Sign Language Detection using LSTM Deep Learning Model

(Action Recognition with Python)

CONTENTS

ABSTRACT

This project was created intending to use computer vision to be able to recognize the Sign Language in real time with high accuracy. The reason for such a project is to help diminish the gap between those who can hear well and those hard of hearing or even deaf. This can be overcome by creating a dataset of images that correspond to the alphabets, digits or signs applied to deep neural networks. These images are labeled according to the letter being signed. They are processed through a neural network using transfer learning to help the machine “learn” what is being signed after already having been taught on larger datasets of many more images and classifications.

Literature Survey

Reference	Contents
Hand sign language recognition using multi-view hand skeleton Published on July-2020 **Razieh Rastgooa, Kourosh Kiania, Sergio Escalerab**	• In this paper, we proposed a novel deep learning-based pipeline architecture using the capabilities of SSD, 2DCNN, 3DCNN, and LSTM for efficient automatic hand sign language recognition from RGB videos. • We used novel hand skeleton features representation in the model for projecting them to three surfaces to feed them to the 3DCNNs in order to get more rich features. Furthermore, we applied 3DCNNs on pixel level and heat map features to obtain discriminative features.

Literature Survey

Reference	Contents
Real-time isolated hand sign language recognition using deep networks and SVD Published on March-2021 **Razieh Rastgoo, Kourosh Kiani & Sergio Escalera**	• One of the challenges in computer vision models, especially sign language, is real-time recognition. In this work, we present a simple yet low-complex and efficient model, comprising single shot detector, 2D convolutional neural network, singular value decomposition (SVD), and long short term memory, to real-time isolated hand sign language recognition (IHSLR) from RGB video. • We employ the SVD method as an efficient, compact, and discriminative feature extractor from the estimated 3D hand keypoints coordinators.

Literature Survey

Reference	Contents
Sign Language Detection from Hand Gesture Images using Deep Multi-layered Convolution Neural Network Published on March 2021 **Rajarshi Bhadra,Subhajit Kar**	• In this paper, a deep CNN architecture consisting 5 layers has been proposed to detect and classify sign languages from hand gesture images. The proposed methodology uses both static (0 - 9 and A - Z) and dynamic (alone, afraid, anger etc.) gestures in the training, validation and blind testing phase to make the system more robust.

Literature Survey

Reference	Contents
A new hybrid deep learning model for human action recognition Published on May-2020 **Neziha Jaouedia, Noureddine Boujnahb,Med Salim Bouhlelc**	• This work presents a novel approach in the domain of human action recognition. This approach is based on the analysis of video content and extraction features. The Motion features are presented by human motion tracking using GMM and KF methods. • Then others features are based on all visual characteristic of each frame on video sequence using Recurrent Neural Networks model with Gated Recurrent Unit. The main advantages of this novel approach are the analysis and the extraction of all features in each time and in each frame of video. This hybrid model has an important role to better human action recognition.

Literature Survey

Reference	Contents
Convolutional and recurrent neural network for human activity recognition: Application on American sign language Published on February-2020 **Vincent Hernandez ,Tomoya Suzuki,Gentiane Venture**	• This study compared various conventional machine learning and deep learning models to classify American sign language. Moreover, a robust user-independent k-fold cross-validation and test phase were provided. • This contrast previous work, where the validation and/or the test phase were not user-independent, or lack of information was provided.

INTRODUCTION

- Sign language is a form of communication that relies heavily on hand kinematics and facial emotions. Hearing impaired persons use it frequently to communicate with one another, but it is rarely utilized by non-hearing-impaired people.
- As a result, they only deal with hearing-impaired people directly, which severely limits social contacts. A Realtime translation with interpreters is an alternative, although they are not always available and can be rather expensive.
- A system that allows for automatic translation would be quite useful. In this subject, many novel strategies have lately been created. We will create a program to translate sign language into OpenCV in this project. It describes a method for recognizing and translating Custom Sign Language into normal text.

METHODOLOGY

- We develop a sign detector in this sign language recognition project that detects custom gestures and can easily be modified to encompass a large array of additional signs and hand motions, including the alphabets.
- The OpenCV, Mediapipe, Tensorflow, and Keras Python modules were used to create this project. The OpenCV feed examines the frames of live video from a camera to detect the action of a person who is being displayed at that moment in time.
- To extract key points from our hands, torso, and face, the video frames are processed with Media Pipe Holistic. The relevant points will then be passed to the prediction algorithm, which will begin the prediction.
- The technology then anticipates the hand sign that is being made in real time. The expected sign will also be displayed.

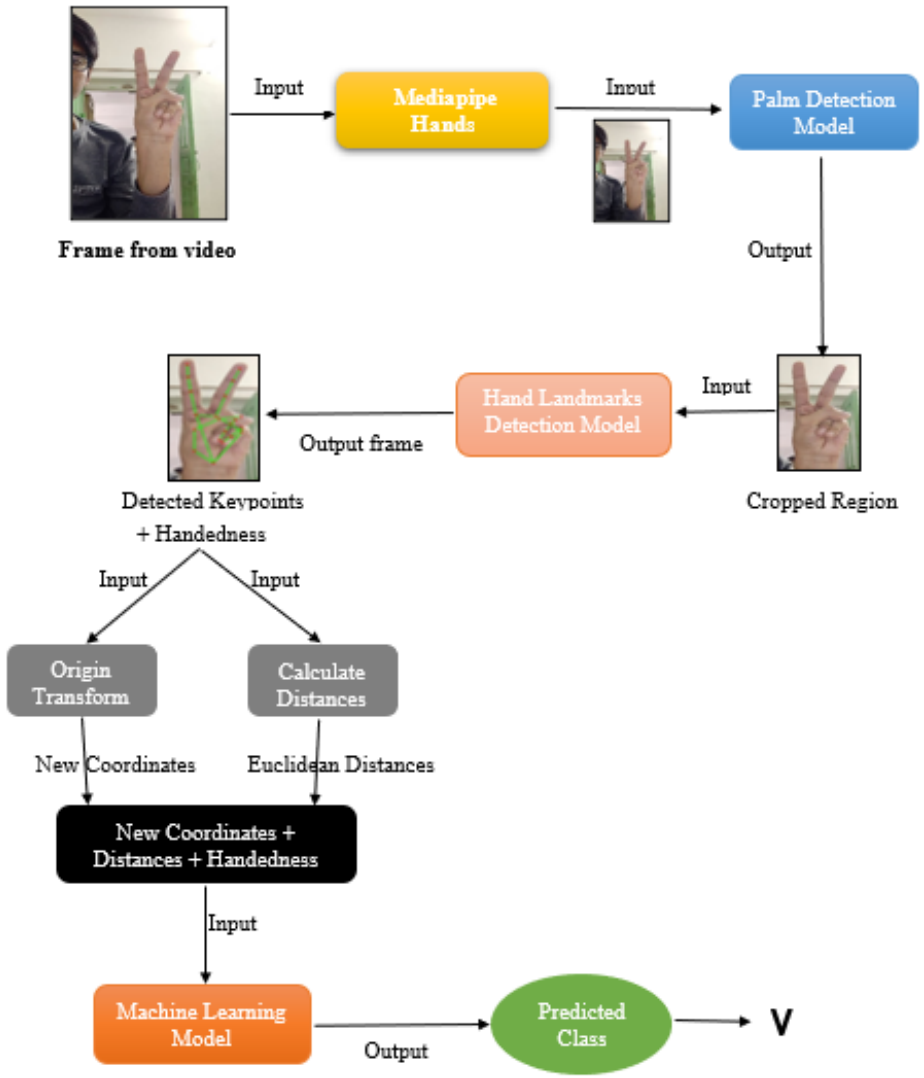

Fig : Workflow of the project in real-time

RESULTS AND DISCUSSION

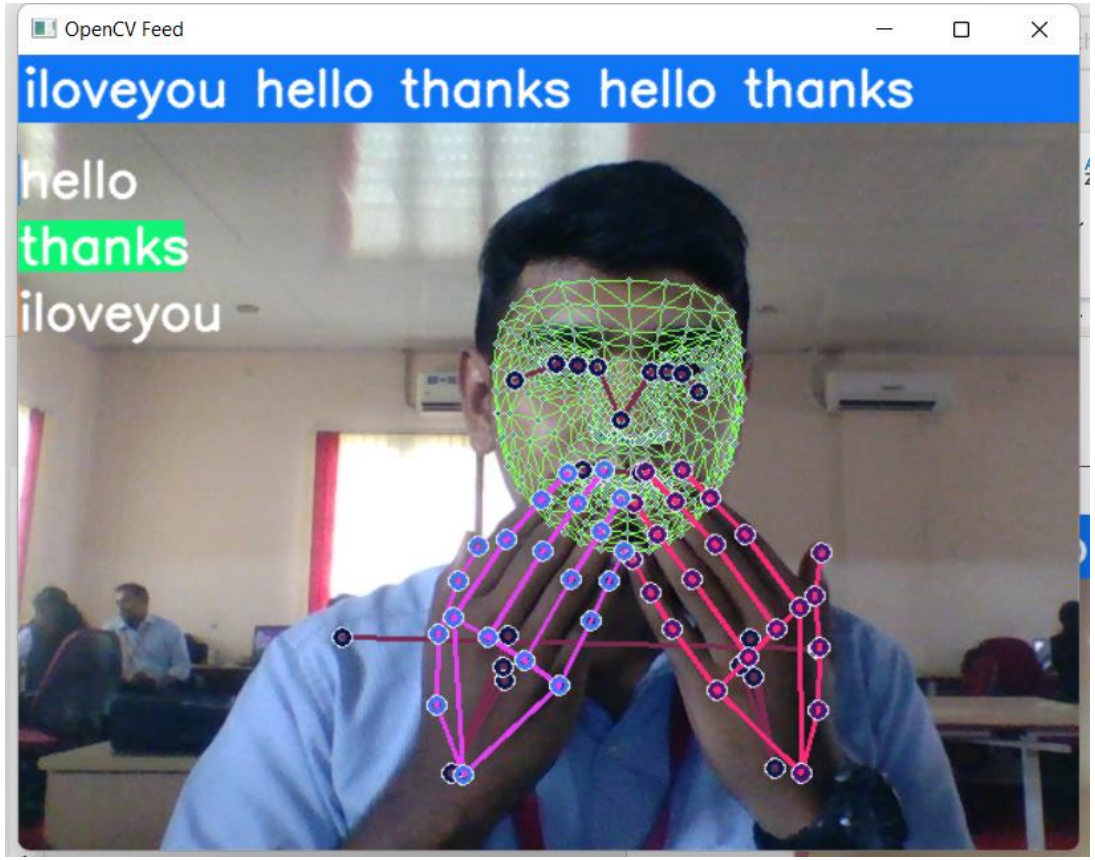

Fig 4.1 Detecting Sign Language for Thanks.

- The aim of this study was to predict signs using deep neural networks with a forearm, hand, and finger kinematics models from key points with Mediapipe Holistic.
- The Mediapipe LSTM with data augmentation produced the best results, with 91.1 (3.8) percent accuracy on the test sets.
- This sign language detector will be able to understand signs and detect hand and produce coordinators. All of the signs will be displayed in real time.

CONCLUSION

- The goal of this study was to improve upon the initial team's design and can perform live video detection and translation of people using SL with high accuracy.
- A future study will hopefully take it one step further and create a mobile application made available for download on Android and Apple platforms that can classify entire word symbols by integrating facial expressions and relative motion of the hand from the face.

REFERENCES

I. Razieh Rastgooa, Kourosh Kiania, Sergio Escalerab,, “Hand sign language recognition using multi-view hand skeleton”,Electrical and Computer Engineering Department, Semnan University, Semnan 3513119111, Iran,Department of Mathematics and Informatics, Universitat de Barcelona and Computer Vision Center, Barcelona, Spain, Received 18 September 2019, Revised 26 December 2019, Accepted 21 February 2020, Available online 22 February 2020, Version of Record 11 March 2020.

https://doi.org/10.1016/j.eswa.2020.113336

II. Rastgoo, R., Kiani, K. & Escalera, S. “Real-time isolated hand sign language recognition using deep networks and SVD.” J Ambient Intell Human Comput 13, 591–611 (2022).

https://doi.org/10.1007/s12652-021-02920-8

III. R. Bhadra and S. Kar, "Sign Language Detection from Hand Gesture Images using Deep Multi-layered Convolution Neural Network," 2021 IEEE Second International Conference on Control, Measurement and Instrumentation ,2021,pp.196-200,doi:10.1109/CMI50323.2021.9362897.

https://ieeexplore.ieee.org/document/9362897

IV. Hernandez V, Suzuki T, Venture G,“Convolutional and recurrent neural network for human activity recognition: Application on American sign language”(2020) PLOS ONE 15(2): e0228869.

https://doi.org/10.1371/journal.pone.0228869

V. Neziha Jaouedia, Noureddine Boujnahb, Med Salim Bouhlelc, “A new hybrid deep learning model for human action recognition.”, National Engineering School, Avenue Omar Ibn El Khattab Zrig Eddakhlania, Gabes 6072, Tunisia,TSSG-Waterford Institute of Technology, West Campus Carriganore Waterford, X91 P0H, Ireland,SETIT Higher Institute of Biotechnology, Soukra Road km 4 BP 261, Sfax 3000, TunisiaReceived 9 January 2019, Revised 4 September 2019, Accepted 8 September 2019, Available online 9 September 2019, Version of Record 11 May 2020.

https://doi.org/10.1016/j.jksuci.2019.09.004

THANK YOU