Journal of Advanced Research in Image Processing and Applications
Volume 6, Print Issue 1 - 2023, Pg. No. 19-24
Peer Reviewed Journal

Review Article

Webcam Paint Application Using Python

Prashansa Jaiswal

Prashansa Jaiswal CSE Department, Chandigarh University Lucknow, Uttar Pradesh, India.

I N F O

E-mail Id:
19bcs1739@cuchd.in
Orcid Id:
https://orcid.org/0009-0001-2240-7034
How to cite this article:
Jaiswal P. Webcam Paint Application Using Python *J Adv Res Image Proc Appl* 2023; 6(1): 19-24.

Date of Submission: 2023-03-20
Date of Acceptance: 2023-04-13

A B S T R A C T

Writing is a cohesive form of communication that can effectively convey our thoughts. Typing and writing are the standard methods of recording information today. Characters or words are written in the free space with a marker or finger. It differs from traditional writing methods in that the pen does not move up and down. With the development of intelligent wearable devices, the digital world can now be controlled by human gestures. These wearable devices can recognize and understand our actions.

Recognizing and interpreting a continuous sequential gesture stream from the given set of input data is gesture recognition. Gestures are non-verbal information used to improve computer language understanding. Human gestures are perceived by vision, computer vision is used to analyze various gestures.

The project takes advantage of this gap and focuses on developing a motion-to-text converter that can potentially serve as a software for intelligent wearable devices for writing from the air. The system will use computer vision to trace the finger's path, in that way, one can write from above. The generated text can also be used for various purposes, such as sending messages, e-mails, etc. It will be a powerful means of communication for the deaf. It is an effective communication method that reduces cellphone and laptop usage by eliminating the need to write.

Keywords: Air Writing, Character Recognition, Object Detection, Real-Time Gesture Control System, Smart Wearables

Introduction

Since its inception, the computer has had a significant impact on human life. Humans need to interact with computers for various reasons, from the daily chore of keeping records to performing complex scientific calculations. Also, a sudden increase in the Internet of Things has led to a rise in computer use.

The art of writing was first discovered around 2000 BC. Neolithic people began to write on the cave wall. These walls were soon replaced by stone, replaced by cloth, today we report on paper. We are gradually moving towards a more digital version based primarily on electronic text documents with recent technological advances. With the help of QWERTY keyboards, we can create and edit these digital text documents. This pair of electronic keyboard and text is slowly replacing traditional pen and paper writing.

Additionally, touchscreen devices have made QWERTY keyboards more portable. However, today we think having a cell phone or tablet to take notes is boring. Additionally, holding a phone in one hand prevents the user from interacting with the physical world.

With the advent of virtual and augmented reality, the need to develop natural human-machine interaction (HCI) systems to replace traditional HCI approaches is growing rapidly. In particular, interfaces involving interaction based on hand gestures have gained popularity in many application areas, viz. Automotive interfaces (Ohn-Bar and Trivedi, 2014), human activity recognition (Rohrbach et al., 2016), various advanced approaches to hand gesture recognition (Molchanov et al., 2015) have been developed; Rautaray and Agrawal, 2015). However, hand gestures as such are not sufficient for entering text. This requires the development of non-contact air writing systems that can replace tactile and electromechanical input fields, leading to a more natural human-computer interaction (HCI) approach. A vision-based system for recognizing handwriting in the air is not a new problem, extensive work has been done over the past two decades. One of the first works by Oka et al. (Oka et al., 2002) used a sophisticated device with an infrared and color sensor to track and detect trajectories of simple geometric shapes with their fingertips. In (Amma et al., 2012), inertial sensors with a glove were used to detect and recognize handwriting from the air continuously. Recently, Misra et al. (Misra et al., 2017) developed a hand gesture recognition framework capable of recognizing letters, numbers, arithmetic operators, 18 printable ASCII characters using a red marker on the tip of the index finger to identify fingertips. Despite satisfactory performance in terms of path recognition accuracy, previous approaches using cumbersome motion detection and detection hardware acknowledge numerous limitations in user behavior. For example, wearing protective gloves can disrupt the user's natural handwriting pattern and is often viewed by many users as an unwanted burden.

