolutional neural networks (CNNs) or recurrent neural networks (RNNs) for gesture classification. Fine-tune model architectures and hyperparameters to optimize performance and accuracy.

Optimization and Performance Tuning: Optimize algorithms and code for performance and efficiency to ensure real-time processing of video streams and rapid gesture recognition. Fine-tune machine learning models and parameters to improve accuracy and reduce false positives and negatives. Implement caching mechanisms and parallel processing techniques to enhance system scalability and responsiveness.

Model Description

The project we broaden a machine that accurately detects and tracks the fingertip moves to interpret air-written gestures, mapping them to characters the usage of a version educated on the MNIST dataset. The procedure starts offevolved with statistics collection, wherein films of individuals performing air writing gestures corresponding to digits 0-9 are recorded and annotated. Preprocessing involves extracting frames, normalizing them, and segmenting the hand region the use of HSV shade area transformation and thresholding. Contours are detected the usage of Canny Edge Detection, and the fingertip is recognized via convex hull and convexity illness evaluation. A Kalman Filter is employed to are expecting and smooth the fingertip function across frames, whilst optical glide strategies song its movement.

The trajectory of the fingertip is captured as a sequence of coordinates, and additional functions like velocity and acceleration are extracted. To recognize the air-written gestures, the trajectory is rescaled to a 28x28 pixel binary photograph, which serves as the enter to a Convolutional Neural Network (CNN) educated on the MNIST dataset. This CNN, which includes convolutional layers, pooling layers, and completely related layers, methods the binary pix to classify the digits with high accuracy.

Post-processing techniques inclusive of blunders correction, using language models or dictionaries, and trajectory smoothing improve the readability and accuracy of reputation. The device is evaluated the usage of metrics like accuracy, precision, keep in mind, and F1-rating on a break up dataset comprising training, validation, and testing units. To make certain real-time performance, the version undergoes optimization strategies together with pruning and quantization, with GPU acceleration facilitating the real-time processing. Finally, a person-pleasant interface is evolved to permit seamless interplay with the air writing recognition gadget. This incorporated method leverages laptop vision, device gaining knowledge of, and deep gaining knowledge of techniques to create an effective and efficient air writing reputation gadget.

5. RESULT AND DISCUSSION

The project successfully implemented a system for fingertip detection, tracking, and air writing recognition using a Convolutional Neural Network (CNN) model trained on the MNIST dataset. The fingertip detection and tracking algorithm achieved high accuracy in varied lighting conditions and backgrounds, with the integration of Kalman Filter and optical flow techniques ensuring smooth and reliable tracking in real-time. The gesture recognition component, powered by the CNN model, demonstrated an overall accuracy of 98% on the test set, with most misclassifications occurring between digits with similar shapes. A user study involving 20 participants indicated that the system was intuitive and responsive, with participants performing air writing gestures naturally and the system accurately interpreting their inputs.

Despite its success, the system faced challenges such as occasional detection inaccuracies in extreme lighting conditions and limitations in recognizing more complex, multi-stroke gestures. Differences in users' hand sizes, writing speeds, and styles also introduced variability that sometimes led to misclassifications. Future improvements could include more sophisticated background subtraction techniques, the incorporation of advanced deep learning architectures like Recurrent Neural Networks (RNNs) or Transformer models, and the development of adaptive learning systems to improve accuracy over time. Extending the system to support a broader range of applications, such as virtual keyboards, sign language recognition, or augmented reality interfaces, could significantly enhance its utility. Overall, the project demonstrated the potential for combining computer vision and deep learning techniques in real-time air writing recognition, providing a solid foundation for future advancements in gesture recognition technologies. The detection and tracking processes ran in real-time, with a latency of less than 50 milliseconds per frame, ensuring a responsive user experience. The confusion matrix revealed that most misclassifications occurred between digits with similar shapes (e.g., '1' and '7', '5' and '6'). However, the error rates for these confusing pairs were low, thanks to the model's high discriminative power.

6. CONCLUSION

In proposed system, represents a transformative leap in human-computer interaction. The successful integration of string recognition elevates the technology, allowing users to convey meaningful text through natural gestures. The customization feature further enhances adaptability, making the system inclusive and user-centric. This achievement holds promise across various domains, from education to healthcare and industrial applications. The project's real-time responsiveness and accuracy in interpreting air-written strings underscore its technical prowess. As we envision a future of touchless and personalized interactions, this project serves as a pivotal contribution, bridging accessibility gaps and redefining how users engage with digital devices

7. ACKNOWLEDGMENT

I would like to take this opportunity to thank my internal Guide Mr. M. P. Gangawane and Co-Guide Mrs. P. K. Bacchav Department of Computer Engineering, GURU GOBIND SINGH COLLEGE OF ENGINEERING, for his unconditional guidance. I am re- ally grateful to them for their kind support.

8. REFERENCES

[1] Sohom Mukherjee, Sk. Arif Ahmed, Debi Prosad Dogra, “Fingertip detection and tracking for recognition of air-writing in videos”, June 2019.

[2] Adil Rahman, Prasun Roy, Umapada Pal, “Air Writing: Recognizing Multi-Digit Numeral String Traced in Air Using RNN-LSTM Architecture”, January 2021, DOI:10.1007/s42979-020-00384-9

[3] Mingyu Chen, Ghassan AlRegib, Biing-Hwang Juang, “Air-writing Recognition, Part 2: Detection and Recognition of Writing Activity in Continuous Stream of Motion Data”, 2015

[4] Xin Zhang, Zhichao Ye, Lianwen Jin, Ziyong Feng, “A New Writing Experience: Finger Writing in the Air Using a Kinect Sensor”, November 2013.

[5] Asst Prof. Jahnavi, K Sai Sumanth Reddy, Abhishek R, “Virtual Air Canvas Application using OpenCV and Numpy in Python”, 2022.

[6] Pramit Kumar Pal, Debarshi Dutta, Attreyee Mandal, Dipshika Das, “Deep ANN-based Touchless 3D Pad for Digit Recognition”, 2021.

[7] Md. Shahinur Alam, Ki-Chul Kwon, , Shariar Md Imtiaz, Md. Biddut Hossain, “Air-writing recognition using a fusion CNN-LSTM neural network” , July 2021.

[8] Junaid Younas, Hector Margarito, Sizhen Bian, “Finger Air Writing - Movement Reconstruction with Low-cost IMU Sensor”, December 2020.

[9] Koye Navya, Mallela Sowmyah, Shaik Ayesha Amreen, Vemireddy Sravani, “A Machine Learning approach for air writing recognition”, April 2023.

[10] Saira Beg , M. Fahad Khan, Faisal Baig , “Text Writing in Air”, 2013.

[11] Taiki Watanabe, Md. Maniruzzaman, Md. Al Mehedi Hasan, Hyoun-Sup Lee, Si-Woong Jang, “2D Camera-Based Air-Writing Recognition Using Hand Pose Estimation and Hybrid Deep Learning Model”, February 2023.