***Sign Language Recognition using Long Short-Term Memory and Locking System***

By Saddam Ahmed Nisar (192101024) and Muhammad Touseef (192101002)

Supervised by: Sir Dr. Usman Ali Gulzari

KICSIT Sub Campus of IST

**Abstract**

In our society, a lot of people are unable to talk and speak and they have a problem to communicate with other peoples in society. To keep this problem in mind, we propose a sign language recognition system to overcome this problem. In present time, machine learning is a technique that authorize the system to acquire knowledge themselves. In now a day, it is vastly used. One of the application of machine learning is hand gesture recognition which works on the basic principles of sign language. In this study we proposed Sign Language Recognition using Long Short-Term Memory and Locking System (SLR-LSTM-LS) for special persons. Sign language is basically a communication through hand gesture **[7]**. It is a natural language that help the persons who are unable to listen or talk. It also helpful for normal person to communicate with person who are unable to talk and listen. As in the world we are facing major problem of disability of hear and talk. The only thing that separate the people who are unable to speak and talk and normal people is communication. The major problem is the communication of disable person who unable talk and listen. Insight of this, we are now making a sign language android app which help them to communicate with others peoples easily because they have right to talk. In previous work, they have a lot of accuracy problems and glitches using CNN **[3]**. Their accuracy was from 68% to 97% **[3]**. And Our proposed work (SLR-LSTM-LS) remove glitches and achieve accuracy from 90% to 99%.

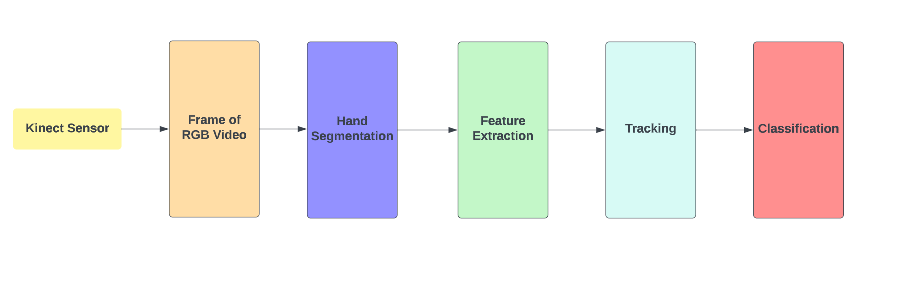
1. **Introduction:**

Artificial intelligence is a development of system that work like human through machine learning **[1]**. It helps the system to think like human. It simplifies the effort of human and help to make decision **[1]**. Machine learning is a sub domain of artificial intelligence **[2]**. It allows the system to learn themselves and from experience **[2]**.Hand gesture is the application of machine learning and it is basically movement of hand to express the word or feelings. By using hand gesture, we can convey our message without speaking or the person who are unable to listen can understand the word or message **[6]**. Sign language is basically communication through body is specially with hand and arm **[7]**. It connects the special persons with social environment. It is used to convey message through sign instead of speak. We use hand gesture to represent numbers, letters, words and sentences, which allows them to communicate among themselves. One hand consist of twenty-one landmarks and by controlling these landmarks we can set the word in sign **[8]**. By using this App, we can easily learn sign language through this app. It will help to communicate with person who are unable to listen and unable to talk. Sign language is important for disable people who are unable to speak and talk. Normal people ignore the importance of sign language but it is loved one for people who are unable to talk and speak. It is the only way for them to communicate with other people. It also for normal persons who are unable to understand the sign when talking with the person who are unable to speak and talk. They can easily learn the sign language from this application to understand what the unspeakable person want to speak, and if normal person wants to ask something from the person who are unable to speak and listen they can use this language to communicate with them.

