**INTRODUCTION TO DATA SCIENCE LAB**

**[CSL-487]**

**Project Name:**

**Hand Gesture Recognition for ASL**

**SEMESTER PROJECT**

**Maximum Marks: 30**

**Submission Due Date: 17th January 2023**

|  |  |  |  |
| --- | --- | --- | --- |
| **Sr.no** | **Name** | **Enrolment** | **Semester** |
| **01** | **Ayesha Shahzad** | **02-134201-083** | **BS(CS)-6B** |
| **02** | **Sadia Urooj** | **02-134201-081** | **BS(CS)-6B** |
| **03** | **Syeda Aqsa Ashraf** | **02-134201-076** | **BS(CS)-6B** |

|  |  |
| --- | --- |
| **Designation** | **Designation** |
| **Course Instructor** | **Ms. Soomal Fatima** |
| **Lab Engineer** | **Ms. Salas Akbar** |

**Acknowledgement**

We would like to express our special thanks of gratitude to our teacher Miss Salas Akbar and Ms. Soomal Fatima for giving us the opportunity to work on this project, as computer vision is one of the most developing fields of AI and is currently used for a variety of purposes to assist humans. While performing this project we got to know how to implement CNN and OpenCV for Real-time image classification. it would not have been possible without the kind support and help of our teachers and family. My appreciations also go to my group members in developing the project and people who have willingly helped us out with their abilities.

**Contents**

**1. Chapter 1**

**1.1. Problem Statement …………………………………………… 4**

**2. Chapter 2**

**2.2. Literature Review ……………………………………………... 4**

**3. Chapter 3**

**3.1. Methodology …………………………………………………… 5**

**3.1. Scope ……………………………………………………………**

**4. Chapter 4**

**4.1. Code Snippet ………………………………………………….. 8**

**5. Chapter 5**

**5.1. Conclusion ……………………………………………………. 16**

**5.2. Future Enhancement ………………………………………… 16**

**6. Chapter 6**

**6.1. References ……………………………………………………. 16**

**1. Chapter 1**

### 1.1. Problem Statement

## There are many applications where hand gestures are used for interaction with the system like videogames and medical equipment etc. These hand gestures can also be used by handicapped people to interact with the systems. Classical interaction tools like keyboard, mouse and touchscreen, etc., may limit the way we use the system. Gestures can interpret the same functionality without physically interacting with the interfacing devices.

## The problem lies in the understanding of these gestures by machine as for different people, the same gestures may look different. This problem can be resolved by the use of Deep learning approaches. Convolution neural networks (CNN) are providing to be the ultimate tool to process the image data for these systems.

1. **Chapter 2**

**2.1. Literature Review**

Zaki, M.M., Shaheen, S.I.: Sign language recognition using a combination of new vision based features. Pattern Recognition Letters 32(4), 572–577 (2011) [1]

It was published on March 2022, by Zaki, M.M., Shaheen, S.I.:Worked on American Language detection is presented in which they have used a vocabulary set of 30 words. The system was appearance-based recognition with a hand tracking recognition to classify. They used hidden Markov Model. An error rate of 10.91% is achieved on the RWTH-BOSTON-50 database.

Cooper, H., Ong, E.J., Pugeault, N., Bowden, R.: Sign language recognition using sub-units. The Journal of Machine Learning Research 13(1), 2205–2231 (2012) [5]. It was published on 1st July 2012, by Cooper, H., Ong, E.J., Pugeault, N., Bowden, R. The approach in this paper uses the Microsoft Kinect to extract appearance-based hand features and track the position in 2D and 3D. The classification results are obtained by comparing a hidden Markov model (HMM) approach with sequential pattern boosting (SP-boosting). This resulted in an accuracy of 99.9% on 20 different isolated gestures on their specifically constructed data set and 85.1% on a more realistic one with 40 gestures.

Another approach that we found most suitable for accuracy in quick measure was from Kazuhito Takahashi hand-gesture-recognition-using-mediapipe[6], In this approach he used mediapipe library from google, we take the hand landmark detected from the mediapipe library and preprocess them for the ML model which a Convolutional Neural Network and then in then real time without images we classify the preprocessed values.

**3. Chapter 3**

### 3.1. Methodology

Our Hand Gesture recognition model is implemented in four steps:

**Data Collection:**

In this step, we collected the real time image data of the hand using the mediapipe hands library for collecting the the landmarks of the hand from a single frame. The data is then stored in a csv file with the respective class of an alphabet, with its specific landmarks points with each ones x and y axis orintation.

**Preprocessing:**

In preprocesing through mediapipe we first take the landmarks of the hand save it in a list and than take out the absolute value of the x, y landmarks taken by the mediapipe. After acquiring the x, y values we tranform it into the relative value for the model. When

**Training:**

In this step we will implement the Convolutional neural network (CNN) model on the preprocessed data. In this state all the layers to be implemented on the data is defined according to the

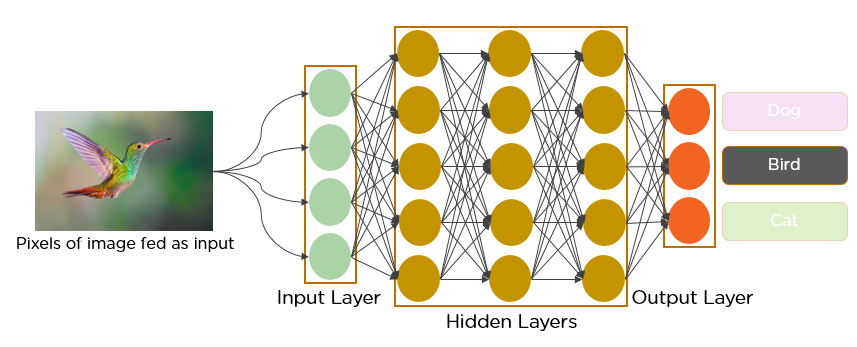


Figure 1: Convolutional neural network (CNN)

Tracking of people provides a significant

Advantage to identification since we can apply the

Principle of continuity of identity [2]. This says that, while

we may only be able to identify a person occasionally

(such as when we have a good view of their face, when

they swipe an ID badge, or when they speak into a

Telephone), if wecan reliablytrack the person, we know

That all identifications associated with the track relate to

The same person and applythroughout the track. Several

(Fallible) identification methods applied at different times

And places can thus be combined and corroborated.

Tracking of people provides a significant

Advantage to identification since we can apply the

Principle of continuity of identity [2]. This says that, while

we may only be able to identify a person occasionally

(such as when we have a good view of their face, when

Tracking of people provides a significant

Advantage to identification since we can apply the

Principle of continuity of identity [2]. This says that, while

we may only be able to identify a person occasionally

(such as when we have a good view of their face, when

**Prediction:**

Using the trained model, we predict the hand gesture and predict the accuracy of the model.

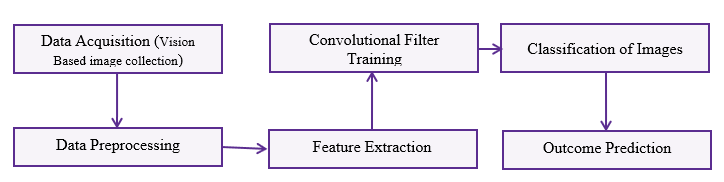


Figure 2: Process of the model training and outcome generation.

### 3.2. Scope

Hand gesture recognition is one of the active research areas in the field of human-computer interface due to its flexibility and user friendliness. The gesture recognition technique is used to develop a system that can be used to convey information among disabled people or for controlling a device. In this project we will be building a model on sign language detection that takes the dataset in real time from the camera and then apply the required preprocessing techniques to filter out all the unnecessary information from the image data after that the pre-processed data is passed through a classifier which predicts the class of the hand gestures.

***Project Objectives:***

The main motives of our project are as follows:

* To build a model that can efficiently collect the data from the camera in real-time.
* To make effective uses of CNN to training the required model for prediction.
* To get approximate predictions from the trained model.

