e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:05/May-2022 Impact Factor- 6.752 www.irjmets.com

SIGN LANGUAGE RECOGNITION

Arugula Sismai[*1], Asritha Diddi[*2], Chavali Anusha[*3],
Chinta Sundara Sreya[*4], Prof. B. Prajna[*5]

[*1,2,3,4,5]Department Of Computer Science And Systems Engineering, Andhra University College Of Engineering For Women, Visakhapatnam, Andhra Pradesh, India.

ABSTRACT

Sign language is the only tool of communication for the person who is not able to speak and hear anything. Sign language is a boon for verbally challenged people to express their thoughts and emotion. Using this Sign Language Recognition system, the communication gap between people with hearing impairments and the general public can be cleared. In this work, a scheme of sign language recognition has been proposed for identifying the gestures in sign language. With the help of computer vision and neural networks, we can detect the signs and give the respective text as output. The major aim of this work is to build a neural network using a Long Short-Term Memory (LSTM) deep learning model using the video frames which offer the translation of gestures into text. The model is trained with the dataset that is collected using holistic key points from the video of the person which detects the pose, face and hand landmarks.

Keywords: Sign Language Recognition; Application Of Sign Language; Data Input; Deaf; Mute.

I. INTRODUCTION

Sign Language is a language that includes gestures with bodily movements made with the hands including facial expressions and postures. It is mainly used by people who are deaf and dumb. There are many different sign languages such as British, Indian and American sign languages. People with disabilities like deaf and dumb use sign language as a tool to express their emotions and thoughts to common people around them. Yet the general public finds it hard to understand the sign and therefore such a trained system like sign language recognition is required during medical and legal appointments, educational and training sessions and for the global meetings being held. A few years ago, there has been an increase in demand for such systems which are formed as video remote human interpreters using high-speed internet connectivity which provided an easy way to translate the sign language that has been used and benefited from yet had a various number of limitations.

To overcome this, we use a Long Short-Term Memory (LSTM) model to detect the actions in sign language. A neural network of six layers is constructed using LSTM deep learning model in which three are LSTM layers and the other three are Dense layers. The dataset we use contains the actions as a specific number of sequences stored as frames which are captured using OpenCV with an interval of time.

II. METHODOLOGY

2.1 REQUIREMENTS

- Software Requirements:

1. Operating system: Windows 7 and above.
2. Language: Python
3. Libraries:

a. OpenCV: An open-source distribution that is primarily focused on real-time applications that provide computational efficiency for managing massive volumes of data. It processes photos and videos to recognize signs and gestures.

b. TensorFlow: An open-source artificial intelligence package that is used to build models using data-flow graphs and to build large-scale neural networks using several layers.

c. Keras: This is used along with TensorFlow to build a neural network that leverages LSTM layers to handle the sequence of key points.

d. MediaPipe: A Framework for building machine learning pipelines for processing time-series data like video, audio, etc. Here we'll be using mediapipe holistic which is one of the pipelines which contains optimized face, hands, and pose landmarks that allow collecting key points, thus enabling the model to simultaneously detect hand and body poses along with face landmarks.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:05/May-2022 Impact Factor- 6.752 www.irjmets.com

e. Sklearn: This is used to evaluate the performance of the system using the built-in metrics and confusion matrix which gives us the accuracy.

- Hardware Requirements:

1. Camera: Good quality, 3MP
2. RAM: Minimum 8GB and higher.
3. GPU: 4GB dedicated.
4. Processor: Intel Pentium 4 or higher.
5. HDD: 10GB or higher.
6. Monitor: 15 inches or 17 inches color monitor.
7. Mouse: Scroll or Optical. / Touchpad.

2.2 PROPOSED SYSTEM

Initially, the video of the person is captured using OpenCV which is taken as an input and then the data is collected using MediaPipe holistic which detects the face, pose and hand landmarks as key points. The dataset to be trained is stored as several sequences put in frames of video format where the key points are pushed into a NumPy array.