1. **Literature review:**

# . Hand recognition by using Kinect sensor [3]:

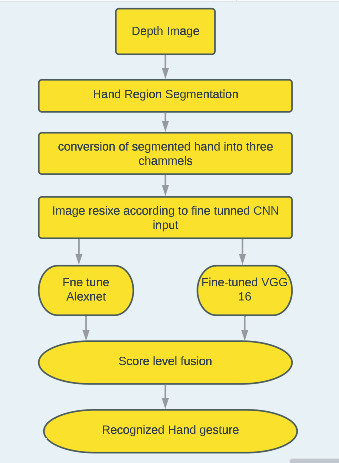
Different researchers researched on hand gesture recognition to overcome the challenges in hand gesture recognition. the main challenge is in accuracy, image acquisition and feature extraction and classification of gestures **[3]**. The researcher works on these problem in classification of gesture and features extraction by using vision based recognition and by using Kinect sensor **[3]**. The main aim of them is to check that the how well the vision based recognition recognize movement of hand and what how well it will be accurate **[3]**. First sensor sense the image then sends the frame of RGB image or video to segmentation process after that extract the features of hand and then tracking the hand and then classify it and then recognize the hand on image or video. After overcome the problems in feature extraction and image acquisition the accuracy of hand gesture is 68 to 97 percent and average accuracy is 86.6 percent and it recognize all type of movement of hand **[3]**. They made working prototype to perform real-time gesture and it has pressure sensor and have computational framework to make it recognition better. by using sensor, it recognizes gestures.



**Figure 1.** **[3]** Standard framework of hand gesture recognition using Kinect sensor

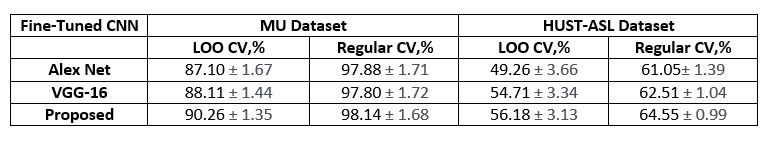
**2.2. Hand gesture recognition using convolution neural network (CNN) [4]:**

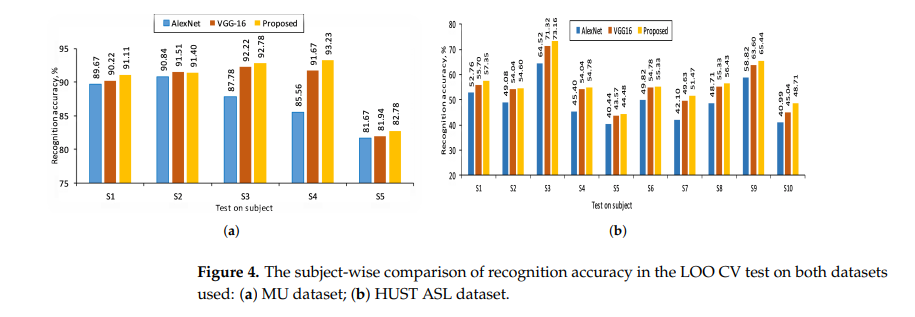
In this system hand gesture recognition is done by using convolution neural network (CNN) **[4]**. CNN has better results in recognition of classification problem of image. In this system the recognition of hand gesture is passed through different process. First is data acquisition, pre-processing and then recognition. In this hand gesture system is developed by using Kinect v2 depth sensor **[4]**. Because in this, from the frame image the hand region is more accurately segmented and in preprocess the it separates the hand area from hand movement image then make this image size equal to the input image size of CNN. Then it will take RGB and depth image from Kinect depth camera **[4]**. Then it takes only depth image for segmentation and after that apply filter process to discover hand region and get out the noise section of segmented image and make bounding box. As CNN works with three channel output and then normalize image to make it one channel **[4]**. After that it passes through recognition process with fined tuned models of CNN, Alex Net and VGG16 for classification. Alex Net and VGG16 both are deep architecture of CNN **[4]**. There are two datasets involved in this work one is HUST-American sign language dataset and other is MU dataset AND then train and test these datasets and recognize hand gesture according to these datasets, and these use two techniques LOO CV AND Regular CV **[4]**. The proposed techniques have accuracy of hand gesture is 64 and 98.14 percent as shown in figure 3 by graph **[4]**. The LOO CV have better performance as compared to Regular CV as shown in Table 2 **[4]**. LOO CV is user independent because this performance of trained model is measured by samples of hands **[4].**