### 

**4. Chapter 4**

### 4.4. Code Snippet

**Training Code:**



Figure 3 CNN model building

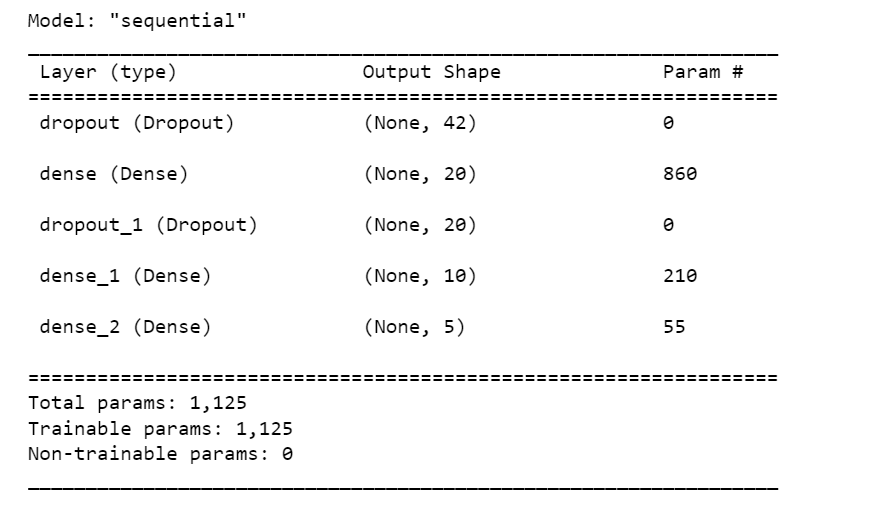


Figure 4 Total parameter after addind the layers

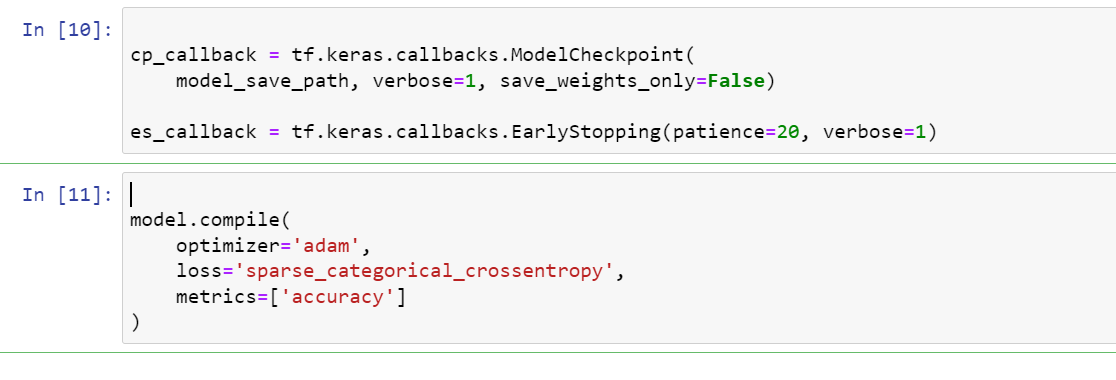


Figure 5 Compiling model for to fit

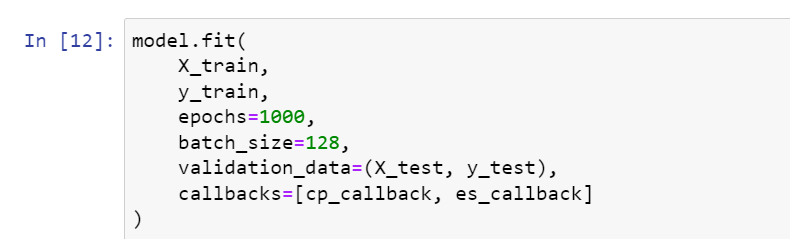


Figure 6 Fitting the train set for model

**Testing Code:**

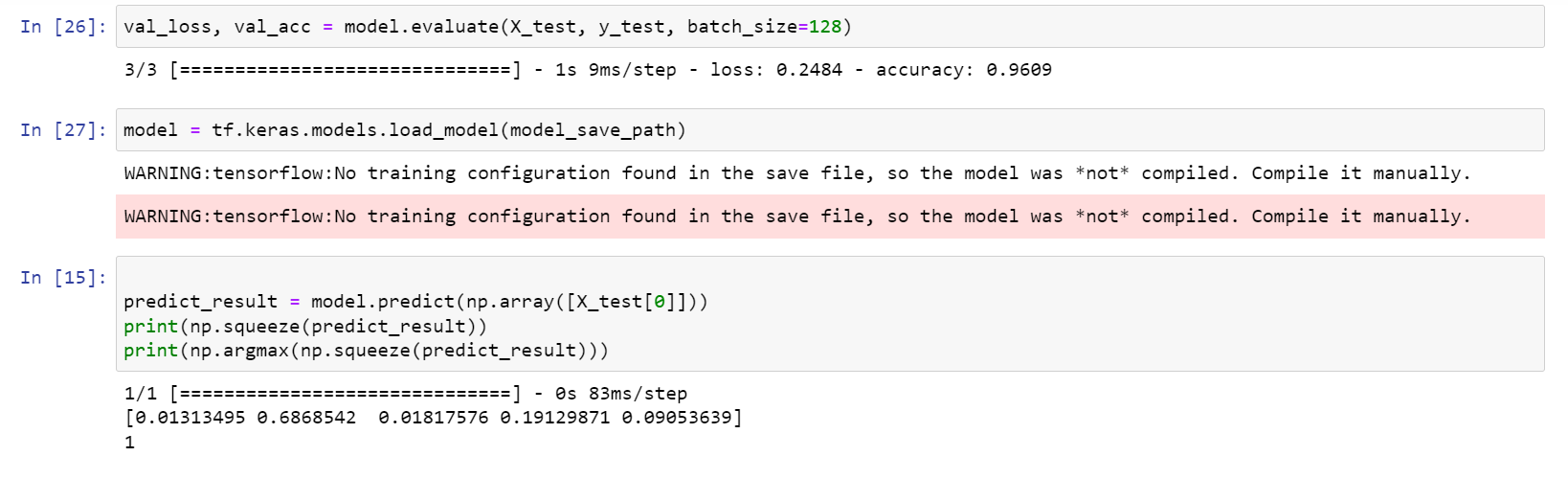


Figure 7 Result Prediction

**Output:**

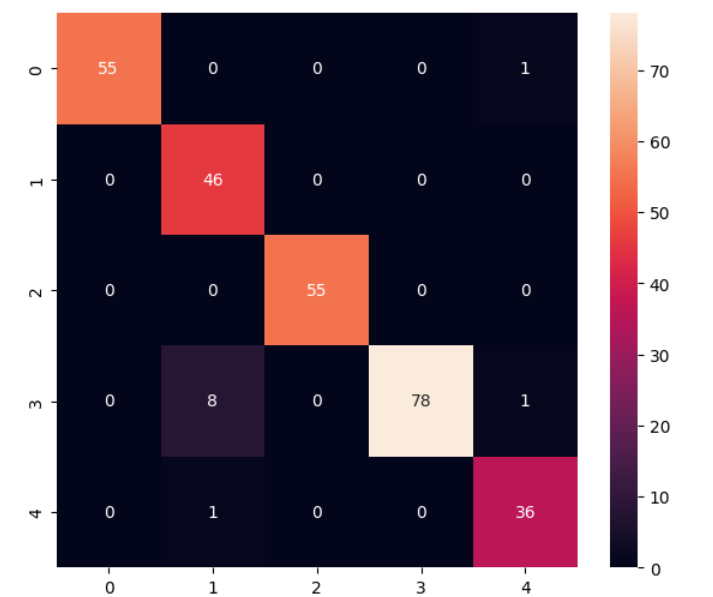


Figure 8 Heat Map of confusion matrix

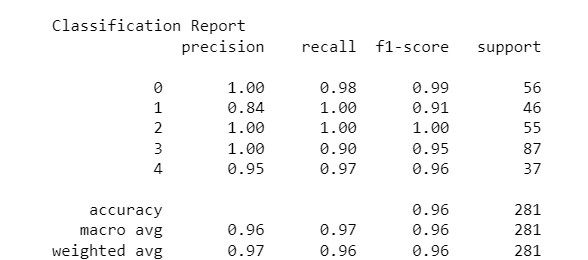


Figure 9 Classification Report

# 5. Chapter 6

## 5.1. Conclusion

In this project, we have implemented hand gesture recognition to detect American sign language(ASL) with Convolutional Neural Networks(CNN) and OpenCV, We have acquired a dataset in real-time consisting of different hand gestures representing a letter in ASL then the data acquired was preprocessed and features have extracted, then the extracted data was trained with convolutional neural network algorithm (such as SqueezNet) to achieve maximum accuracy, especially in the gestures that are similar to each other.

### 5.2. Future Work

We are planning to achieve higher accuracy even in case of complex backgrounds by trying out various background subtraction algorithms. We are also thinking of improving the preprocessing to predict gestures in low light conditions with a higher accuracy.