Hereafter, the system is trained and built using Long Short-Term Memory (LSTM) deep learning model which is constructed using three LSTM layers and three Dense layers. This model was trained for 2000 epochs on a batch size of 128 using the dataset extracted. The model was trained using the dataset to minimize the loss by categorical cross-entropy using the Adam optimizer.

Finally, after building the neural network, real-time sign language recognition is performed using OpenCV where the gestures are recognized and displayed as text within the highlighted section.

The working of the proposed system is shown as a network architecture below:

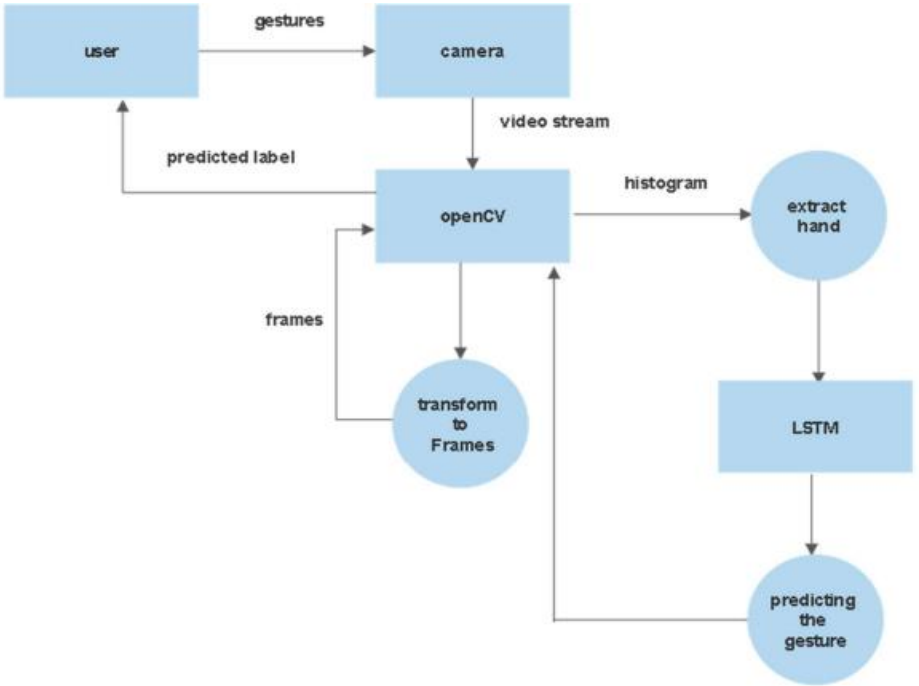

Figure 1: Network architecture of the proposed system.

2.3 ALGORITHM SPECIFICATION

The step by step procedure to construct this system:

1. Installing and importing dependencies.
2. Key points using MediaPipe holistic.
3. Extract Key Points.
4. Setup folders for data collection.
5. Collect Keypoint sequences.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:05/May-2022 Impact Factor- 6.752 www.irjmets.com

6. Preprocess data and create labels.
7. Build and Train an LSTM deep learning model.
8. Make sign language predictions.
9. Save model weights.
10. Evaluation using confusion matrix and accuracy score.
11. Test in real-time.

2.4 SYSTEM MODULES

1. Install and import dependencies:

Here we do install all the required tools i.e. TensorFlow, OpenCV, MediaPipe, Sklearn, matplotlib and import dependencies of NumPy, os, time and pyplot from matplotlib.

2. Collecting key points from MediaPipe holistic:

We detect Hand, Face, and Pose Landmarks and extract all the key points detected using MediaPipe holistic.

3. Collecting and pre-processing data:

We create folders to export the data to be stored as NumPy arrays and create labels.

4. Training and Testing:

Using TensorFlow and Keras we build and train the model using an LSTM deep learning neural network where the model summary and accuracy are defined and tested in real-time.

III. TESTING

Software testing is an examination that provides us with information about the quality of the system being evaluated. Product testing can also provide an objective, unbiased view of the software, allowing them to appreciate the usage and understand the limitations associated with its implementation. One example of a test technique is the practice of executing a program or application to find software problems such as errors and many more. Program testing can provide users with objective, unbiased information regarding software quality and the risk of failure.

