Smart Drawing for Online Teaching

D. T. D. M. DAHANAYAKA, A. R. LOKUGE, J.A.D.E. JAYAKODY,
I. U. ATTHANAYAKE

(*Mechanical Engineering*, *Department of Mechanical Engineering*, *The Open University of Sri Lanka*, *Nawala*, *Nugegoda*)

Abstract: Teaching and learning-related activities have embraced digital technology especially, with the current global pandemic restrictions pervaded during the last two years. On the other hand, most of the academic and professional presentations are conducted using online platforms. But the presenter-audience interaction is hindered to a certain extent in an online case in contrary to face-to-face where real-time writing is beneficial when sketching is involved. The use of digital pens and pads is a solution for such instances, though the cost of acquiring such hardware is high. Economical solutions are essential as affordability is concerned. In this study, a real-time user-friendly, innovative drawing system is developed to address the issues related to the problems confronted with online presentations. This paper presents the development of an algorithm using Hand Landmark Detection, which replaces chalks, markers and regular ballpoint pens used in conventional communication and presentation, with online platforms. The proposed application in this study is acquired by Python and OpenCV libraries. The letters or the sketches drawn in the air were taken straight to the computer screen by this algorithm. The proposed algorithm continuously identifies the Hand Landmark using the images fed by the web camera, and text or drawing patterns are displayed on the screen according to the movements made by Hand Landmark on the image space. The developed user interface is also user-friendly. Hence the communication of letters and sketches were enabled. Although the concept has been developed and tested, with further research the versatility and accuracy of communication can be enhanced.

Keywords: Hand Landmark Detection, Off-line Recognition, Online Teaching, Smart Drawing

1 Introduction

Teaching and learning using digital platforms, or online teaching and learning, had become popular more than ever before with the restrictions imposed to conducting physical meetings and interaction imposed to curb the pervasion of the covid-19 pandemic through the world. Besides the flexibility of learning and teaching over time, live communication online saves transportation costs, time, and other advantages. The challenge with online teaching and learning is, inability to maintaining eye contact between the presenter and the students, tacking the hand guestures that support presentation which leads to less attention during the presentation or education. Real-time writing to explain the lesson's content is one of the methods that has been practiced by teachers habitually in face-to-face teaching and learning. Various tools have been developed in digital platforms to suffice the lack of the physical presence of the presenter and to maintain their virtual movements, such as whiteboard marker in the well-known ZOOM platform. But manipulations of most of these tools need reasonable control. Besides, the digital touch pens and the related accessories are costly, though they serve real-time writing. Many people cannot afford such expensive products. Also the chemical wastes from chalks and marker pens cause Environmental issues, while usingconventional pens and papers are consumes money too[1-2]. One of the primary objectives of the present study is to develop a zero-budget real-time writing mechanism for online classes, where no hassle to the user and traditional writing skills only become beneficial when using the technology. With the dawn of intelligent equipment, there has been a more significant motivation on hand gesture recognition systems those have the ability to interact efficiently with the machine via human-computer interaction. There are different types of methods in hand gesture recognition. For example, find contours, convexity hull or train a Convolution Neural Network (CNN). But it requires special conditions when using these methods like a suitable background and light condition used in training or similar situations. Therefore, recognize hand gestures using landmark features is the most acceptable solution.

2 Literature Review

Hand Gesture Recognition keeps gaining attention in recent years due to its efficient cooperation with the machine through various applications. The Literature review suggests diverse methods to distinguish the hand gesture, leading to diverse applications. Hand Gesture Recognition provides a revolutionized alternative to the complex interfaces of human-computer interaction. Two approaches are commonly used for Hand Gesture Recognition for human-computer interaction. They are sensor-based glove innovation and computer vision-based approaches [3-5]. Sensor-based access requires sensors and equipment to detect hand movement and position [6, 7], which leads to several disadvantages. They are the high cost, complexity, and difficulty of maintenance. Fig. 1 shows two approaches to detect Hand Gestures.A Block diagram for a Sensor-based system is shown in Fig.1a, and a circuit diagram for the sign language translator device is shown in Fig.1b.