This system eliminates the need to carry cell phones to take notes. Instead, portable devices are needed daily to enable writing in the air. Finger and movement recognition techniques and aerial character recognition are used to recognize the drawn word. With the help of OpenCV and CNN, the fingertip is identified first. The trajectory of this fingertip is then recorded and drawn on a black background. Our aerial photography model is not a handwriting model. It can then be turned into a handwriting recognition model.

Problem Statement

The existing system works with your fingers only, there are no highlighters, paints, or relatives. Identifying and characterizing an object such as a finger from an RGB image without a depth sensor is a great challenge. Another problem is the lack of up and down movement of the pen. The system uses a single RGB camera to write from above. Since depth detection is not possible, the up and down activities of the pen cannot be tracked. Therefore, the entire path of the fingertip is drawn, the resulting image would be absurd and not recognized by the model. Using real-time hand gestures to change the system from one state to another requires much code care. In addition, the user must know many movements to control his plan adequately. The project focuses on solving some critical social problems.

First, people with hearing problems have many problems in daily life. Although listening and listening are taken for granted, people with this disability communicate using sign language. Most countries in the world cannot understand your feelings and emotions without an intermediate translator. Second, the excessive use of smartphones: causes accidents, depression, distractions, other illnesses that humans can still discover. Although its portability and ease of use are deeply admired, its drawbacks include life- threatening events. Waste of paper is not uncommon. A lot of paper is wasted scribbling, writing, drawing, etc. The production of A4 paper requires an average of about 5 liters of water. 93% of sources are from trees, 50% of commercial waste is paper, 25% of landfills are paper, the list goes on. Wasted paper damages the environment through the use of water and trees and generates tons of waste. Air writing can solve these problems quickly. It will serve as a communication tool for the deaf. Your text written in the air can be presented in AR or converted to speech. One can write on-air quickly and keep working without too many distractions.

Also, writing in the air requires no paper. Everything is archived electronically.

Literature Survey

Automatic object tracking has many uses in computing, such as Computer vision and human- machine interaction.[1-2] Various applications of tracking algorithms are suggested in the literature. One group of researchers used it to translate sign languages,[5-9] others for hand gesture recognition,[6-7] another group for text tracking and recognition,[8,18] to monitor the body movement of objects for virtual reality[10] and character recognition based on finger tracking,[19-21] etc. Bragatto et al. developed a method that automatically translates Brazilian Sign Language from video input. They used a multilayer NN (Neural Perceptron) network with a piecewise linear approximate trigger function for real-time video processing. This activation function reduces the mean complexity time of NN. In addition, they used NN in two stages: color recognition and steps to assess hand posture. Their results show that the proposed method works well with a detection rate of 99.2%.[5] Cooper also introduced a method for managing 3D cell bioprinting that is more complex than the generalized set. Cooper developed a technique that reduces tracing by identifying errors in his thesis's classification and tracing processes. Cooper used two pretreatment steps; One is for movement, the other is used to determine the shape of the hand. He also used the screen to expand his vocabulary. The viseme is an essential position of the mouth and face in the pronunciation of a phoneme and the visual representation of phonemes. Over time, he develops a poorly structured learning method that identifies characters.

Araga et al. proposed a hand gesture recognition system using Jordan's Recurrent Neural Network work even with poor segmentation. They combined both the segmentation process and recognition using a cross-cluster process. Their results show better performance with a performance loss of 5% for both models. Neumann et al. developed a method to find and recognize text in authentic images. In their article, they used a hypothesis framework that can handle multiple lines of text. They also use synthetic characters to train the algorithm,, finally, they use maximum stable areas (MSER), which offer robustness to geometric and lighting conditions.[8]

Furthermore, Wang et al. discussed the color-based motion detection system for indoor and outdoor environments. In the proposed method, they used a webcam and a colored T-shirt to track the object. The result of the proposed method shows that the proposed method can be used for virtual reality applications. Jari Hannuksela et al.[19] Toshio Asano et al.[20] Sharad Vikram et al.[21] have character recognition systems based on finger tracing. In,[19] the author presents a motion-based tracking algorithm that combines two Kalman filtering techniques and Expectation Maximization (EM) methods to estimate two different motions. Finger and camera movement. The estimate is based on the motion properties that the scene calculates for each image. Its main idea is to control mobile devices by simply swiping a finger in front of a camera. In,[20] the authors discuss a visual interface that recognizes Japanese katakana characters in the air. To follow the movement of the hand, they used the LED pen and a camera. They convert the signal of the pencil into direction codes. Codes are normalized to 100 data items to eliminate the effect of typing speed, where 46 Japanese characters are defined. For a single camera, they achieve a character recognition accuracy of 92.9%, for multiple cameras, they accomplish a gesture- directing accuracy of 9°. (JRNN). Their system modeled the 5 and 9 different IV.

