iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** IV **Month of publication:** April 2021

DOI: https://doi.org/10.22214/ijraset.2021.33661

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue IV Apr 2021- Available at www.ijraset.com

Heterogeneous Hand Gesture Recognition using 3D Dynamic Skeletal Data

C. Anitha[1], S. Sethulekshmi[2], A. Shanmathy Shree[3], S. Vinodhini[4]
[1] *Assistant professor,* [2, 3, 4] *UG Students, Electronics and Communication Engineering, Jeppiaar Engineering College, Tamil Nadu, India.*

Abstract: ***Hand gestures area unite the foremost natural and intuitive non-verbal communication medium whereas interacting with a pc, and connected analysis have recently boosted interest. To boot, the distinctive options of the hand provided by current business cheap depth – camera are often exploited in numerous gesture recognition based systems, for human – computer interaction.***
This paper builds a sturdy hand form options from two modalities of depth and skeleton form approach, we have a tendency to use the movements, the rotations of the hand joints with relevance to their neighbors, and also the skeleton point cloud to find out the 3D geometric transformation.
For the hand depth form approach, we have tendency to use the feature illustration from the hand element segmentation model. Finally, we propose a multi-level feature LSTM with CONV1D and CONV2D algorithm where LSTM is used to manage the range of hand options.
Therefore, we tend to propose a completely unique technique by exploiting skeletal point clouds from skeletal form as well depth features from real hand depth form in order for the LSTM model to benefit from both. Our projected technique achieves the best result with skeletal and depth data.
Keywords: Hand gesture, Detection, LSTM, CONV1D, CONV2D.

I. INTRODUCTION

Besides the common language modalities, hand gestures are also typically utilized in our daily lives to speak with one another. Additionally hand gesture recognition is one of the way in which computers will interact with humans by translating the human hand gestures into commands. Especially the 3D hand estimation combined with depth camera has contributed to winning launch of computer game and increased reality applications like linguistic communication recognition, computer game, robotics and many interaction systems.

The background of this paper is to induce economical and easy operating methodology for the deaf and dumb people to speak with others and it involves use in real world application. This paper involves two easy nonetheless effective ways that has been used they're i) TEMPORAL PYRAMID and ii) LSTM with CONV1D and CONV2D Algorithm. The Temporal Pyramid splits the skeletal form and real hand depth form. Then the neural network model that is trained with our dataset will be used to detect the hand gesture.

The LSTM with CONV1D and CONV2D is employed to extract the feature, decrypt the unwanted information and identify the proper hand gesture. Then the raspberry pi is dumped with python code for each gesture respectively. Then a speaker is connected to the raspberry pi through that audio for every gesture is recited. The main goal of our paper is to assist dumb, deaf to others through hand gestures.

II. EXISTING SYSTEM

Existing technology of hand gesture detection: On having the input sequence of images through in- depth camera it uses pre-processing steps for removal of background noise and employs K-means clustering for the removal of the background and separate the hand gesture, so that only segmented significant cluster or hand object is to be processed in order that solely divided vital cluster or hand object is processed so as to calculate the primarily based options and provide the detected hand gesture output.

A. Drawbacks of Existing System

1) Simple shape identification is only possible, which leads to the low efficiency in real-time gesture identification.
2) The algorithm used is slow and time-consuming.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue IV Apr 2021- Available at www.ijraset.com

III. III.PROPOSED SYSTEM

The proposed system works on two methods each having their own use cases; they are TEMPORAL PYRAMID and LSTM WITH CONV1D AND CONV2D. TEMPORAL PYRAMID is used to split the hand gesture into the skeletal view and hand depth view respectively. LSTM with CONV1D and CONV2D algorithm is used to detect the hand gesture and deal with the diversity of hand features. Python is used to code the respective audio output for each gesture and dumped in raspberry pi. And the audio is recited through a speaker connected to the raspberry pi. The input image is used to detect what gesture has been shown with the help of pre trained model. The model detects the hand gesture shown in front of the in-depth camera. And will recite the audio for the gesture respectively.

Advantages of proposed system

1) This requires less training time and requires less memory space.
2) Simple, fast and easy to implement.
3) Cheap and effective solution for the bed ridden and deaf, dumb people.