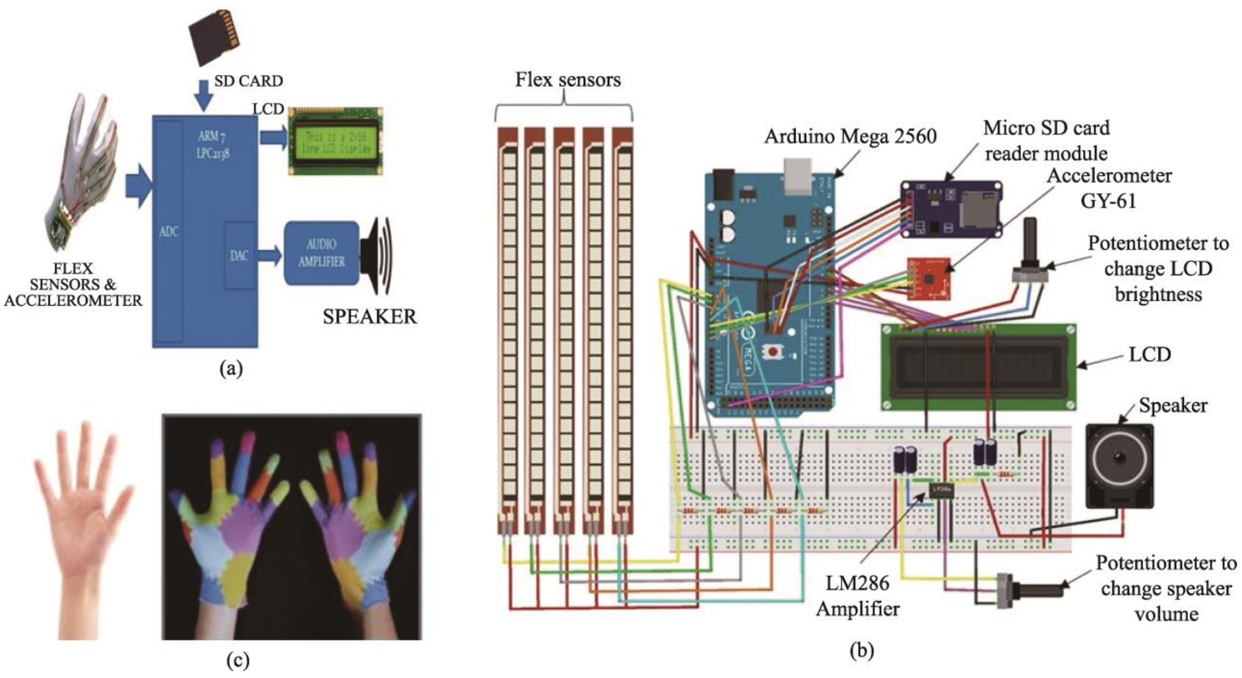

Fig.1 Approaches to Detect Hand Gestures (a) Block diagram for Sensor-based system(b) Circuit Diagram for Sign Language Translator Device(c) Skin Colour-based and Color Glove-based Hand Gesture Recognition (from Left to Right)

The vision-based approach is a stimulating research area, which involves computer vision, image processing and machine learning, etc., and this can be used in various engineering applications. The advantages of the Vision-based approach are simple and fast operation, high reliability, better precision of the final output, a wider range of use, and low cost. Gesture recognition typically has skin colour-based hand gesture recognition and colour glove-based hand gesture recognition, which were used to recognize the hand gesture [8]. The skin colour-based and glove-based hand gestures recognition techniques are shown in Fig.1c. Gesture recognition consists of important stages. They are Detection and Segmentation, Tracking, and Classification. The gray-scale, RGB, HSV colour models were used to classify colour surfaces. If the standard RGB colour space is used, colour detection can be complicated under variable lighting and contrast levels[9]. Therefore, an HSV colour model which is insensitive to lighting is used. The advantages of HSV over the RGB are separated colours information. HSV color space is secure to represent a color than RGB [10-13]. Once the scanning of the webcam feed was started, the RGB colours were converted into HSV colours. After that, the given colours are identified by the proposed algorithm in the colour space. Then it detects a certain points of the colour surface [14]. In OpenCV, there are several color space conversions, RGB ↔ GRAY, RGB ↔ CIE, RGB ↔ YCrCb, RGB ↔ HSV, RGB ↔ HSL etc. Dissimilar to the RGB colour model, the HSV model separates the image intensity from the colour information, convenient for separating colour components from the intensity. For HSV, hue range is [0,179], Saturation range is [0,255] and Value range is [0,255]. The range of the HSV values between low and high selected will be used to create a mask around the colour profile to be detected [15]. Then Bitwise-AND-Mask was used to display filtered colour. An example for the HSV colour mode application is shown in Fig.2. HSV Cylinder is shown in Fig.2a.The original frame is shown in Fig.2b and the filtered frame using the HSV colour model and Bitwise-AND mask is shown in Fig.2c.

The constant tracking of hand gestures is a complicated exercise. Therefore, numerous Algorithms were initiated to gain improved results, for example, Optical Flow, Cam shift, Particle Filtering, etc. Classification is the concluding step for Gesture identification. The different approaches for gesture recognition can be roughly divided into two categories based on the techniques. They are the mathematical-based models and the soft-computer-based. Hidden Markov Model (HMM) and the Finite State Machine (FSM) are examples of Mathematical based models, and Artificial Neural Networks (ANN) are soft computer-based methods [16- 19]. Generally, ANN is used for classification. [20- 22]. This is the core of deep learning algorithms, which is a subset of machine learning. With a Higher number of required neurons to function, the computer becomes inefficient, where the Convolution Neural Network (CNN) becomes dominant with better performance [23, 24]. The algorithms like

(a) (b) (c)

Fig.2 HSV Colour Mode Application (a) HSV Cylinder (b) Original Frame (c) Filtered Frame Using HSV Colour Model and Bitwise-AND Mask

You Only Look Once (YOLO), and Single Shot Detector (SSD) can be used to detect hand gestures. [25-28]. The architecture of the neural network for hand gesture recognition is shown in Fig.3.

There have been several approaches for Hand Landmarks. Most of the current technology is limited to recognizing the fingertips [29]. A hand consists of 27 bones, including 14 phalanges. Hand Landmark is shown in Fig.4.