Approach

Hand positions through a sequence of representative static images. He then takes the footage as input and begins to classify the hand poses. JRNN finds the input gesture after the temporal behavior of the posture sequence has been recognized. In addition, they are developing a new training method. The proposed method shows a precision of 99.0% for five different hand positions, while it achieves an accuracy of 94.3% for nine gestures.[6] In,[7] Yang et al.; discussed an alternative solution to matching a sequence of images to a pattern, which usually occurs with hand gesture recognition. The proposed method is not based on skin color patterns and can Writing Hand Pose Detection Recognizing the position of the writing hand and recognizing it through other gestures is an essential step in initializing aerial writing. Unlike traditional writing, when the pen moves down, the pen moves up, writing in the air is not outlined as a writing sequence. Events. The system recognizes the position of a writing hand and distinguishes it from a non-writing hand by counting the number of raised fingers.

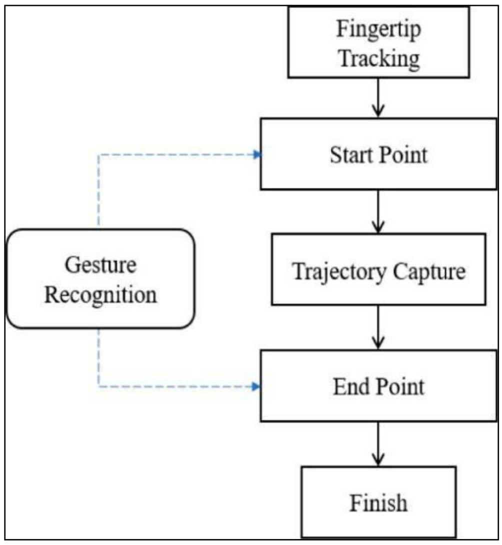

Figure 1.Flow chart of Proposed System

Hand Region Segmentation

Once we have accurately captured the hand using the above technique, the segmentation of the hand area is done using a two-step approach, viz. The skin segmentation and the subtraction of the background and the final binary image of the hand are obtained as an aggregation of the two. The proposed algorithm works well in real-time and provides relatively accurate segmentation results. Although skin colors vary greatly from breed to breed, it has been observed that skin color has a small area between different skin types, while skin luminosity differs significantly.

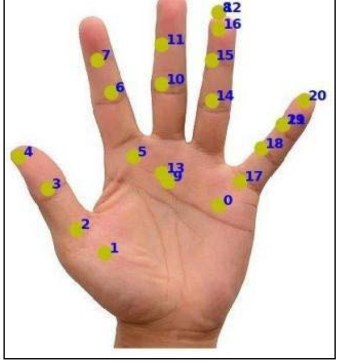

Figure 2.Hand Region Segmentation

Background Subtraction

Since accurate hand detection with the Faster R-CNN handheld detector followed by filtering of skin color at the boundary of the candidate's hand provides a reasonably good segmentation result, the subtraction step of background is only used to remove skin-colored objects (not part of the hand) that are in the bounding box of the recognized hand may be present.

Hand Centroid Localization

Since determining the exact center of gravity of the hand is critical in the following steps, the system uses two algorithms to determine the initial estimates of the center of gravity, the final center of gravity is calculated as the average of the two. The distance transformation method is used to obtain the first estimate of the center of gravity (xc1, yc1). In the distance transform image, each pixel is represented by its distance from the next edge pixel. Euclidean distance was used to measure the distance between a pixel and its closest edge pixel. Therefore, the pixel with the highest intensity in the distance transform image is taken as the center of gravity.