A. Block Diagram of Proposed System

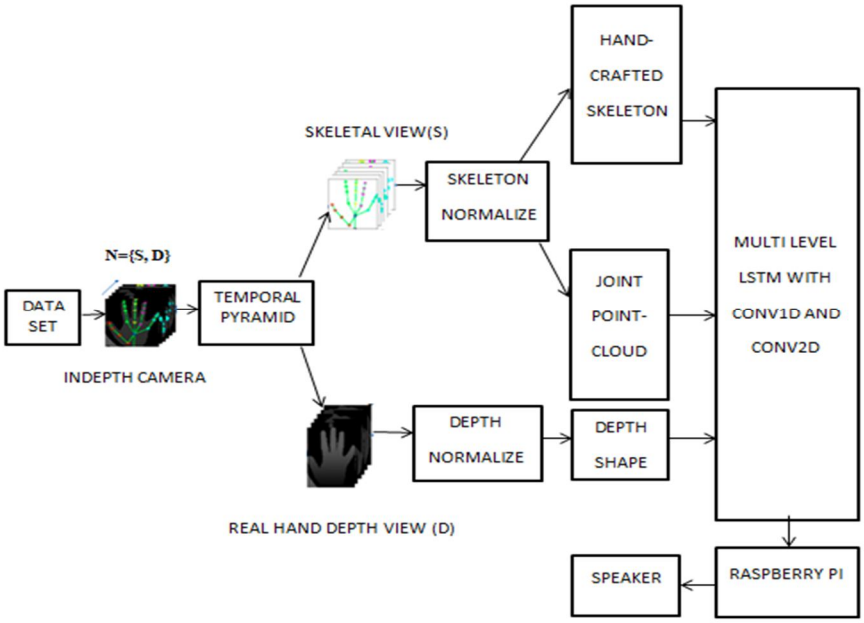


Figure 1: Block diagram of the system

The block diagram of our proposed system is shown in Figure 1. The Dataset for hand gestures is collected and also the system is pre-trained using the Datasets. Then when the gesture is shown in front of the in-depth camera. We apply the temporal frame sub-sampling for every dynamic gesture to the specific length. We split the dynamic gesture into hand skeletal data represented as S and hand depth data represented as D as the input to the feature extraction step. Within the normalize section, It normalizes the skeletal and depth data information to the particular size in the dataset and extract the features of the views respectively. In our projected framework, the feature extraction section consists of feature varieties to exploit the robust features for the dynamic hand gesture. First two features such as hand skeleton and join point cloud are extracted from skeletal data. Through hand skeletal information movements, folding of fingers, internal angles between the fingers are found. Through the joint point-cloud we can extract the 3D joint points of normal hand. With real hand depth information depth and shape of the hand are found. We propose the multi-level feature LSTM model to train on every hand gesture feature the multi-level feature LSTM with Conv1D, the Conv2D pyramid is used to deal with the diversity of hand features and detect the gesture correctly. Our design allows the LSTM filter layer to take advantage of the semi-permanent dependencies between the frames and remove the complexity and exploit the character of the input feature. Finally; Audio output for specified gesture is recited through speaker which is connected to raspberry pi where the program is dumped.

IV. RESULTS AND DISCUSSION

The software that is used in the project is mainly python based Anaconda 3 and python 3.9

A. Server Initiation

Here the anaconda prompt is initiated using the proper command and the camera is switched on to recognize the gesture shown. The video feed is taken and used in image by image basis for the algorithm. The initiation of the anaconda turns on the in-depth camera is shown in Figure 2 and live video feed is taken. And this is used for further feature extraction and detection.

Figure 2: Server initiation

B. Temporal Pyramid And LSTM with Conv1d and Conv2d

The temporal pyramid divides the gesture shown in front of the camera into skeletal view and hand depth view in order to improve the accuracy of the output. After the skeletal view is obtained it is normalized where the unwanted noise and background are removed and features like hand skeleton, joint point cloud is obtained. Similarly the real hand depth view is normalized and features like depth and shape of the hands are extracted. Among the hand gesture here two detection output of the hand gesture is shown in Figure 3 and Figure 4 respectively. The gestures are detected using the LSTM, CONV1D and CONV2D algorithm. The LSTM decodes the unwanted part of the image and extract the features to detect the gesture and provide the correct detection output.