Feasible results can be obtained via trained CNN, contour, or convex hull train with less difficulties. Yet special conditions were required, for example, the appropriate background or similar conditions used in training. In this study, MediaPipe, which is a reliable hand gesture recognition solution. was used to recognize the Hand Landmark [30]. It finds the keypoint features of hands from both images and videos. MediaPipe was first introduced in June 2019 to provide a combined computer vision and some machine learning features. It effectively manages resources CPU and GPU to achieve low latency performance and handles audio and video frames like synchronization of time

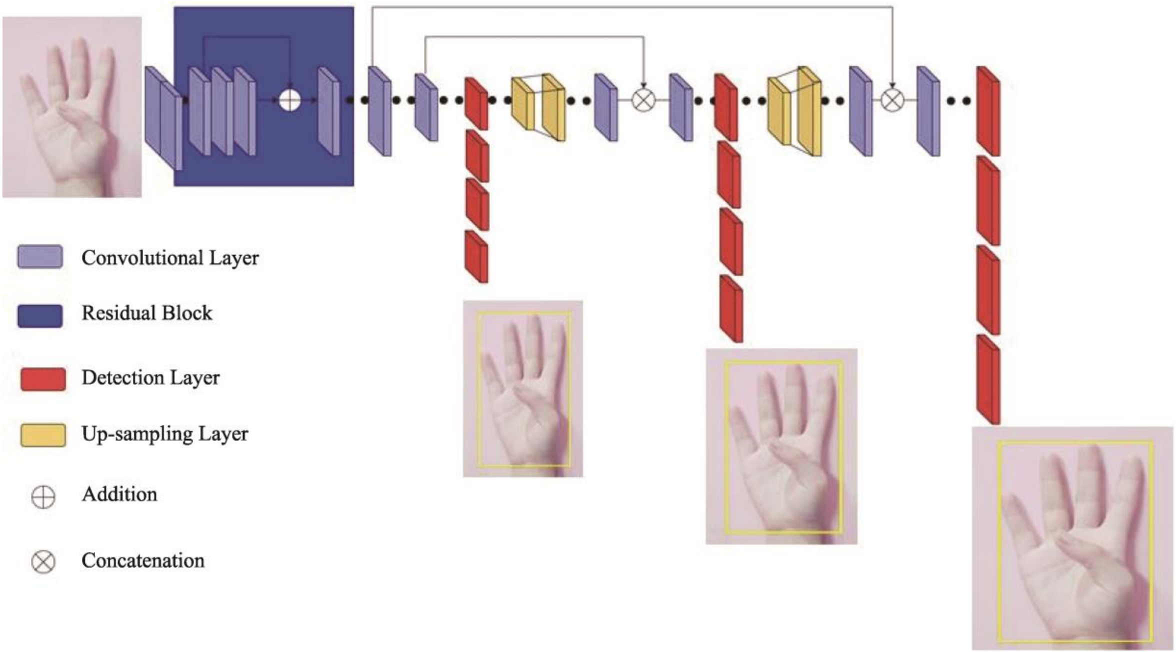

Fig.3 The Architecture of the Neural Network for Hand Gesture Recognition

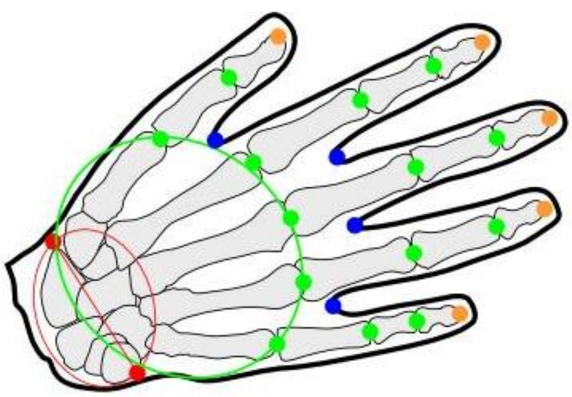

Fig.4 Hand Landmark

series data. When considering the architecture, this consists of two models working together. They are Palm Detection Model and Hand Landmark Model. First, a palm detector is trained instead of a hand detector. Then, an encoder-decoder feature extractor is used for more significant scene-context responsiveness, even for small objects. Subsequently, the focal loss is minimized during training to support many anchors resulting from the high-scale modification. Palm detector architecture is shown in Fig.5.

After activating the palm recognition throughout the whole image, it detects hand regions, and the hand landmark model performs precise landmark localization of 21 coordinates. The model has three outputs. The first one is x, y, and relative depth 21 hand landmarks, whilst the second one is the probability of hand presence in the input image indication, and the third one is left or right-hand classification of the hand. Hand landmark model architecture is shown in Fig.6.

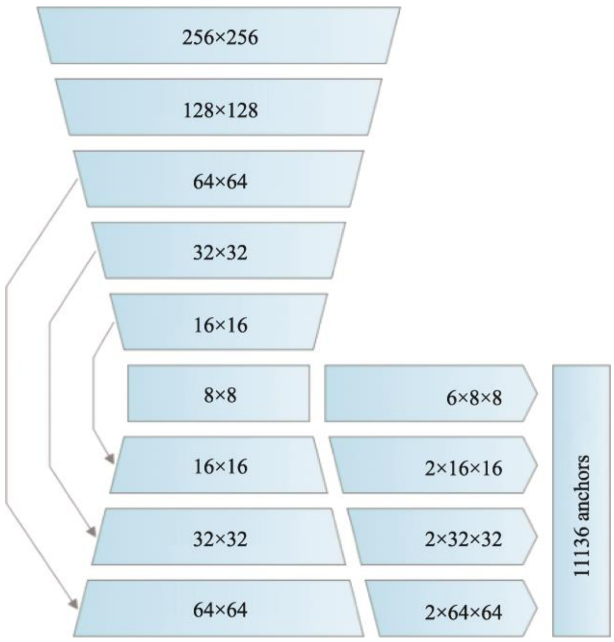

Fig.5 Palm Detector Model Architecture

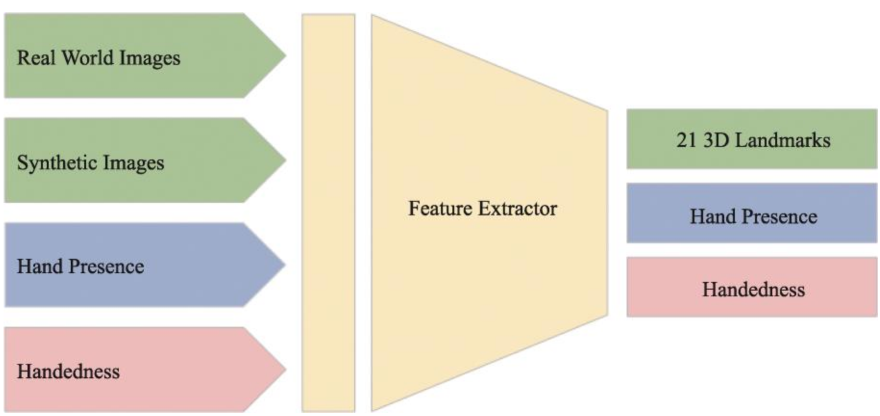

Fig.6 Hand Landmark Model Architecture

The MediaPipe graph for hand tracking is shown in Fig.7. The graph consists of two subgraphs, one for hand detection and another for landmarks computation. One key optimization that MediaPipe provides is palm detector only and it runs as needed (reasonably infrequently), saving significant computation. This is achieved by deriving the hand location in the current video frames from the computed hand landmarks in the previous frame, eliminating the need to apply the palm detector on every frame. For robustness, the hand tracker model also outputs an additional scalar capturing with the confidence that a hand is present and reasonably aligned in the input crop. The hand detection model reapplied to the next frame only when the confidence falls below a certain threshold. Examples of screenshots of real-time gesture recognition are shown in Fig.8.

3 Methodology

An algorithm is developed first to identify the landmarks, and then the writing can be started. This algorithm detects a certain point of the fingers then drawing can be done with the point that makes any movements using movement recognition. There are some options to facilitate the colours selection and delete drawn sketches in this application, and the user can use left or right hand to write. This interface and the proposed algorithm are user-friendly. The proposed algorithm is developed using Python and OpenCV libraries. Using this application, the user can draw an image by moving the fingers which appear on the screen. For this application, the basic block diagram is shown in Fig.9. As shown in the figure, the first step is feature extraction using video

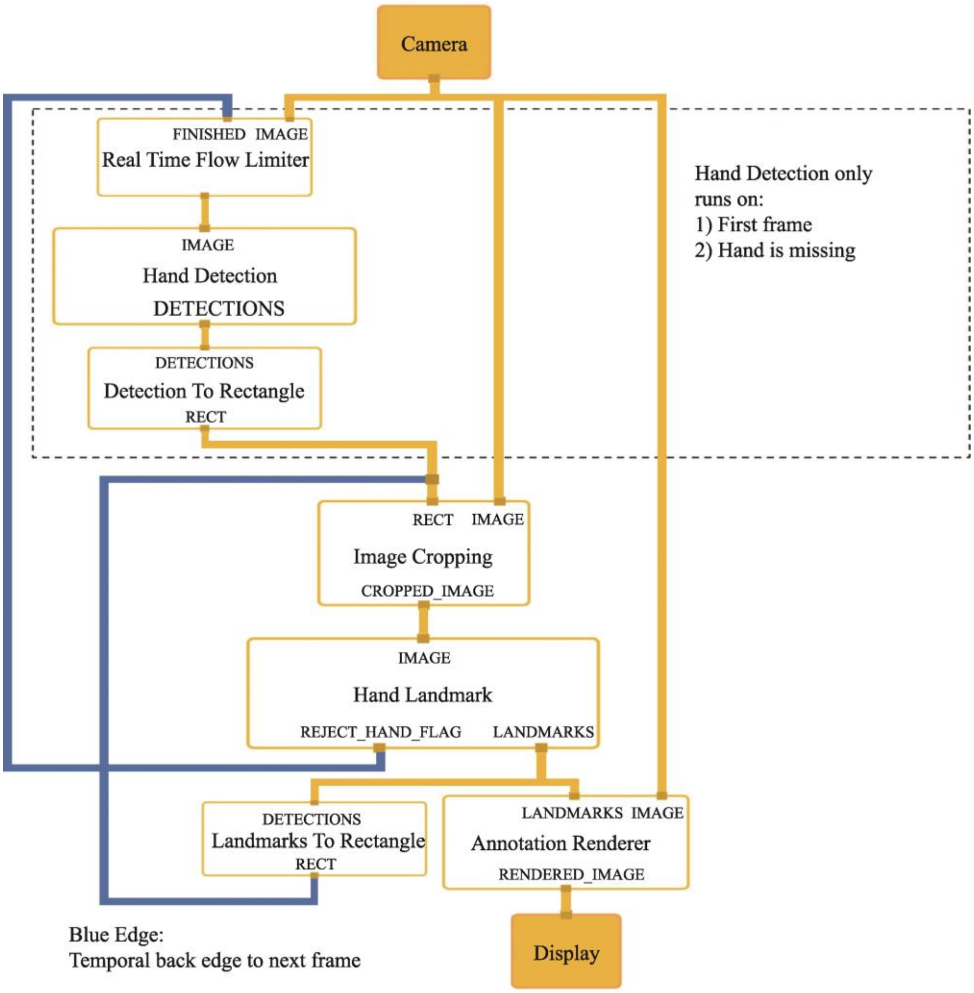

Fig.7 MediaPipe Graph for Hand Tracking

tracking with a web camera. Then it detects a certain point of the fingers. The user-selected colour is assigned for the pointer to draw. Then the movement is tracked. Ultimately the movement path is demonstrated on the screen.

