SIGN LANGUAGE DETECTION

B.Sc.

IN

DATA SCIENCE AND ANALYTICS

SUBMITTED BY

3ND YEAR STUDENTS

BHAVANA R	**20BSR18008**
JEEVAN S	**20BSR18014**
SHREE SHANGAAVI N	**20BSR18030**
Y ARCHANA	**20BSR18039**
YATIN KHANDE	**20BSR18040**

Table of Contents

DECLARATION OF AUTHORSHIP AND COMPLIANCE OF ACADEMIC ETHICS 3
ABSTARCT 3
INTRODUCTION 5
LITERATURE SURVEY 7
PROPOSED WORK 9
RESULTS 11
SYSTEM REQUIREMENTS 14
ADVANTAGES AND DISADVANTAGES OF SIGN LANGUAGE DETECTION 14
CONCLUSION 14
FUTURE WORK 15
REFERENCE 16

DECLARATION OF AUTHORSHIP AND COMPLIANCE OF ACADEMIC ETHICS

We, Bhavana R, Jeevan S, Shree Shangaavi N, Y Archana, Yatin Kande of second year B.Sc Honours, in the department of Data Science and Analytics from Jain University School of Science, Bengaluru, Karnataka, hereby declare that the project work entitled SIGN LANGUAGE DETECTION is carried out by us. We also declare that no chapter of this manuscript in whole or part has been incorporated in this report from any earlier work done by others or by us. However, extracts of any literature which has been used for this report has been duly acknowledged providing details of such literature in references.

Name

USN no

Bhavana R
20BSR18008

Jeevan S
20BSR18014

Shree Shangaavi N
20BSR18030

Y Archana

20BSR18039

Yatin Kande
20BSR18040

ABSTARCT

Sign language is a boon for the physically challenged people to express their thoughts and emotion. In this work, a novel scheme of sign language recognition has been proposed for identifying the alphabets and gestures in sign language. With the help of computer vision and neural networks we can detect the signs and give the respective text output.

The project aims at building a machine learning model that will be able to classify the various hand gestures used for fingerspelling in sign language which is of real time. In this user independent model, classification machine learning algorithms are trained using a set of image data and testing is done on a completely different set of data. In this sign language detection project, we create a sign detector, which detects hello, thankyou and I love you sign language. We have developed this project using OpenCV and Keras modules of python.

The steps followed here for sign language detection are

1. Extract holistic key points
2. Train an LSTM DL Model
3. Making real time predictions using sequences.

INTRODUCTION

1. IMAGEPROCESSING

Image processing is a method to perform some operations on an image, in order to get an enhanced image or to extract some useful information from it. Nowadays, image processing is among rapidly growing technologies. It forms core research area within engineering and computer science disciplines too. Image processing basically includes the following three steps:

• Importing the image via image acquisition tools.

• Analysing and manipulating the image.

• Output in which result can be altered image or report that is based on image analysis.

There are two types of methods used for image processing namely analogue and digital image processing. Analogue image processing can be used for the hard copies like printouts and photographs. Digital image processing techniques help in manipulation of the digital images by using computers.

Our concern is on Digital image processing which consists of the manipulation of images using digital computers.

The processing of digital images can be divided into several classes such as image enhancement, image restoration, image analysis, and image compression. Digital image processing is to process images by computer. Digital image processing consists of the conversion of a physical image into a corresponding digital image and the extraction of significant information from the digital image by applying various algorithms. Under the conditions that an image includes several objects, the pattern detection consists of three phases.

Phases of pattern detection:

The first phase includes the image segmentation and object separation. In this phase, different objects are detected and separate from other background. The second phase is the feature extraction. In this phase, objects are measured. The measuring feature is to quantitatively estimate some important features of objects, and a group

of the features are combined to make up a feature vector during feature extraction. The third phase is classification. In this phase, the output is just a decision to determine 3 which category every object belongs to.

2. Sign Language

It is a language that includes gestures made with the hands and other body parts, including facial expressions and postures of the body. It used primarily by people who are deaf and dumb. There are many different sign languages as, British, Indian and American sign languages. A functioning signing recognition system could provide a chance for the inattentive communicate with non-signing people without the necessity for an interpreter. It generates speech or text making the deaf more independent.

3. Sign language and hand gesture detection

The process of converting the signs and gestures shown by the user into text is called sign language recognition. It bridges the communication gap between people who cannot speak and the general public. Sign Language is the primary means of communication in the deaf and dumb community. As like any other language it has also got grammar and vocabulary but uses visual modality for exchanging information. Interesting technologies are being developed for speech recognition but no real commercial product for sign recognition is actually there in the current market. The idea is to make computers to understand human language and develop a user-friendly human computer interface.

LITERATURE SURVEY

1. TensorFlow:

TensorFlow is a free and open-source software library for dataflow and differentiable programming across a range of tasks. It is a symbolic math library,and is also used for machine learning applications such as neural networks. It is used for both research and production at Google.

2. Opencv:

OpenCV (Open-Source Computer Vision Library) is a library of programming functions mainly aimed at real-time computer vision. Originally developed by Intel, it was later supported by Willow Garage then It seez (which was later acquired by Intel). The library is cross-platform and free for use under the open-source BSD license.

3. Keras:

Keras is an open-source neural-network library written in Python. It is capable of running on top of TensorFlow, Microsoft Cognitive Toolkit, R, Theano, or PlaidML. Designed to enable fast experimentation with deep neural networks, it focuses on being user-friendly, modular, and extensible. It was developed as part of the research effort of project ONEIROS (Open-ended Neuro-Electronic Intelligent Robot Operating System), and its primary author and maintainer is François Chollet, a Google engineer. Chollet also is the author of the XCeption deep neural network model.

4. Numpy:

NumPy is a library for the Python programming language, adding support for large, multidimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays. The ancestor of NumPy, Numeric, 12 was originally created by Jim Hugunin with contributions from several other developers.

5. Neural Network:

A neural network is a series of algorithms that endeavors to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates. In this sense, neural networks refer to systems of neurons, either organic or artificial in nature. Neural networks can adapt to changing input; so the network generates the best possible resultwithout needing to redesign the output criteria. The concept of neural networks, which has its roots in artificial intelligence, is swiftly gaining popularity in the development of trading systems.

PROPOSED WORK

3.1. Import and Install Dependencies

This includes importing libraries like cv2, numpy, os, matplotlib, time, mediapipe and installing packages tensorflow and opencv2

3.2. Keypoints using MP Holistic

This aims to enable a holistic, simultaneous perception of body language, gesture and facial expressions. Its blended approach enables remote gesture interfaces, as well as full-body AR, sports analytics, and sign languagedetection.

3.3. Extract Keypoint Values

Keypoint extraction is the first task in various computer vision algorithms, where the keypoints are then stored in a database used as the basis for detection and one important thing to understand is that after extracting the keypoints, we only obtain information about the position.

Here the input data used for this detection model is a series of 30 arrays each of which contains 1662

values(30, 1662), each of the 30 arrays represents the landmark values(1662 values) in a single frame.

3.4. Setup Folders for Collection

The Action Detection takes place here which is a key difference between action detection and other computer vision tasks is that a sequence of data rather than a single frame is used for detection.

3.5. Collect Keypoint Values for Training and Testing

Here we are going to collect 30 videos per action i.e, hello, thank-you and I love you.

Then each one of those sequences are going to contain 30 frames of data. Each frame will contain 1662 landmark values i.e, 3*30 sequence, 30 frames, 1662 landmarks.

3.6. Preprocess Data and Create Labels and Features

The figure shows a collection of two-dimensional data, colored according to two different class labels and the process of cleaning raw data will be took place.

Collection Break Having breaks between each sequence collection allows you to reset and reposition yourself to collect the action from start to finish. Shorten Breaks in order to shorten the break between sequences collected change the value in cv2.waitTime() to a lower value. To increase the time increase the value.

3.7. Build and Train LSTM Neutral Network

LSTMs provide us with a large range of parameters such as learning rates, and input and output biases. Hence, no need for fine adjustments. The complexity to update each weight is reduced to O(1) with LSTMs, similar to that of Back Propagation Through Time (BPTT).

3.8. Make Predictions

This is predicting the values through observation which will help us understand the behaviour of the data.

3.9. Save Weights

This step saves training history and weights of your model

3.10. Evaluating using Confusion Matrix and Accuracy

Confusion matrices can help with side-by-side comparisons of different classification methods and the best accuracy is 1.0 and our accuracy is

```
In [118]: 1 multilabel_confusion_matrix(ytrue, yhat)

Out[118]: array([[[56,  0],
                  [ 0, 29]],

                 [[57,  0],
                  [ 0, 28]],

                 [[57,  0],
                  [ 0, 28]]], dtype=int64)

In [119]: 1 accuracy_score(ytrue, yhat)

Out[119]: 1.0
```

Here conversion takes place which mean Running these cells converts the predictions from their one-hot encoded representation to a categorical label e.g. 0,1 or 2 as opposed to [1,0,0] [0,1,0] or [0, 0, 1]

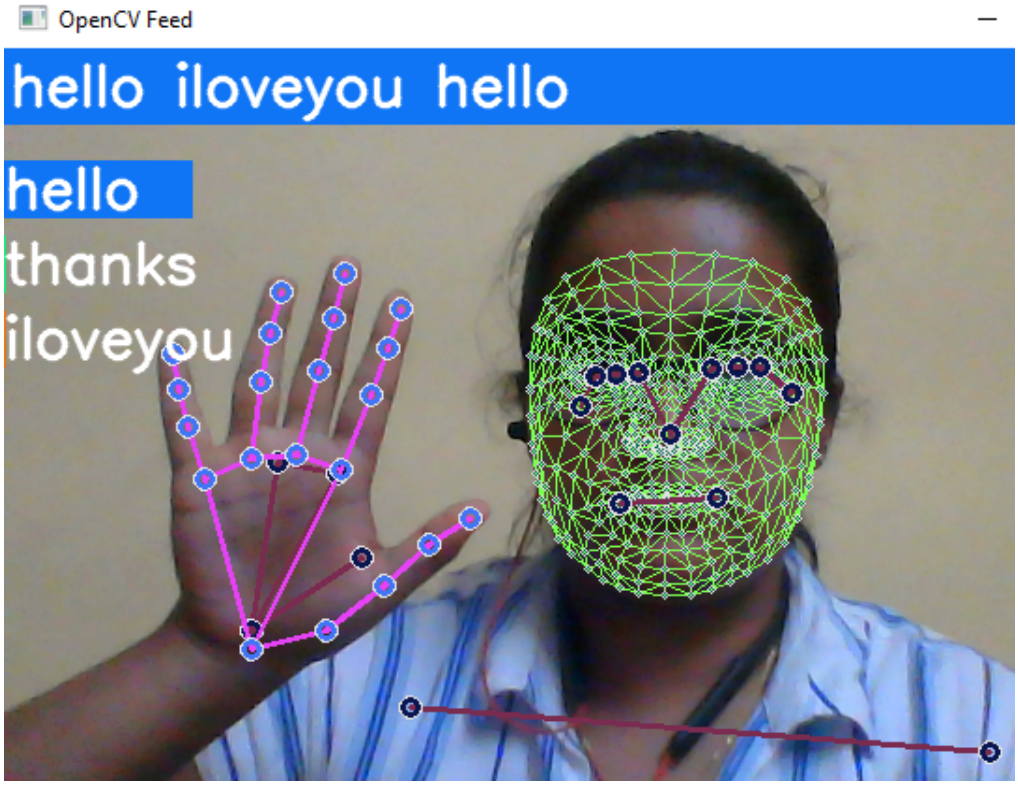
RESULTS

Test in Real Time

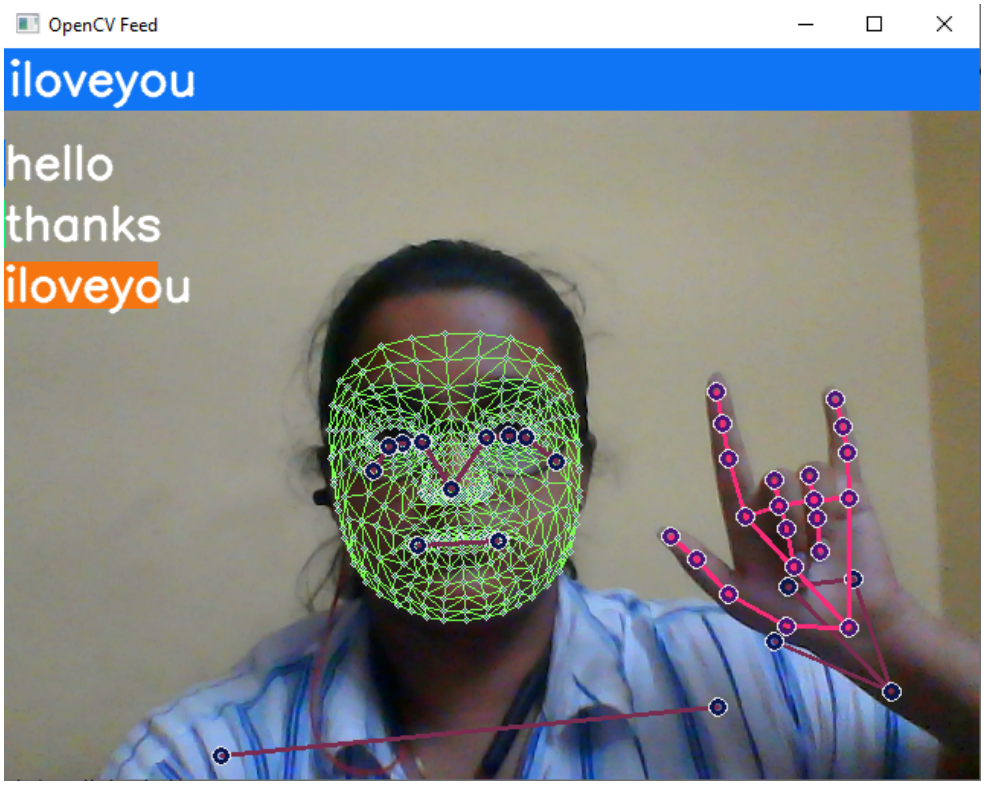
Upon training the image dataset without any augmentation, the training accuracy achieved and the real time performance was also up to the mark with 1.0 accuracy.

Here in real time when we show sign language it detects the sign which we are showing and colours on the labels i.e., Hello, Thank- you and I Love You in different colours and the pictures screenshot will takes place in OpenCV feed. For Hello we labelling will be highlighted through blue colour, thanks throughgreen colour and I Love You through red colour.

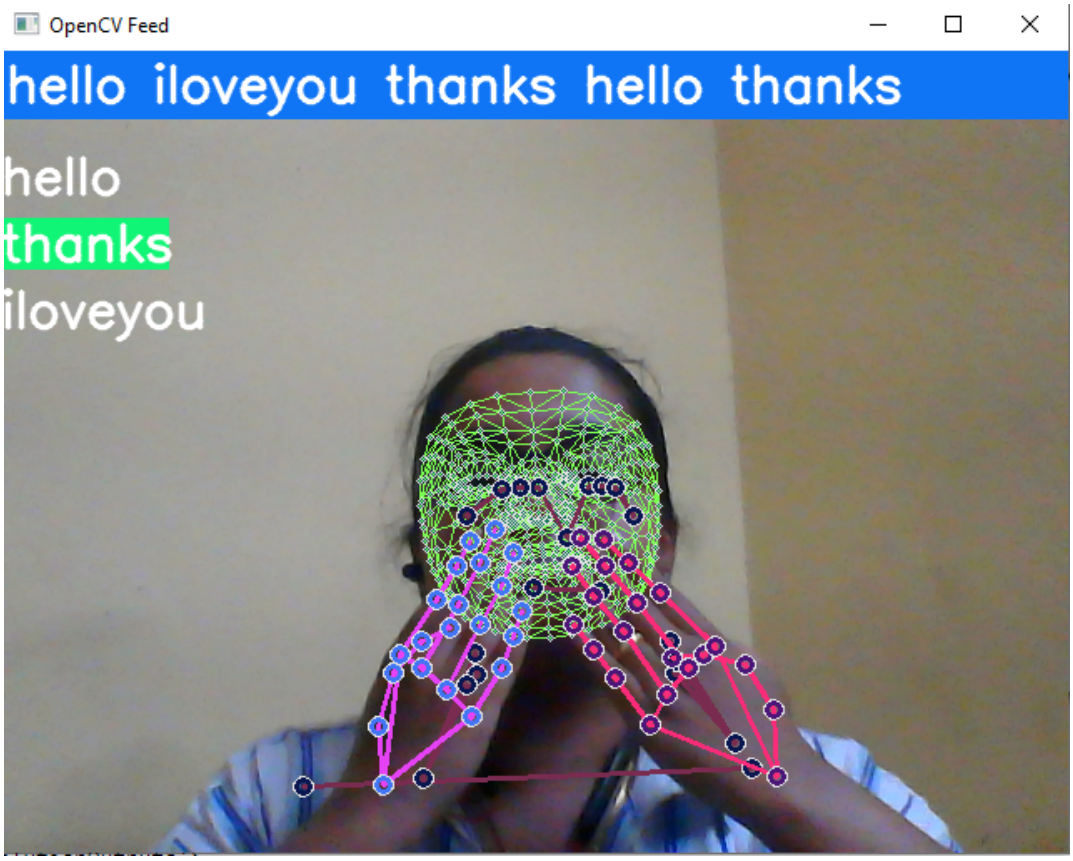
Screenshots Of Results



Hello sign picture highlighted with blue colour



I Love you sign picture highlighted with red colour



Thanks sign picture highlighted with green colour

SYSTEM REQUIREMENTS

- Python
- IDE (Jupyter)
- Numpy
- cv2 (openCV)
- Keras
- Tensorflow (as keras uses tensorflow in backend and for image preprocessing)
- Mediapipe

ADVANTAGES AND DISADVANTAGES OF SIGN LANGUAGE DETECTION

1. **Advantages**

   It provides an efficient and accurate way to convert sign language into text or voice has aids for the hearing impaired for example or enabling very young children to interact with computers (recognizing sign language). It enriches and enhances children's cognitive processes, leading to higher abstract and creative thinking, better problem-solving skills, greater cognitive flexibility, better listening skills, greater academic achievement, and much more. It also promotes cultural awareness, literacy, and other intellectual benefits.

2. **Disadvantages**

   The disadvantages are that they are costly and are difficult to be used commercially. Classification methods are also varying from researchers. Researchers tend to develop their own concept, based on known methods, to give better result in recognizing the sign language.

CONCLUSION

The applications need several kinds of images as sources of information for elucidation and analysis. Several features are to be extracted so it as to perform various applications. When an image is transformed from one form to another such as digitizing, scanning, and communicating, storing, etc. degradation occurs. Therefore, the output image as to undertake a process called image enhancement, which contains of a group of methods that seek to develop the visual presence of an image. Image enhancement is fundamentally enlightening the interpretability or awareness of information in images for human listeners and providing better input for other automatic image processing systems. Image then undergoes feature extraction using various methods to make the image more readable by the computer. Sign language recognition system is a powerful tool to prepare an expert knowledge, edge detect and the combination of inaccurate information from different sources. The intend of convolution neural network is to get the appropriate classification.

FUTURE WORK

The proposed sign language recognition system used to recognize sign language letters can be further extended to recognize gestures facial expressions. Instead of displaying letter labels it will be more appropriate to display sentences as more appropriate translation of language. This also increases readability. The scope of different sign languages can be increased. More training data can be added to detect the letter with more

accuracy. This project can further be extended to convert the signs to speech.

REFERENCE

[1] https://www.academia.edu/download/31388985/D0361015021.pdf

[2] https://www.academia.edu/download/31388985/D0361015021.pdf

[3] https://link.springer.com/chapter/10.1007/978-3-030-72073-5_9

[4] https://data-flair.training/blogs/sign-language-recognition-python-ml-opencv/

[5] https://arxiv.org/ftp/arxiv/papers/2201/2201.01486.pdf