**Figure 2.** **[4]** Real-time gesture recognition development flowchart

**Table 2**. **[4]** Accuracy of both MU Dataset and HUST-ASL Dataset



;

**Figure 3. [4]**The subject-wise comparison of recognition accuracy in the LOO CV test on datasets used: (a)MU dataset;(b)HUST ASL dataset

**Proposed System:**

1. **Methodology:**

Our proposed system is basically a sign language hand gesture recognition application. firstly, we make our own dataset of hand gestures by taking image of hand and train it by using LSTM algorithm. LSTM woks in layers to recognize pattern for best performance. First we adopted the hand detector to recognize the hand gestures to generate our dataset for hand recognition gestures.

**3.1. Data collection:**

Different cameras are available to take images. The analysis of camera with advantage and limitation are displayed in Table 1.the table shows that the camera is free of cost the user can use it to take image. First, we take hand gesture images from webcam, the images are RGB in blue color bounding box then after that it detect the coordinates of hand until the coordinates are correct. After the Complete acquisition of hand gesture, the image of hands gesture is passed through the pre-processing process.

**Table 1.** Analysis of components used for hand recognition

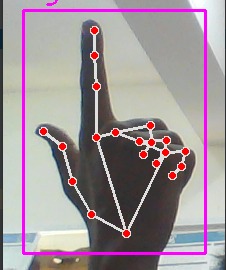
|  |  |  |
| --- | --- | --- |
| **Data acquisition component** | **Advantage** | **Limitations** |
| Webcam | Free to use | Affect by background |

**3.2. Pre-processing process:**

in this process, the images are basically being in frames. This process will detect the coordinates of hand in a window that we made. Then subtract the area of hand according to the coordinates show in the image of hand from the background and then resize the image of hand according to the input size of image of pre-trained LSTM.

**3.3. Extraction of feature and classification of hand gesture:**

After the pre-processing process, the input image in preprocess module is in raw form then it pass through the feature extraction to extract the features of the image and then convert it into array and then make dot mpy file format after that, train the dot mpy file of image with dataset that we make earlier and then recognize the sign, then recognize the sign and check the coordinates as shown in figure 4. If sign is not correct this show that the sign is invalid and go back to take image and detect the hand gesture again. If the is correct the pass through the lock process.

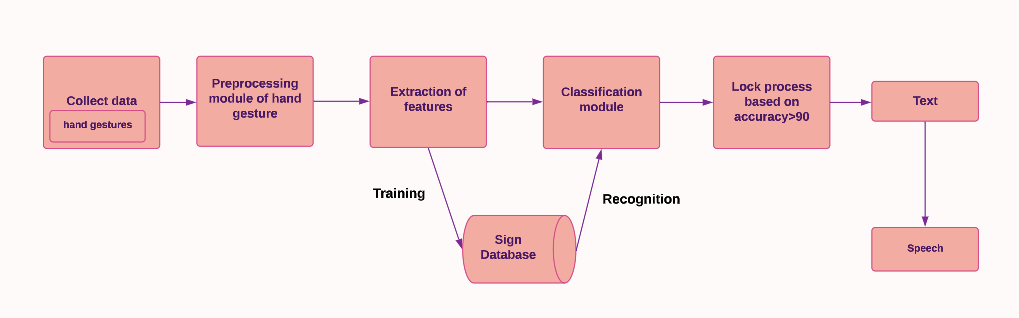


**Figure 4.** Feature extraction and recognition of hand gesture

**3.4. Lock process:**

By keeping in view of accuracy problem in existing system and to remove glitches in hand gesture recognition we develop a lock system on hand gesture. This system helps us to make accuracy better. In this system we make a condition that if the accuracy of sign is greater than 90 than lock the sign otherwise take detect and match the sign until the accuracy greater than 90. If the sign gives the accuracy greater than 90 than the process locks the sign and show right and show the text of the sign and show blue color line on correct text and also and also give accuracy of sign. And when the text is displayed, then speech the text until the hand is shown in the window.