The hand's module creates a 21 points based localization of a hand. In this study, the index finger, middle finger and ring finger were used to give commands and 8th, 12th and 16th LandMark were used. When the index finger, middle finger and ring finger are detected, drawing can be done. Hand landmarks are shown in Fig.10.

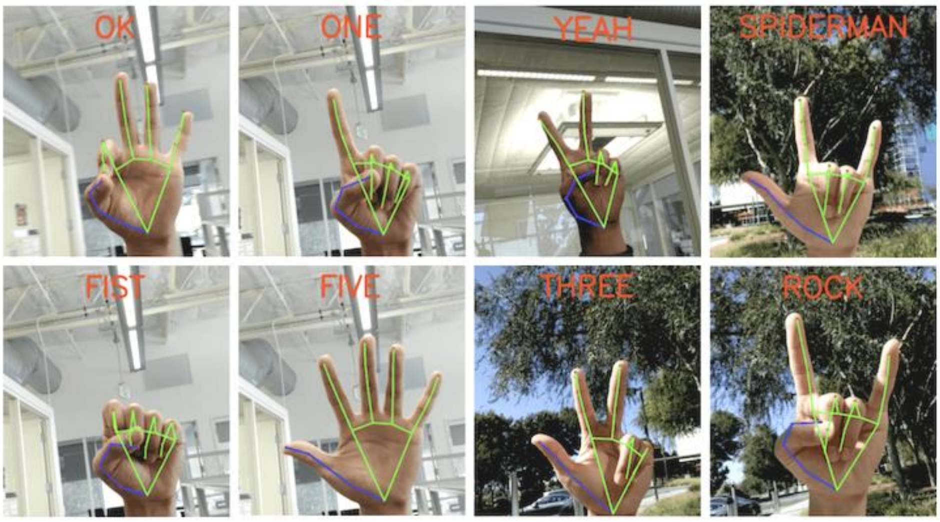

Fig.8 Examples of Screenshots of Real-time Gesture Recognition

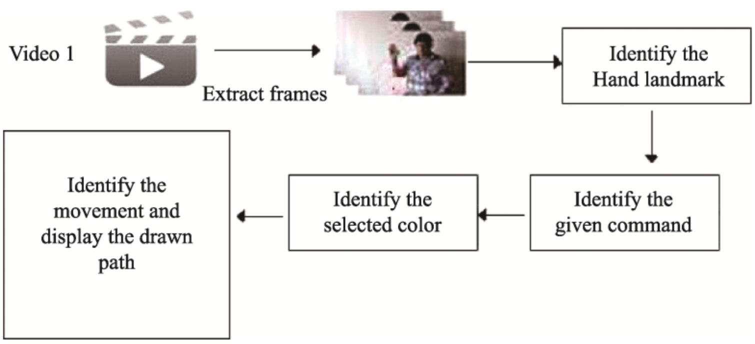

Fig.9 The Block Diagram of the Process

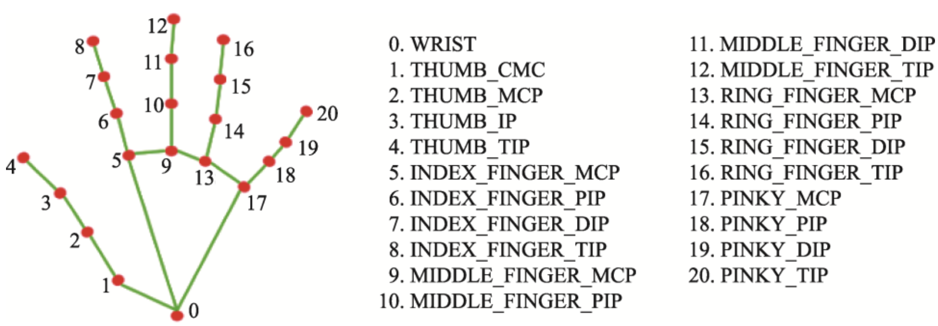

Fig.10 Hand Landmarks

4 Results

The writing technique is shown in Fig.11, and the method for stopping the writing is demonstrated in Fig.12. There are a variety of algorithms, but in this study, novelties used were Hand Landmark. Because if coloured objects are used to detect, there is a possibility to detect another object as a drawing item. When the index finger, middle finger and ring finger were detected only, the writing can be started, which otherwise is stopped. The accuracy was checked by writing 100 letters and drawing patterns. Then an accuracy of 87% was obtained. Examples for real-time writing are shown in Fig.13 and 14.