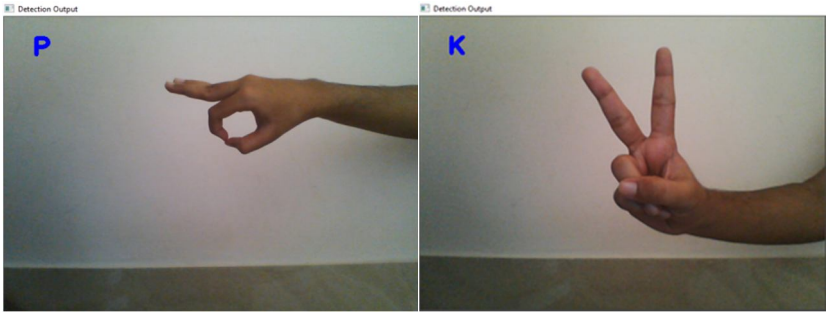


Figure 3: Detection output of gesture P Figure 4: Detection output of gesture K

C. Processing Module Raspberry PI

The reboot command is used to initiate the raspberry pi where the code for each gesture is dumped. The initiation of the raspberry pi is shown in Figure 5. After the detection output from the LSTM block , in raspberry pi an audio output that is a task is been programmed using the python code and dumped in it. So when we show a gesture in front of the camera the gesture is detected using the algorithm. And according to the detected output the audio output that is the task commands dumped in the raspberry pi is recited through the speaker connected to the port of raspberry pi. The Figure 5 shows the overall setup.

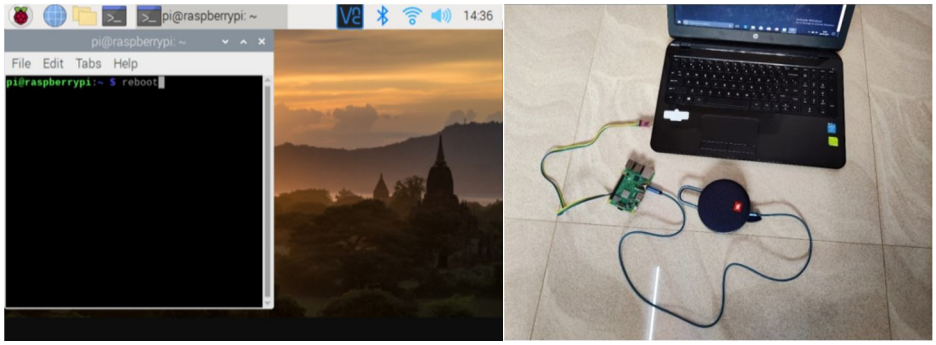


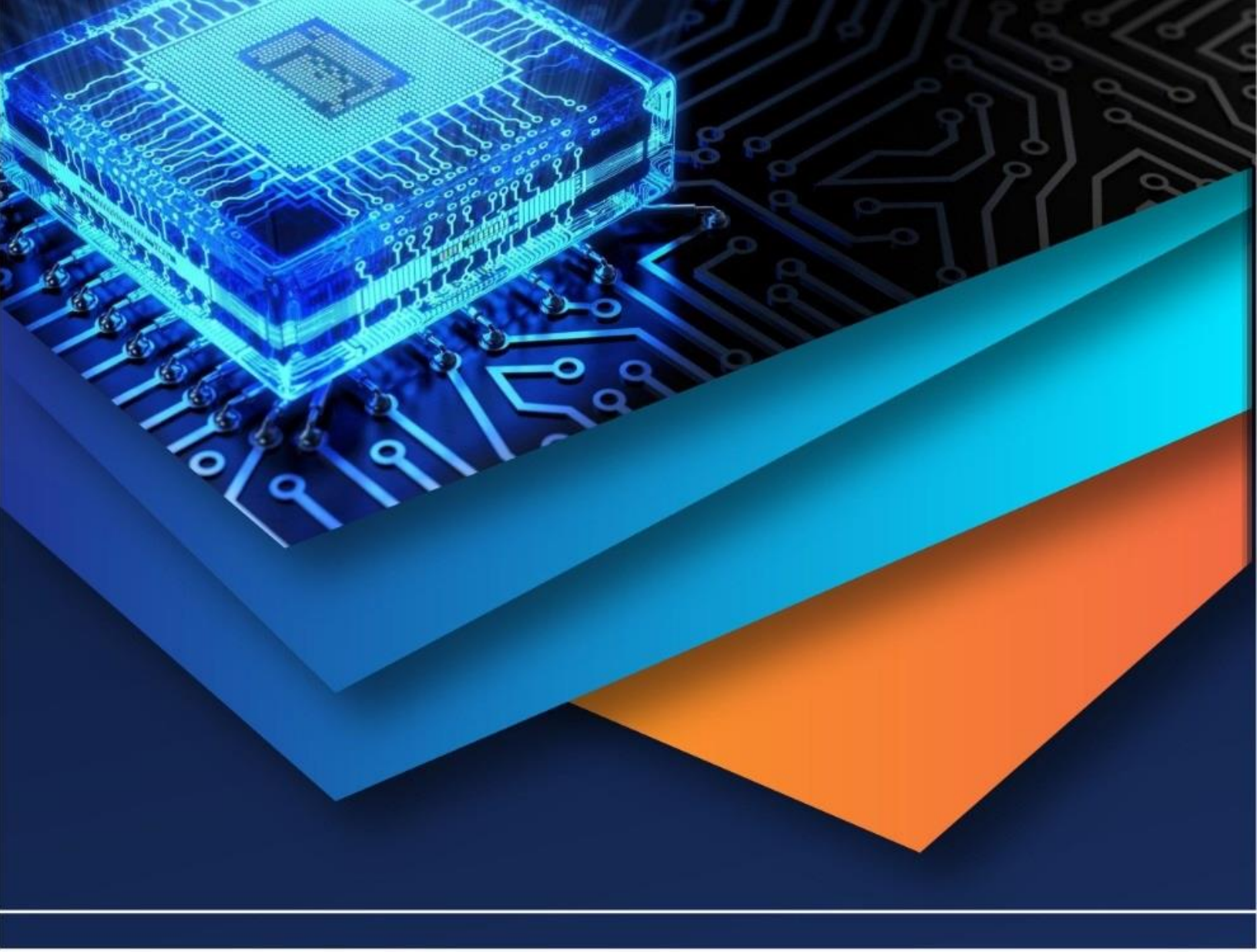
Figure 4 Initiation of Raspberry Pi Figure 5 Overall Setup of the system