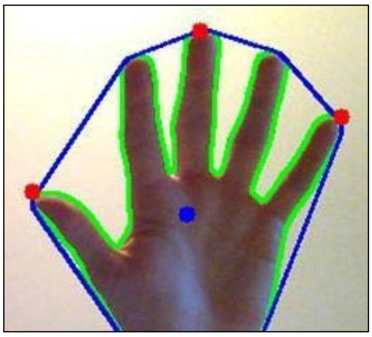

Figure 3.Hand Centroid Localization

Fingertip Tracking

Hand detection and tracking on consecutive images is an essential part of fingertip detection and tracking. Experiments show that using the Faster R- CNN handheld detector for each frame is computationally intensive and produces frame rates well below real-time performance. Therefore, the KCF tracking algorithm (Henriques et al., 2015) is used to track the area of the captured hand.

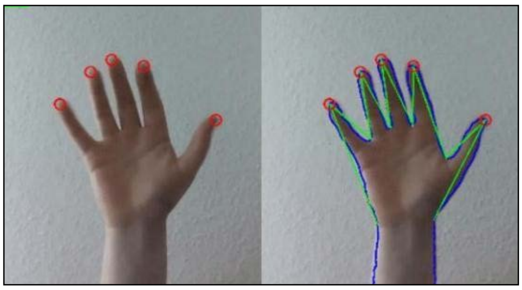

Figure 4.Fingertip Tracking

Colour Tracking of Object at fingertip

Incoming webcam images must be converted to HSV color space to recognize hand-colored objects.

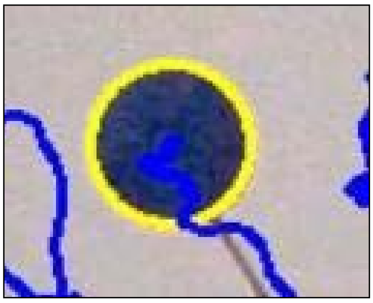

Figure 5.ColourTracking of Object at Finger tip

The algorithm code converts the incoming image into HSV space, which is a very suitable color space and perfect for color tracking.

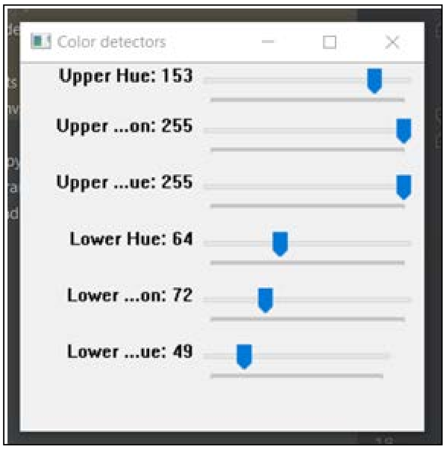

Figure 6.Colour Trackbar

The tracking bars have organized the HSV values into the required color range of the colored object you placed on your finger.

Contour Detection of the Mask of Colour Object

After detecting the Mask in Air Canvas, it is now time to locate its center position for drawing the Line. The system will perform some morphological operations on the Mask to make it free of impurities and detect contour easily.

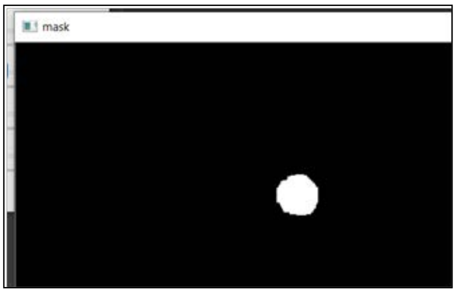

Figure 7.Mask Detection

Drawing the Line using the position of the contour

The actual logic behind this computer vision project is to create a Python deque (a data structure). The deque will memorize the position of the outline in each subsequent frame, we will use these accumulated points to create a line using OpenCV's drawing capabilities.

Now use the outline position to decide whether to click a button or draw on the provided sheet. Some of the buttons are at the top of the canvas. When the pointer enters this area, it is activated according to the method present in this area.