**3.5. Block diagram:**

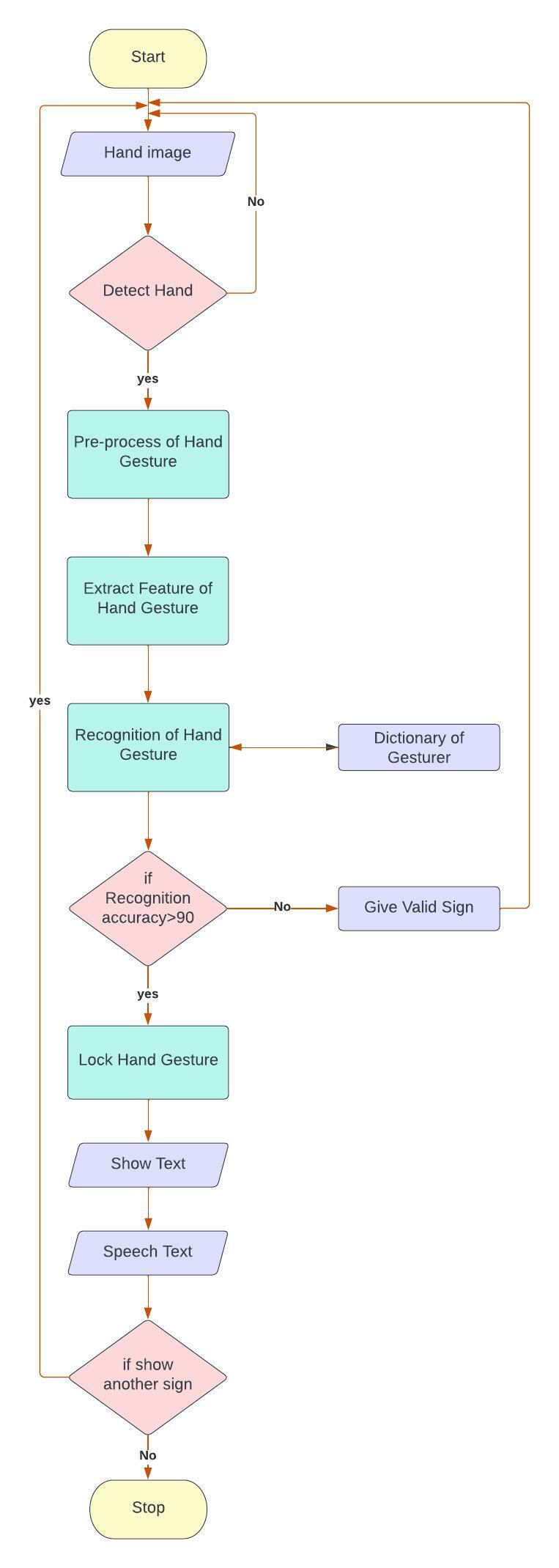


**Figure 5.**  Hand Gesture Recognition with lock system using webcam

**3.6. FLOW CHART:**

**3.6.1 Description:**

As show in figure 5 we have real time hand gesture image then detect the hand and check the coordinates, as we make a bounding box of blue where we show the sign if sign is not in the bounding box it continuously detects the hand gestures and if the hand is in the region of blue color bounding box then it passes through the pre-processing process. In this it detects the coordinates of hand and remove the background of image only take hand coordinates area. After preprocessing process, extract the feature of hand gesture and change it into array and make dot mpy file of image. After that, recognize the hand gesture from trained model. If the recognition accuracy greater than ninety, then lock the sign and show text of sign and then speech the text until another sign otherwise show the sign is invalid and then again take image of hand.