V. CONCLUSION

In our proposed model, we have designed the prototype model for deaf and dumb people. This is used to effectively improve the livelihood of the deaf and dumb people. The system successfully performs hand gesture recognition and gives an audio output of the results. So, the system has a very large scope in terms of helping the hearing aided people and old people. The results are more accurate compared to the other existing works.

REFERENCES

[1] Alina Kuznetsova, Laura Leal-Taixé, Bodo Rosenhanh, (2013) "Real-Time Sign Language Recognition Using a Consumer Depth Camera", 2015.
[2] Andrew Fitzgibbon, Mat Cook, Toby Sharp, Mark Finocchio, Richard, Richard Moore, Alex kipman, Andrew Blake (2011) "Real-time pose recognition in parts from single depth images" , IEEE, pp.55-58(2011)
[3] Chen, G. AlRegib, and B.-H. Juang, (2018) "Hand gesture recognition using Leap Motion via deterministic learning", 36th Chinese Control Conference, pp. 22-27, 2018.
[4] Jie Huang, Wengang Zhou, Houqiang Li, Weiping Li,(2015) "Sign Language Recognition using 3D convolutional neural networks" IEEE International Conference on Multimedia and Expo, pp. 14-29, 2015.
[5] Linchu Yang, Ji'an Chen, and Weihang Zhu**,** (2020)"Dynamic Hand Gesture recognition based on leap motion controller and two layer bidirectional recurrent neural network", 2020;
[6] Ren, S.; He, K.; Girshick, R.; Sun, J, "Towards Real-Time Object Detection with Region Proposal Networks", FasterR-CNN IEEE Trans. Pattern Anal. Mach. Intel. 2017, 39, 1137–1149.
[7] Sang-Heon Lee, Myoung-Kyu Sohn, Dong-Ju Kim, Byungmin Kim, Hyunduk Kim (2013) **, "**Smart TV interaction system using face and hand gesture recognition" ,IEEE International Conference on Consumer Electronics, 11-14 Jan.2013; pp.279-295.
[8] Simonyan, K.; Zisserma A (2015) "Very deep convolutional networks for large-scale image recognition". In Proceedings of the 3rd International Conference on Learning Representations, Santiago, CA, USA, 7–9 May 2015; pp. 1–14.
[9] Tomas Bravenec, Tomas Fryza (2019), "Multiplatform System for Hand Gesture Recognition" IEEE International Symposium on Signal Processing and Information Technology, 2019.
[10] Zhi-hua Chen, Jung-Tae Kim, Gaining Liang, Jing Zhang, and Yu-Bo Yua(2014),"Real-Time Hand Gesture Recognition Using Finger Segmentation" ,25 Jun 2014; pp. 29-33.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY