HUMAN ACTION IMITATION USING GAIT CLASSIFIER

P S LathaKalyampudi[1], N Akshitha Reddy[2], ChSrivalli[3], T Mahalakshmi[4]

Bvrit Hyderabad College of Engineering for Women, Hyderabad, Telangana, India 500090

Article History : **Received:** 18.04.2023 **Revised:**07.05.2023 **Accepted:** 16.06.2023

Abstract: *A computer vision problem termed as Human ac- tion imitation aims to estimate human body movements and demonstrates how the joints are related. Because it has somany implications for the automatic retrieval of films of indi- vidual behaviors using visual elements, human action imitation has gained in popularity. Human categorization, extraction of features, and activity detection are among the most frequent action imitation phases. In the domains of video surveillance, human-computer interaction, virtual reality, etc., human action imitation has important theoretical implications and numerous application possibilities. To distinguish human actions, thefootage is exhibited as Content-Based Video Retrieval (CBVR). The Gait Classifier mechanism and DBSCAN (Density-Based Spatial Clustering of Applications with Noise) are utilised to replicatethe gestures of the individual. Even though DBSCAN is being used to optimize clustering performance, gait analysis is frequently used within medical care to assistance in comprehending the gaitactions.*

Keywords: *Gait Classifiaction, DBSCAN, Mediapipe, Hu- man action, Live*

DOI: 10.48047/ECB/2023.12.8.588

INTRODUCTION

In interpersonal interactions and application of human cre- ative skill, imitation is essential. perhaps as result of the infor- mation it provides further about a person's identity, personal characteristics, and psychological state. One of the significant researchimportantsubjectsinthedomainsofmachinelearning and computer vision is the human tendency to acknowledge another person's actions. Determining a person's kinetic states is important when seeking to comprehend human activities so that the technology can clearly identify the behaviors [3][4]. Walking and running are examples of common human actions that are relatively simple to recognize and occur frequently in daily life. However, it can be more challenging to recognise more intricate movements, such "peeling an apple." Smaller, more recognisable actions can often be created by breaking down larger, more complex activities. Typically, the discovery of objects in a scene can aid in the comprehension of human behaviour by offering important details about the current situation. Humans naturally have a three-level classification that indicates the various levels of the behaviours we engage in. This hierarchy serves as a guide for us as we carry out our daily tasks. A basic element is specified by an atomic element, and these action primitives express more intricate human behaviours[19][20]. After the primitive level, action is divided into two segments: action or activity. Complex interactions, with their most fundamental level, involve human behaviours including more than contracting individuals or elements. This study includes a three classification model foraction components ,actions oractivities, and interactions.

This three-level classification retains a constant concept while differing slightly from earlier polls. The atomic movements of the limbs described as "lifting the right leg" and "extending the left arm" are examples of action primitives. [21][22]. Aindividual body parts, includes hands, arms, or upper body, is necessary for handling out atomic activities. In this review, the words "actions" and "activities" are used commonly used to describe whole-body movements made up of various of action primitives done in a temporally sequential order by a single person without the assistance of any other people or objects. All movements of the three layers, as well as all activities and actions at the intermediate level, are

particularly referred to as human actions. Individual actions such as walking, running, and waving hands are described at the level of actions or activities. [13][14]. Carrying out atomic activities requires a specific bodily component, such as the hands, arms, or upper torso. The phrase ”actions” and ”activities” are frequently used in this review to indicate to whole-body movements consisting of several action primitives accomplished in a temporally sequential order by a specific individual without the aid of any other individuals or things. The phrase ”human actions” refers specifically to all movement of the three levels as well as all responsibilities and operations at the intermediate level. At the level of actions or activities, specific actions like walking, running, and waving your hands are described.
Gestures are often seen as body movements that a person makes in response to a specific activity. The seatomic activities of an individual define a specific motion that could be a component of a more complex activity. Human activities that include two or more individuals or things are referred to as human or human-to-human interactions. Activities carried out by a group of people are termed as group actions. Personality characteristics are physical actions that represent aindividual’s emotional responses, personality, and psychological state. Ul- timately, occurrences are high-level activities which describes interpersonal interactions and identify a person’s preferences or role insociety.
In human action imitation, the character carries out a situation with an emphasis on characters against a plain background. Due to issues including backdrop clutter, partial occlusion, variations in scale, viewpoint, lighting, look, and frame reso- lution, it is difficult to create a fully automated human activity recognition system that can accurately categorise a person’s activities. Moreover, it takes time and requires knowledge of the particular event to annotate behavioural roles. The endeavour is particularly difficult because of the similarities within and between classes[19][20].That seemstobe, activities
between classes may be particularly complex, because they may be represented by comparable information, and activities within the same class may be expressed by multiple individuals with distinct body movements. The approach in which humans conduct an activity relies on their traditions, which makes it challenging to pin point the under lying activity.

Basic Architecture

Fundamental components of the suggested strategy for human action imitation as shown in the Fig.1. Give video as an input to the image pre-processing. As part of image pre-processing, it removes, integrates, and reduces frames that are unnecessary in the video. Feature extraction is carried out using the output of data preprocessing. In the feature extraction step, it will extract the key points from the frames. And then it will move on to the next step called Human action imitation. To compare the imitation of human action, a gait classifier mechanism and a media pipe were used.

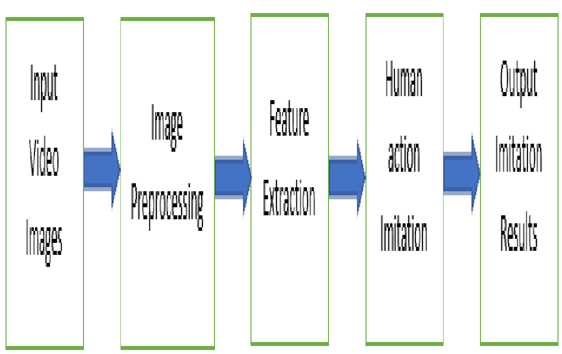

Fig. 1: Fundamental components of the suggested strategy for human action imitation

MODULES

The modules included in Human Action Imitation Using Gait Classifier are

Pre-Processing

The idea of 3D posture information, an indistinguishable posture can be communicated distinctively relying upon the survey point and the level or state of an individual. This trademark becomes an issue while learning since even in a similar posture, it very well may be perceived as an alternate posture in the grouping process which prompts execution debasement of the model. Hence, continue to standardize and adjust present so that posture can be perceived under similar circumstances prior to grouping. The equation communicating standardize is as per the following (1) .

$$norm=(x_i - min_x)/(max_x - min_x) \quad (1)$$

The posture can be standardize by applying the standardization proportion to every one of the information. Additionally, pivot the information in light of the left and right pelvic joints of posture information to deal with all information to have a similar point. Through these strategies, an issue in which a comparable posture is perceived as an alternate posture because of a point can be tackled

Feature Extraction

The result of the pre-processing module is given as the contribution to the component feature extraction module. Crude video succession comprises of gigantic spatio-transient pixel power varieties that don't contribute anything to the actual activity, for example, pixels connected with the shade of garments and jumbled foundation. Include extractionis an interaction that identifies and concentrates most delegate data from crude information as elements. Any video succession will produce a particular number of highlights, and different video groupings will have unmistakable number of elements. Include portrayal is an interaction to give aone of a kind portrayal for each video grouping in view of the extricated highlights. The last portrayal ought tobe of similar aspect among various recordings. The video successions is partitioned in 255 casings and the result of this module is given to present recognition module, to prepare the information.

Pose Detection

The result of the component extraction module is given as the contribution to present discovery module and the information is be prepared with Gait Classifier and DBSCAN methods and the mirroring the activities of the individual is the result of this module as displayed in Figure2.

Fig. 2: Action Imiation of the Person

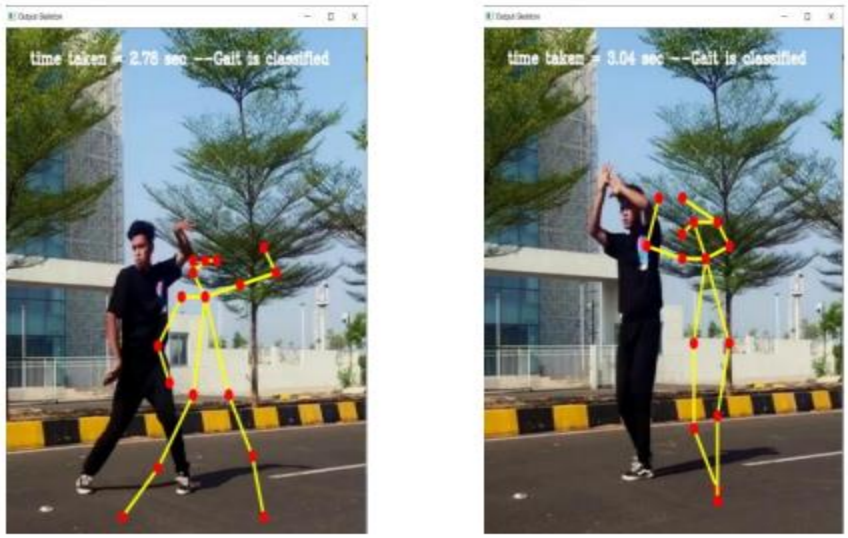

Fig. 3: Gait Classification Results

TECHNIQUES

Gait Classifiation

Gait can be classified from videos. This classificationoutput can help people understand whether their rehabilitation moves are correct [7][8]. To implement this project, the module used different limb movements such as knee movement, toe, ankle, etc.

1. Gait Pattern Extraction: From videos, this module will be used to extract the limb movements during walking phases andusethosefeaturestoperformgaitphaseclassification.
2. Gait pattern clustering: The listed module features will be clustered using DBSCAN clustering algorithms, andsimilar movements will automatically go to the same cluster, such as the ankle, knee, and toe will have one cluster, and their movementwillgotoitsappropriatecluster[9][10].
3. Gait Phase Reconstruction: The model can analyze clusters to find movement features that are similar and can thenpredictgaitmovementusingthismodule.Usingtheabove three modules' algorithms, a gait classification application is being developed as shown in Figure 3, which will accept video input from the user and then begin categorization of gaitmovement.

DBSCAN(Density based spatial clustering application with Noise)

Cluster analysis is an unsupervised learning method for understanding the properties of a huge database by classifying data based on their similarity. Because of its efficiency, the k- means clustering method, which sets the number of groups in advance, is commonly used for most cluster analysis [11][12]. DB-SCAN has the advantage of automatically forming a suitable cluster.

The process of tuning parameters to match datasets has con- tinued in the past because [11][12], as shown in Figures 4 and 5, it

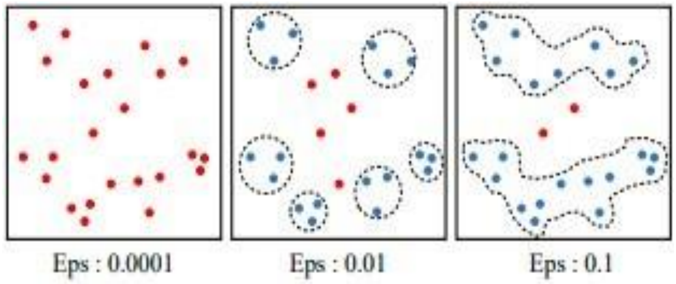

Fig. 4: The tendency of clustering according to Eps.

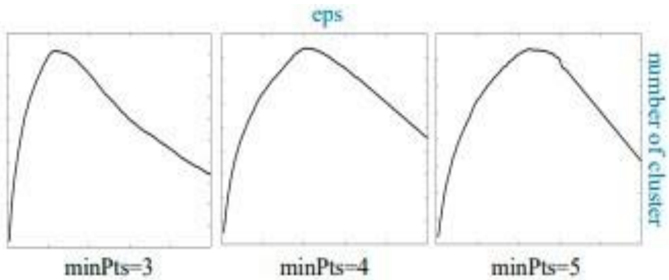

Fig. 5: The tendency of clustering according to minPts

is critical to appropriately tune two parameters (minPts, Eps) in order for DBSCAN to work successfully.

Media pipe

Utilizing RGB video, Media Pipe Posture is a machine learning technique for high-fidelity body pose continuous monitoring. [1][2] frames to infer 33 3D landmarks as shown in the Figure 6 and Table 1 and a background segmentation mask over the entire body.