****

**Figure .6** Flow chart of hand Gesture Recognition using lock process

1. **Pre-trained LSTM:**

The long short-term memory algorithm is basically a sub domain of RNN (recurrent neural network) **[9]**. This algorithm is basically for problems like speech to text and text to speech recognition and feature extraction, preprocess problems. This algorithm is capable to process entire sequence of data. It has better performance. So by keeping in view of its performance in accuracy we use this algorithm in our proposed work.

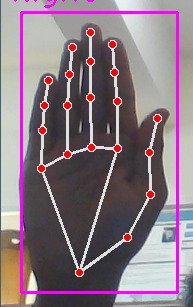
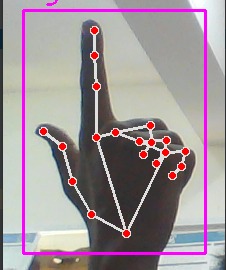
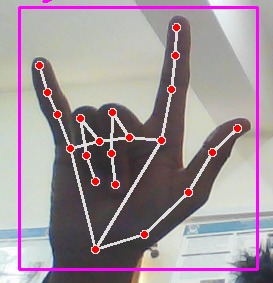
1. **Evaluation:**

**5.1. Dataset:**

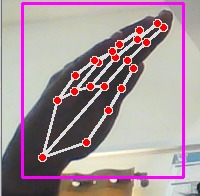
We use our own dataset to check the productiveness of our preferred system. We made our own dataset and we use sequential model to train the dataset. We take 5 signs for different text for recognition as shown in figure 8. Then we make collect images module to collect images of five gestures as shown in figure 7. The collect module takes 30 images of every gesture of hand. Then we made data module to convert these images into dot-npy format. As the image we take was in raw format. This module extracts the features of gestures and convert into dot-npy format.in this way we can easily process these images. Then we make train model module that inputs the npy format images in sequential model. After that sequential model perform training in given data and produce two files one is model.h5 and second is model.h, both files contain the trained supervised learning which meaning these files contains label data. By using this dataset, we evaluate the effectiveness of our proposed system

****

**Figure 7.** Take image of gesture to train model

 ** **

**(a) (b) (c)**

** ** 

**(d) (e) (f)**

**Figure 8.** Signs for Recognition. (a) Thank you (b) How are you (c) Nice to meet you

(d) Hello (e) What is your name (f) Invalid sign

**6. Result and Analysis:**

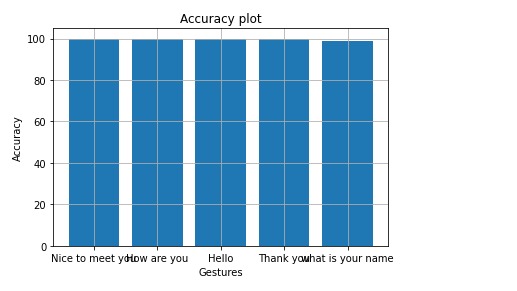
In our proposed system we use webcam and performed task using intel core i7 2.60GHz CPU with 8GB RAM. The main evaluation of our proposed system is accuracy. Accuracy is basically a rate of correct data samples over all data samples **[5]**. The accuracy formula to find accuracy is **[5]**.

**Accuracy**= TN+TP/TN+FP+TP+FN

The accuracy is basically depending upon dataset if the dataset is unstable the accuracy may not be better but if the dataset is stable then it gives better accuracy **[5]**. The accuracy of different gesture or sign is shown in Table 2. In table we have accuracy of different sign that are locked by locking process. As the mean accuracy of our proposed system is , and it is better than the accuracy of vision based hand recognition which is from 68% to 97% and with the mean accuracy of 86.6% **[3].** And the accuracy of different gestures is mapped in graph as shown in Figure 9.