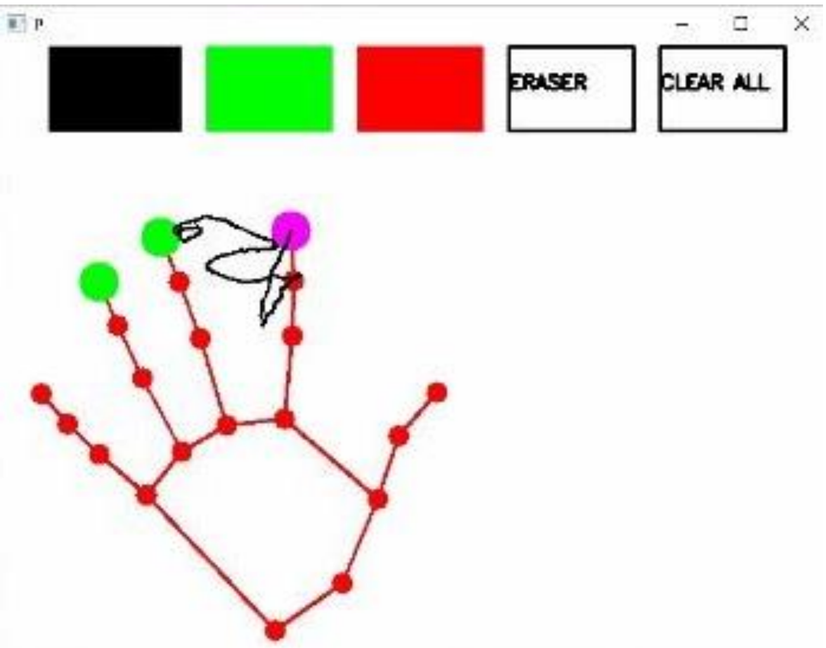

Fig.11 Writing Technique

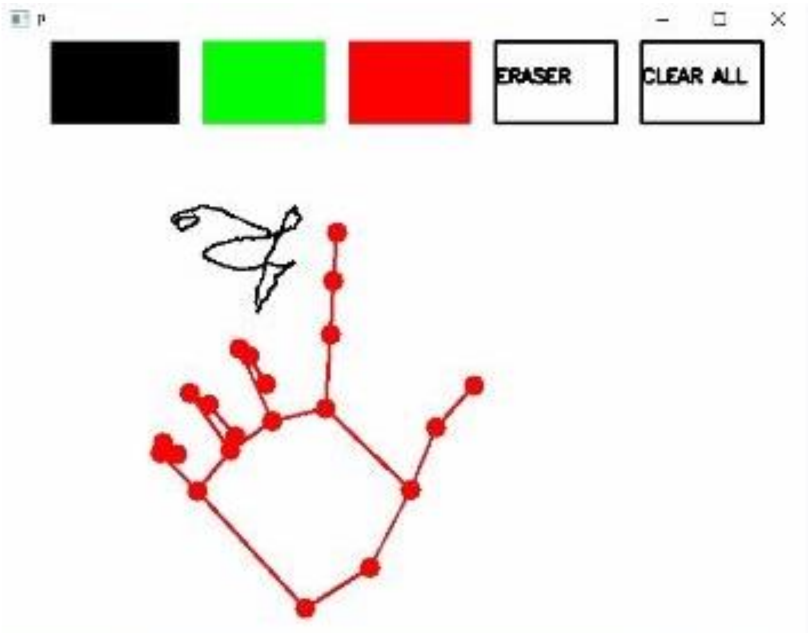

Fig.12 The Way of Stop Writing

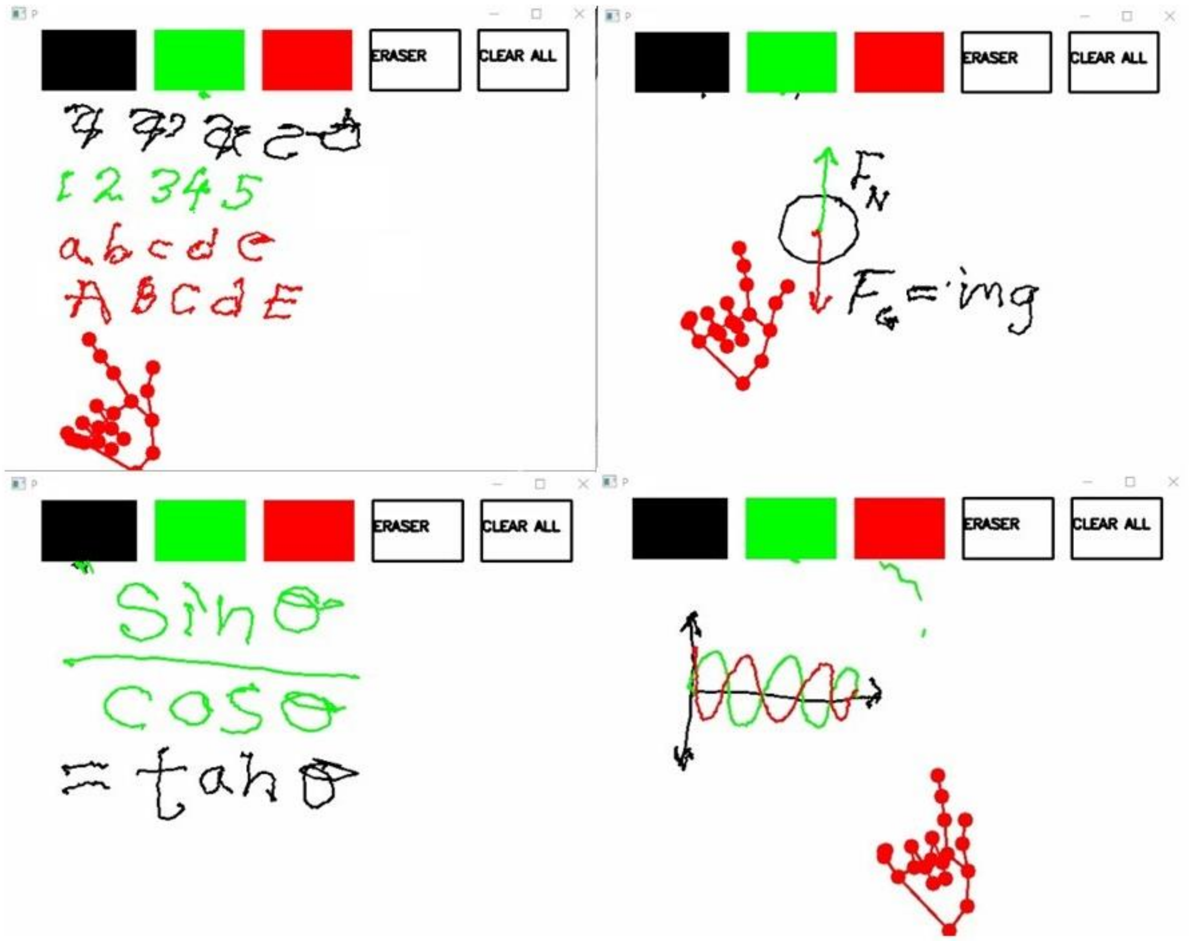

Fig.13 Examples for Real-time Writing

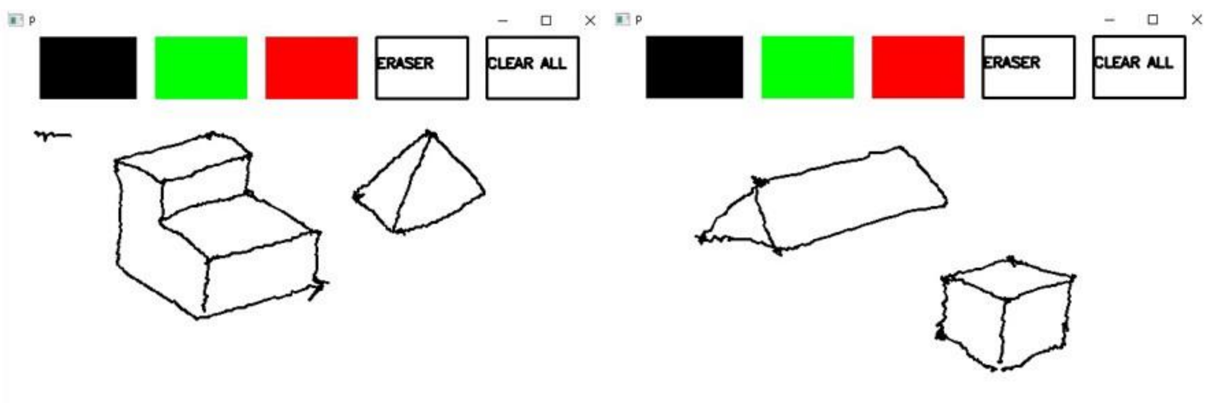

Fig.14 Examples for Real-time Writing

5 Conclusions

This paper has proposed a method for real-time writing that results in an image on the screen with improving teacher-learner interaction. The key advantage of the proposed algorithm is the requirement of t zero training and skill to handle the platform. This low-cost smart drawing is a versatile technique in the context of online meetings, forums and classroom teaching. Online teaching and learning have become an essential component in every education institute, especially for school teachers in non-technical areas who are not familiar with handling computers and accessories. In this study, video tracking is done using the web camera and Hand Landmarks are identified by the proposed programme. It detects the particular point of the fingers. Then the drawn path is exhibited on the screen. This method being user-friendly can be used even by school children for peer learning. In future, an algorithm with more functions with more colours is expected to be created to improve the application to enhance the quality of the results.

References

[1] Majumdar, D. and William, S. (2008). Chalk dustfall during classroom teaching: particle size distribution and morphological characteristics. Environmental Monitoring and Assessment, 148(1-4), pp.343-351.