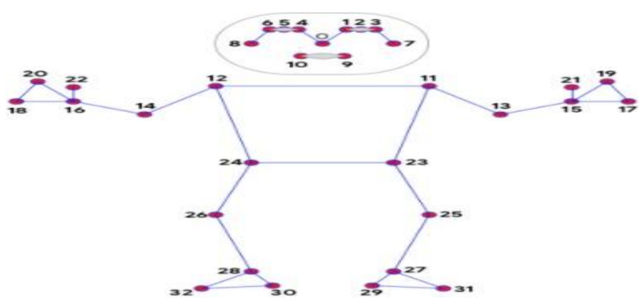

Fig. 6: 33 3D landsmarks using mediapipe

A Graph is the name given to the MediaPipe perception pipeline. Take, for example, the first solution, the human body [1][2]. A 265-frames video as input [15][16], and the result is human posture landmarks rendered beside the human body as shown in Figure 7.

Live Streaming

While performing live for human action imitation, it will replica the human's actions as they are performed via the human. Those movements will be captured alongside the human body as shown in Figure 8.

Table I: Key Points And Its Gestures Of Mediapipe

Key Points	Gesture
0	nose
1	left eye inner
2	left eye
3	left eye outer
4	right eye inner
5	right eye
6	right eye outer
7	left ear
8	right eye
9	mouth left
10	mouth right
11	left shoulder
12	right shoulder
13	left elbow
14	right elbow
15	left wrist
16	right wrist
17	left pinky
18	right pinky
19	left index
20	right index
21	left thumb
22	right thumb
23	left hip
24	right hip
25	left knee
26	right knee
27	left ankle
28	right ankle
29	left heel
30	right heel
31	left foot index
32	right foot index

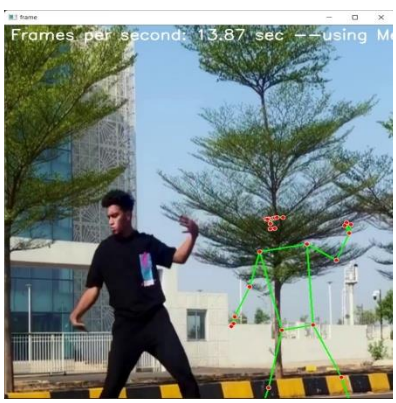

Fig. 7: Mediapipe Result

RESULTS

The User Interface(UI) of human action imitation contains five buttons, which are

1. Upload VideoFile
2. Start Gait Phase Classification
3. UsingMediapipe
4. Live demo
5. Exit

Firstly, the user will gather all the necessary videos for the projec t and then the user will create a directory called

Fig. 8: Live Stream Result

Fig. 9: Uploading video

Input Videos in the system. Then, the user will move all the gathered videos to the directory. Now, the user will upload a video file from the directory into the first button. After uploading the video successfully, a message will be appeared on the screen. It shows the video path with a message 'loaded' as shown in the Figure9.

After successful completion of video uploading, then the user will start next step called Gait Classification Phase. In this phase, the streaming video will processed into 256 frames. Because splitting of video into 256 frames, there will be a pause between each frame, so human eye can capture every

Fig. 10: Gait classification Result

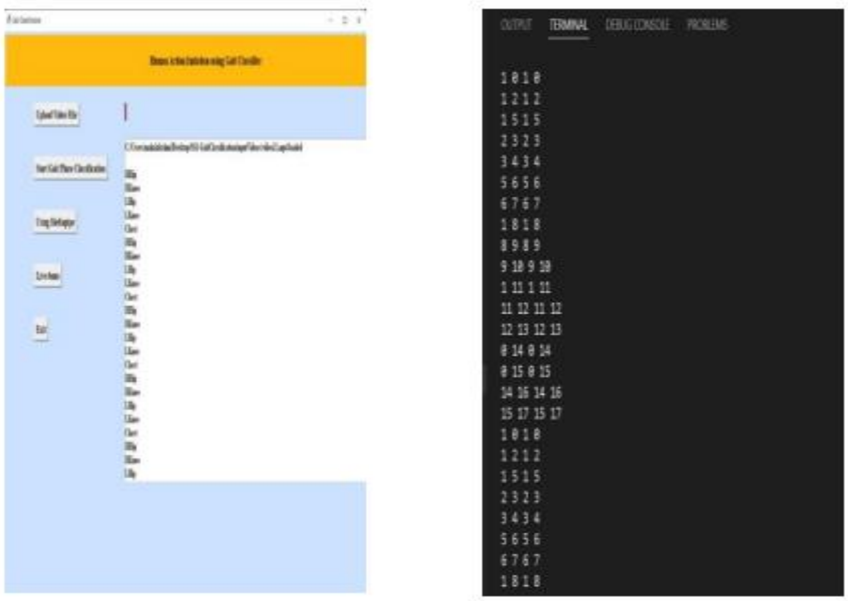

Fig. 11: Live Stream Result

pose in the video easily. After classification of Gait, the key points will appear alongside the human body as shown in the Figure 10. After, completion of the gait classification, video will be automatically stopped and disappeared from the screen. Whatever actions perfomed by the human in the video, from that specific key movements captured and displayed on the screen as shown in the Figure 11 and the key points from the video will be displayed on the terminal as shown in the Figure 11.