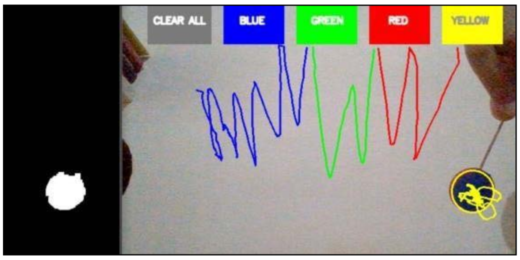

Figure 8.Drawing Line using the position of the contour

Object Position: Image extracted from the video sequence. Extract the color image from the reference: This proposed method tracks the index finger movement, colored blue. We don't have a reference image, so any previous image is a reference to the next. Now take the difference in the images and extract the color and movement of the object.

Edge Enhancement (EE): Edge enhancement technique makes the object location algorithm robust against noise, different lighting conditions, darkening, fading of objects, even in low contrast images.

Algorithm of Workflow

The most exciting part of this system is the workflow phase. Writing involves a lot of functionalities. So, the number of gestures used for controlling this particular system is equal to these number of actions involved. The basic functionalities that are included in this system are:

Writing Mode: In this state, the system willtrace the fingertip coordinates and stores them.

Colour Mode: The user can change the color of the text among the various available colors.

Backspace: Say if the user goes wrong, we need a gesture to add a quick backspace.

Results

Based on the algorithm defined in the upper section, we will write a code to take input for the system via webcam, once the color tracking of the object held by the user is done, the designed system can be used.

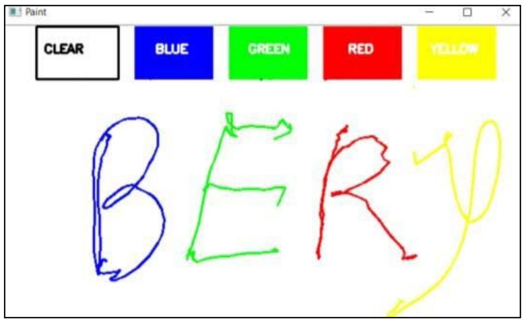

Figure 9.System Output

Project Scope

The scope of this system is mainly used as a powerful means of communication for the deaf, which means implementing this project can help in:

1. An effective communication method that reduces mobile and laptop usage by eliminating the need to write.
2. It helps people with hearing impairments to communicate well.
3. Various purposes, such as sending messages, e-mails, etc., as the generated text can also be used for that.

Conclusion

This system has the potential to challenge traditional writing methods. It eradicates the need to carry a mobile phone in hand to jot down notes, providing a simple on-the-go way to do the same. It will also serve a great purpose in helping especially abled people communicate easily. Even senior citizens or people who find it difficult to use keyboards can effortlessly use the system. Extending the functionality, this system can also be used to control IoT devices shortly. Drawing in the air can also be made possible. This system will be an excellent software for smart wearables using which people could better interact with the digital world. Augmented reality can make text come alive. Air- writing systems should only obey their master's control gestures and should not be misled by people around. Upcoming object detection algorithms such as YOLO v3 can improve fingertip recognition accuracy and speed. In the future, advancements in Artificial Intelligence will enhance the efficiency of air-writing.

References

1. Huang Y, Liu X, Zhang X. A Pointing Gesture Based Egocentric Interaction System: Dataset, Approach, Application," 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Las Vegas, NV, 2016; 370-377.
2. Ramasamy P, Prabhu G, Srinivasan R. An economical air writing system is converting finger movements to text using a web camera *International Conference*