[2] Lin, C., Lee, M. and Huang, H. (2015). Effects of Chalk Use on Dust Exposure and Classroom Air Quality. Aerosol and Air Quality Research, 15(7), pp.2596-2608.

[3] Sunitha, K.A., Anitha Saraswathi, P., Aarthi, M., Jayapriya, K. and Lingam Sunny. (2016). Deaf-Mute Communication Interpreter-A Review. International Journal of Applied Engineering Research, 11, pp.290-296.

[4] Cheok, M.J., Omar, Z. and Jaward, M.H. (2019). A review of hand gesture and sign language recognition techniques. International Journal of Machine Learning and Cybernetics, 10(1), pp.131-153.

[5] Rajan, R.G. and Judith Leo, M. (2019). A Comprehensive Analysis on Sign Language Recognition System. International Journal of Recent Technology and Engineering (IJRTE)

[6] Lokhande, P., Prajapati, R. and Pansare, S. (2015). Data gloves for sign language recognition system. International Journal of Computer Applications, 975, p.8887.

[7] Ambar, R., Fai, C.K., Wahab, M.H.A., Jamil, M.M.A. and Ma'radzi, A.A. (2018). Development of a Wearable Device for Sign Language Recognition. In Journal of Physics: Conference Series (Vol. 1019, No. 1, p. 012017). IOP Publishing.

[8] Zhu, Y., Yang, Z. and Yuan, B. (2013). Vision-Based Hand Gesture Recognition. In: International Conference on Service Science. Shenzhen: 2013 International Conference on Service Sciences (ICSS), p.261.

[9] Majumdar, D. and William, S. (2008). Chalk dustfall during classroom teaching: particle size distribution and morphological characteristics. Environmental Monitoring and Assessment, 148(1-4), pp.343-351.

[10] Patil, Y., Paun, M., Paun, D., Singh, K. and Borate, V. (2020). Virtual Painting with OpenCV Using Python. International Journal of Scientific Research in Science and Technology, 5(8), pp.189-194.

[11] Hemalatha, C., Muruganand, S. and Maheswaran, R. (2014). A Survey on Real-Time Object Detection, Tracking and Recognition in Image Processing. International Journal of Computer Applications, 91(16), pp.38-42.

[12] Dr. Rachna Verma (2017). A REVIEW OF OBJECT DETECTION AND TRACKING METHODS. International Journal of Advanced Engineering and Research Development, 4(10).

[13] Vishwakarma, A. and Khare, A. (2018). Moving Object Tracking System Using Morphological Image Processing. International Journal of Computer Sciences and Engineering, 6(7), pp.88-93.

[14] Shilpa, Prathap, H.L. and Sunitha, M.R. (2016). A Survey on Moving Object Detection and Tracking Techniques. International Journal of Engineering and Computer Science.

[15] Gonzalez, R.C and Woods, R.E (2008). Digital Image Processing, Taipei: Pearson Education Taiwan.

[16] Ahmed, M.A., Zaidan, B.B., Zaidan, A.A., Salih, M.M. and Lakulu, M.M.B. (2018). A review on systems-based sensory gloves for sign language recognition state of the art between 2007 and 2017. Sensors, 18(7), p.2208.

[17] Manandhar, S., Bajracharya, S., Karki, S. and Jha, A. (2019). Hand Gesture Vocalizer for Dumb and Deaf People. SCITECH Nepal, 14(1), pp.22-29.

[18] Kaur, C. and Gill, N. (2015). An Automated System for Indian Sign Language Recognition. International Journal of Advanced Research in Computer Science and Software Engineering.

[19] Shivashankara, S. and Srinath, S. (2018). American Sign Language Recognition System: An Optimal Approach. International Journal of Image, Graphics & Signal Processing, 10(8).

[20] Ivanko, D., Ryumin, D. and Karpov, A. (2019). Automatic Lip-Reading of Hearing Impaired People. Interna-

tional Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42(2/W12).

[21] Rahim, M.A., Islam, M.R. and Shin, J. (2019). Non-Touch Sign Word Recognition Based on Dynamic Hand Gesture Using Hybrid Segmentation and CNN Feature Fusion. Applied Sciences, 9(18), p.3790.

[22] Ramkumar, L. K., Premchand. S and Vijayakumar, G. K. (2019). Sign Language Recognition using Depth Data and CNN. SSRG International Journal of Computer Science and Engineering (SSRG – IJCSE)

[23] Nyirarugira, C.; Choi, H.-R.; Kim, J.; Hayes, M.; Kim, T. Modified Levenshtein distance for real-time gesture recognition. In Proceedings of the 6th International Congress on Image and Signal Processing (CISP), Hangzhou, China, 16-18 December 2013.

[24] Chen, L.; Fu, J.; Wu, Y.; Li, H.; Zheng, B. Hand Gesture Recognition Using Compact CNN via Surface Electromyography Signals. Sensors 2020, 20, 672.

[25] Pedoeem, J.; Huang, R. YOLO-LITE: A Real-Time Object Detection Algorithm Optimized for Non-GPU Computers. In Proceedings of the IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 10–13 December 2018; pp. 2503–2510.

[26] Mujahid, A., Awan, M., Yasin, A., Mohammed, M., Damaševičius, R., Maskeliūnas, R. and Abdulkareem, K. (2021). Real-Time Hand Gesture Recognition Based on Deep Learning YOLOv3 Model. Applied Sciences, 11(9), p.4164.

[27] Shinde, S., Kothari, A. and Gupta, V., 2018. YOLO-based Human Action Recognition and Localization. Procedia computer science, 133, pp.831-838.

[28] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. and Berg, A., 2016. SSD: Single Shot MultiBox Detector. Computer Vision – ECCV 2016, pp.21-37.

[29] Grzejszczak, T., Kawulok, M. and Galuszka, A., 2015. Hand landmarks detection and localization in colour images. Multimedia Tools and Applications, 75(23), pp.16363-16387.

[30] Zhang, F., Bazarevsky, V., Vakunov, A., Tkachenka, A., Sung, G., Chang, C., Grundmann, M., MediaPipe Hands: On-device Real-time Hand Tracking. arXiv: 2006.10214 [cs.CV].

Author Biographies

D. T. D. M. DAHANAYAKA. received her B.Sc. Honours degree in Mechatronics Engineering field from the Open University of Sri Lanka.

Email: tharudahanayaka@gmail.com

A. R. LOKUGE. received his B.Sc. Honours degree in Mechatronics Engineering field from the Open University of Sri Lanka.

Email: randimaaruna@gmail.com

J.A.D.E. JAYAKODY. received his B.Sc. Honours degree in Mechatronics Engineering field from the Open University of Sri Lanka.

Email: duminduerandajayakody@gmail.com

I. U. ATTHANAYAKE. received her B.Sc. (Hon.) degree from the University of Peradeniya and MPhil from University of Moratuwa Sri Lanka. Her PhD is from University of Warwick in the UK. She is currently a senior lecturer at the Department of Mechanical Engineering. The Open University of Sri Lanka. Her research interest includes molecular communication in fluid and rotating fluids.

Email: iuatt@ou.ac.lk