After completion of gait phase, then the user used Mediapipe technique. In this technique the output will be same as gait classification technique as shown in the Figure 12, but there is no splitting of frames in the streaming video. So, there will be no pausing for frame to frame. So human can't understand the poses in the video easily.

Fig. 12: Media pipe Result

Fig. 13: Live Result-1

Live is the last step in this project. While performing live for human action imitation, it will replica the human’s actions as they are performed via the human Figure 13. Those movements will be captured alongside the human body as shown in the Figure 14. After performing live, the user uses the exit button to come out from the main UIpage.

CONCULSION

The utilization of Machine Learning in gait analysis is a de- veloping field in Computer Vision.In the fields of surveillance

Fig. 14: Live Result-2

footage, interpersonal communication, VR(Virtual Reality) technology, etc., human action imitation has strong theoretical significance and a wide range of potential applications. In this project, the human action imitation is capturedusing gait classifier and mediapipe techniques. Both techniques gave output alongside the human. But the captured outputwith gait classifier is more accurate with human activity because it carried out actions for each frames of the streaming video. The captured output with media pipe is not much accurate because the output is a non-stop streaming authentic video, which can't be hold close by the human. This work can be extended in future by including detection of multiple persons in the video surveillance and other non-living things movements detection in the streamingvideo.

REFERENCES

1. D. Singh, S. Panthri and P. Venkateshwari, ”Human Body Parts Measure- ment using Human Pose Estimation,” 2022 9th International Conference on Computing for Sustainable Global Development (INDIACom), 2022, pp. 288-292, doi: 10.23919/INDIACom54597.2022.9763292.
2. Pradeep Bheemavarapu, P S LathaKalyampudi and T V Madhusudhana Rao, “An Efficient Method for Coronavirus Detection Through X-rays using deep Neural Network”, Journal of Current Medical Imaging,[online Available] Vol.18, No. 6, with ISSN: 1875-6603,2022.
3. D. C. Luvizon, D. Picard and H. Tabia, ”Multi-Task Deep Learn- ing for Real-Time 3D Human Pose Estimation and Action Recog- nition,” in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 8, pp. 2752-2764, 1 Aug. 2021, doi: 10.1109/TPAMI.2020.2976014.
4. S Satyanarayana ,“Privacy Preserving Data Publishing Based On Sensitivity in Context of Big Data Using Hive”,Journal of Bigdata(Springer), Volume:5,Issue:20, ISSN: 2196-1115, July 2018.
5. M. Wang, J. Tighe and D. Modolo, ”Combining Detection and Tracking for Human Pose Estimation in Videos,” 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 11085- 11093, doi: 10.1109/CVPR42600.2020.01110.
6. P.Mahesh Kumar et al “Frequent Pattern Retrieval on Data Streams by using Sliding Window”, EAI Endorsed Transactions on Energy web,Volume:5,issue:35,2021.
7. A. Rohan, M. Rabah, T. Hosny and S. H. Kim, ”Human Pose Estimation- Based Real-Time Gait Analysis Using Convolutional Neural Network,” in IEEE Access, vol. 8, pp. 191542-191550, 2020, doi: 10.1109/AC- CESS.2020.3030086.