### 5.3. Tools

* **Google Colab**

We have used Google Collaboratory for our project, popularly known as Colab, which is a web IDE for python.

* **Tensor Flow**

Tensorflow is an open source software library for numerical computation. TensorFlow is widely used in Machine Learning.

* ***OpenCV***

OpenCV(Open Source Computer Vision) is an open source library of programming functions used for real-time computer-vision. It is mainly used for image processing, video capture and analysis for features like face and object recognition. It is written in C++ which is its primary interface, however bindings are available for Python, Java and etc.

**6. Chapter 6**

# 6.1. References

[1] Zaki, M.M., Shaheen, S.I.: Sign language recognition using a combination of new vision-based features. Pattern Recognition Letters 32(4), 572–577 (2011)

[2]<https://gongster.medium.com/how-does-a-neural-network-work-intuitively-in-code-f51f7b2c1e3f>

[3]https://medium.com/nybles/a-brief-guide-to-convolutional-neural-network-cnn-642f47e88ed4

[5] Cooper, H., Ong, E.J., Pugeault, N., Bowden, R.: Sign language recognition using sub-units. The Journal of Machine Learning Research 13(1), 2205–2231 (2012)

[6]https://github.com/Kazuhito00/hand-gesture-recognition-using-mediapipe/blob/main/README\_EN.md