**Table 2.** Accuracy of different signs or hand gestures

|  |  |
| --- | --- |
| **Signs/ gestures** | **Accuracy** |
| Nice to meet you | 99.9% |
| How are you | 99.93% |
| Hello | 99.9% |
| Thank you | 99.8% |
| What is your name | 99% |



**Figure 9.** Accuracy Graph of different Signs or Hand Gestures

1. **Goals, objective and Equipment:**

**7.1. Objective:**

This app will help special persons to deliver their message properly. This app will help normal person to understand special person message after learning from about sign and their meaning from this app. The difficulty for special person will decrease to deliver their message.it will provide us a platform to judge their hand gestures and getting information of words. This app will covert our hand gestures into words.

**7.2. Goals:**

The main goal is to develop a platform/app for special persons who can’t speak. It provides a way for special persons to communicate and also for normal persons to understand sign language. This app convert will convert hand gestures to text through yolo algorithm.

* 1. **Benefits:**
* Sign language is for everyone
* It is helpful for children to remember words
* It helps the deaf people to communicate with other people
  1. **Tools and Equipment:**
     1. **Language:**
     2. Python
     3. **Software:**
     4. Pycharm
     5. Android studio

1. **Conclusion:**

The sign language app is generating for special person who are unable to speak and it is very difficult for them to deliver their message also other persons cannot understand them .so this app will be able to extract their hand gestures and compare them with the store information and extract that what they want to say. In this way, both will be satisfied. In future this system is more powerful in solving the problem of extracting features of frames, and more gestures are recognized for further use in future.

**9. References:**

[1]Frankenfield, J. (2022) *Artificial Intelligence: What it is and how it is used*, *Investopedia*. Investopedia. Available at: https://www.investopedia.com/terms/a/artificial-intelligence-ai.asp (Accessed: January 12, 2023).

[2]Selig, J. (2022) *What is machine learning? A definition.*, *expert.ai*. Available at: https://www.expert.ai/blog/machine-learning-definition/ (Accessed: January 12, 2023).

[3]Al Farid, F., Hashim, N., Abdullah, J., Bhuiyan, M. R., Shahida Mohd Isa, W. N., Uddin, J., ... & Husen, M. N. (2022). A Structured and Methodological Review on Vision-Based Hand Gesture Recognition System. *Journal of Imaging*, *8*(6), 153.

[4]Sahoo, J. P., Prakash, A. J., Pławiak, P., & Samantray, S. (2022). Real-Time Hand Gesture Recognition Using Fine-Tuned Convolutional Neural Network. *Sensors*, *22*(3), 706.

[5]B, H.N. (2020) *Confusion matrix, accuracy, precision, recall, F1 score*, *Medium*. Analytics Vidhya. Available at: https://medium.com/analytics-vidhya/confusion-matrix-accuracy-precision-recall-f1-score-ade299cf63cd (Accessed: January 12, 2023).

*[6] Gesture definition & meaning* (no date) *Dictionary.com*. Dictionary.com. Available at: https://www.dictionary.com/browse/gesture#:~:text=A%20gesture%20is%20a%20movement,to%20say%20hello%20to%20someone. (Accessed: January 12, 2023).

*[7] What is sign language?* (no date) *What is sign language? | AccessComputing*. Available at: https://www.washington.edu/accesscomputing/what-sign-language#:~:text=Sign%20language%20is%20manual%20communication,gesture%20is%20called%20a%20sign (Accessed: January 12, 2023).

*[8]Gesture recognition guide for web  |  mediapipe  |  google developers* (no date) *Google*. Google. Available at: https://developers.google.com/mediapipe/solutions/vision/gesture\_recognizer/web\_js#:~:text=The%20recognized%20gesture%20categories%20of%20the%20detected%20hands.&text=There%20are%2021%20hand%20landmarks,x%20%2C%20y%20and%20z%20coordinates. (Accessed: January 12, 2023).

[9]says:, K.L. (2022) *What is LSTM - introduction to long short term memory*, *Intellipaat Blog*. Available at: https://intellipaat.com/blog/what-is-lstm/#:~:text='%20LSTM%20stands%20for%20long%20short,especially%20in%20sequence%20prediction%20problems. (Accessed: January 12, 2023).