8. T.V. Madhusudhana Rao, Suresh Kurumalla, Bethapudi Prakash, "Matrix Factorization Based Recommendation System using Hybrid Optimization Technique, EAI Endorsed Transactions on Energy Web,Volume:5,issue:35,2021.
9. S. U. Yunas, A. Alharthi and K. B. Ozanyan, "Multi-modality sensor fusion for gait classification using deep learning," 2020 IEEE Sensors Applications Symposium (SAS), 2020, pp. 1-6, doi: 10.1109/SAS48726.2020.9220037.
10. T.V. Madhusudhana Rao, P.S. LathaKalyampudi, "Iridology based Vital Organs Malfunctioning identification using Machine learning Techniques", International Journal of Advanced Science and Technology, Volume: 29, No. 5,PP: 5544 – 5554,2020.
11. D. Deng, "DBSCAN Clustering Algorithm Based on Density," 2020 7th International Forum on Electrical Engineering and Automation (IFEEA), 2020, pp. 949-953, doi: 2020.
12. S.VidyasagarAppaji, P. V. Lakshmi, "Maximizing Joint Probability in Visual Question Answering Models", International Journal of Advanced Science and Technology Vol. 29, No. 3, pp. 3914 – 3923,2020.
13. C. J. Dhamsania and T. V. Ratanpara, "A survey on Human action recognition from videos," 2016 Online International Conference on Green Engineering and Technologies (IC-GET), 2016, pp. 1-5, doi: 10.1109/GET.2016.7916717.
14. VidyasagarAppajisetti , "A Novel Scheme For Red Eye Removal With Image Matching", Journal of Advanced Research in Dynamical & Control Systems, Vol. 10, 13-Special Issue, 2018.
15. B. Antic, T. Milbich and B. Ommer, "Less is more: video trimming for action recognition", International Conference on Computer Vision Workshops(ICCVW),IEEE,pp.515-521,2013.
16. Krishna Prasad, P.E.S.N, "A Secure and Efficient Temporal Features Based Framework for Cloud Using MapReduce", springer, 17th International Conference on Intelligent Systems Design and Applications (ISDA 2017),Volume:736,pp:114-123, ISSN 2194-5357 Held in Delhi,India, December 14–16, 2017.
17. N. Nguyen and A. Yoshitaka, "Human interaction recognition using independent subspace analysis algorithm", International Symposium on Multimedia (ISM), IEEE, pp. 40-46, 2014.
18. Madhusudhana Rao, T.V., Srinivas, Y, "A Secure Framework For Cloud Using Map Reduce",Journal Of Advanced Research In Dynamical And Control Systems(IJARDCS), Volume:9,Sp-14,Pp:1850-1861,ISSN:1943-023x,Dec, 2017.
19. H. Wang, A. Kl a¨ser, C. Schmid and C. Liu. "Action recognition by dense trajectories", Computer Vision and Pattern Recognition, IEEE, pp. 3169-3176, 2011.
20. Sushma Rani N, "An Efficient Statistical Computation Technique for Health Care Big Data using R", Scopus, IOP Conference Series: Materials Science and Engineering, Volume: 225, ISSN:1757-8981,ISSUE NO :012159,2017.
21. K. Yun, J. Honorio, D. Chattopadhyay, T. Berg and D. Samaras, "Two- person interaction detection using body-pose features and multiple instance learning", Computer Society Conference on Computer Vision and Pattern Recognition Workshops, IEEE, pp. 28-35, 2012.
22. Krishna Prasad, M.H.M., Thammi Reddy, K , "A Efficient Data Integration Framework in Hadoop Using MapReduce" Published in Computational Intelligence Techniques for Comparative Genomics , Springer Briefs in Applied Sciences and Technology , ISSN:2191-530X,PP 129-137, October 2014
23. NageshVadaparhi, Srinivas Yarramalle , "A Novel clustering approach using Hadoop Distributed Environment", Springer, (Applied Science and Technology), ISSN:2191-530X,Volume:9, pp:113-119,October 2014.
24. BalajeeMaram, Guru KesavaDasuGopisetty, P Srinivasa Rao, "A Framework for Data Security using Cryptography and Image Steganography",International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-11,September,2019.

Fig. 12. Mediapipe Result

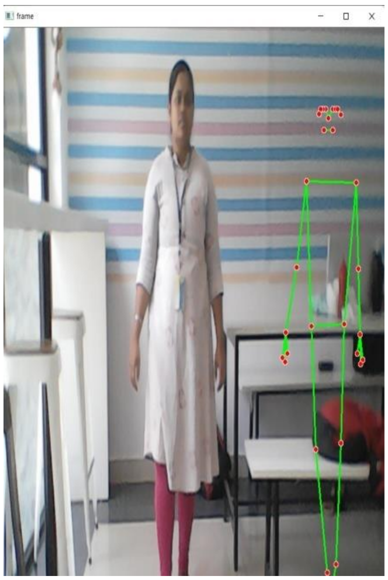

Fig. 13. Live Result-1

Fig. 14. Live Result-2