on Recent Trends in Information Technology (ICRTIT) Chennai 2016; 1- 6: 2016.
3. Saira Beg, M. Fahad Khan and Faisal Baig, "Text Writing in Air Journal of Information Display Volume 14, Issue 4, 2013
4. Alper Yilmaz, Omar Javed, Mubarak Shah, "Object Tracking: A Survey", ACM Computer Survey. Vol. 38, Issue. 4, Article 13, Pp. 1-45, 2006
5. Yuan-Hsiang Chang, Chen-Ming Chang, "Automatic Hand-Pose Trajectory Tracking System Using Video Sequences", INTECH, pp. 132- 152, Croatia, 2010
6. Erik B. Sudderth, Michael I. Mandel, William T. Freeman, Alan S. Willsky, "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, Pp. 1-9, 2004
7. T. Grossman, R. Balakrishnan, G. Kurtenbach, G. Fitzmaurice, A. Khan, B. Buxton, "Creating Principal 3D Curves with Digital Tape Drawing," Proc. Conf. Human Factors Computing Systems (CHI' 02), pp. 121-128, 2002.
8. T. A. C. Bragatto, G. I. S. Ruas, M. V. Lamar, "Real-time Video-Based Finger Spelling Recognition System Using Low Computational Complexity Artificial Neural Networks", IEEE ITS, pp. 393-397, 2006
9. Yusuke Araga, Makoto Shirabayashi, Keishi Kaida, Hiroomi Hikawa, "Real Time Gesture Recognition System Using Posture Classifier and Jordan Recurrent Neural Network", IEEE World Congress on Computational Intelligence, Brisbane, Australia, 2012
10. Ruiduo Yang, Sudeep Sarkar, "Coupled grouping and matching for sign and gesture recognition", Computer Vision and Image Understanding, Elsevier, 2008
11. R. Wang, S. Paris, J. Popovic, "6D hands: markerless hand-tracking for computer-aided design," in Proc. 24th Ann. ACM Symp. User Interface Softw. Technol., 2011, pp. 549–558.
12. Maryam Khosravi Nahouji, "2D Finger Motion Tracking, Implementation For Android Based Smartphones", Master's Thesis, CHALMERS Applied Information Technology,2012, pp 1-48
13. EshedOhn-Bar, Mohan ManubhaiTrivedi, "Hand Gesture Recognition In Real-Time For Automotive Interfaces," IEEE Transactions on Intelligent Transportation Systems, VOL. 15, NO. 6, December 2014, pp 2368-2377
14. P. Ramasamy, G. Prabhu, R. Srinivasan, "An economical air writing system is converting finger movements to text using a web camera," 2016 International Conference on Recent Trends in Information Technology (ICRTIT), Chennai, 2016, pp. 1-6.
15. Kenji Oka, Yoichi Sato, Hideki Koike, "Real-Time Fingertip Tracking and Gesture Recognition," IEEE Computer Graphics and Applications, 2002, pp.64-71.
16. H.M. Cooper, "Sign Language Recognition: Generalising to More Complex Corpora", Ph.D. Thesis, Centre for Vision, Speech and Signal Processing Faculty of Engineering and Physical Sciences, University of Surrey, UK, 2012
17. Y. Huang, X. Liu, X. Zhang, L. Jin, "A Pointing Gesture Based Egocentric Interaction System: Dataset, Approach, Application," 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Las Vegas, NV, pp. 370-377, 2016
18. Vladimir I. Pavlovic, Rajeev Sharma, Thomas S. Huang, "Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review," IEEE Transactions on Pattern Analysis and Machine Intelligence, VOL. 19, NO. 7, JULY 1997, pp.677-695
19. Guo-Zhen Wang, Yi-Pai Huang, Tian-Sheeran Chang, Tsu-Han Chen, "Bare Finger 3D Air- Touch System Using an Embedded Optical Sensor Array for Mobile Displays", Journal Of Display Technology, VOL. 10, NO. 1, JANUARY 2014, pp.13-18
20. Napa Sae-Bae, Kowsar Ahmed, Katherine Isbister, NasirMemon, "Biometric-rich gestures: a novel approach to authentication on multi-touch devices," Proc. SIGCHI Conference on Human Factors in Computing System,2005, pp.977-986
21. W. Makela, "Working 3D Meshes and Particles with Finger Tips, towards an Immersive Artists' Interface," Proc. IEEE Virtual Reality Workshop, pp. 77-80, 2005.
22. A.D. Gregory, S.A. Ehmann, M.C. Lin, "inTouch: Interactive Multiresolution Modeling and 3D Painting with a Haptic Interface," Proc. IEEE Virtual Reality (VR' 02), pp. 45-52, 2000.
23. W. C. Westerman, H. Lamiraux, M. E. Dreisbach, "Swipe gestures for touch screen keyboards," Nov. 15 2011, US Patent 8,059,101
24. S. Vikram, L. Li, S. Russell, "Handwriting and gestures in the air, recognizing on the fly," in Proceedings of the CHI, vol. 13, 2013, pp. 1179– 1184.
25. X. Liu, Y. Huang, X. Zhang, L. Jin. "Fingertip in the eye: A cascaded CNN pipeline for the real-time fingertip detection in egocentric videos," CoRR, abs/1511.02282, 2015.