Table 1: Test Cases Representation

S.No	Test Case	Input Description	Expected Output	Result (Pass/Fail)
1.	Run the code in Jupyter Notebook without havinga dataset	Delete the downloaded dataset onthe testing machine & Run the Jupyter notebook cells	Error is thrown	Pass
2.	Loading model	Initializing the trained model and loading it into ON	Loaded modelwithout errors	Pass
3.	Converting video to frames	Capturing video and converting it intoframes	Image frames ofthe captured video stream	Pass
4.	Recognize hand gesture	Image frame that contains hand object	label	Pass

As soon as there is an executable program, software testing can be started even if it is completed partially. The whole strategy for software testing or development usually determines when and how testing is done, as well as the results. The majority of testing, for example, occurs after system requirements have been created and then implemented in testable code in a planned process.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:05/May-2022 Impact Factor- 6.752 www.irjmets.com

IV. RESULTS

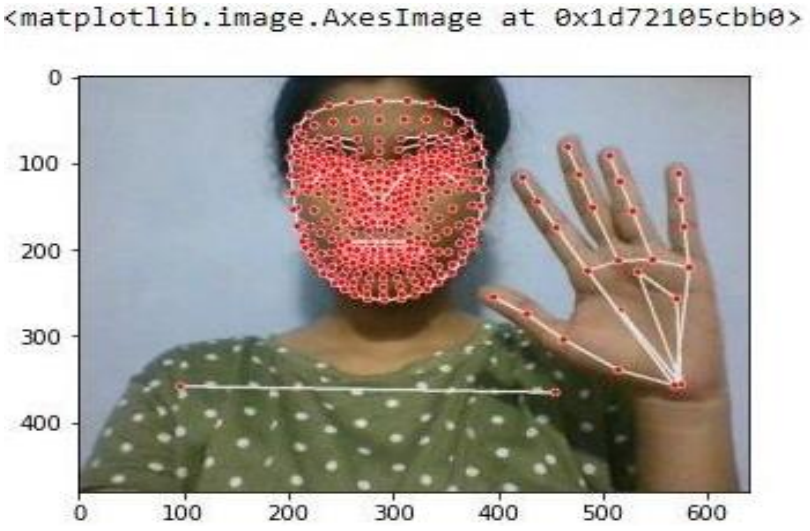

Figure 2: Extracting key points using MediaPipe holistic.

```
In [45]: model.fit(X_train, y_train, epochs=2000, callbacks=[tb_callback])

Epoch 1/2000
3/3 [==============================] - 1s 101ms/step - loss: 1.3544e-05 - categorical_accuracy: 1.0000
Epoch 2/2000
3/3 [==============================] - 0s 113ms/step - loss: 1.3529e-05 - categorical_accuracy: 1.0000
Epoch 3/2000
3/3 [==============================] - 0s 113ms/step - loss: 1.3508e-05 - categorical_accuracy: 1.0000
Epoch 4/2000
3/3 [==============================] - 0s 109ms/step - loss: 1.3501e-05 - categorical_accuracy: 1.0000
Epoch 5/2000
3/3 [==============================] - 0s 114ms/step - loss: 1.3470e-05 - categorical_accuracy: 1.0000
Epoch 6/2000
3/3 [==============================] - 0s 111ms/step - loss: 1.3436e-05 - categorical_accuracy: 1.0000
Epoch 7/2000
3/3 [==============================] - 0s 126ms/step - loss: 1.3414e-05 - categorical_accuracy: 1.0000
Epoch 8/2000
3/3 [==============================] - 0s 115ms/step - loss: 1.3420e-05 - categorical_accuracy: 1.0000
Epoch 9/2000
3/3 [==============================] - 0s 124ms/step - loss: 1.3396e-05 - categorical_accuracy: 1.0000
Epoch 10/2000
```

Figure 3: Training the LSTM model for 2000 epochs.

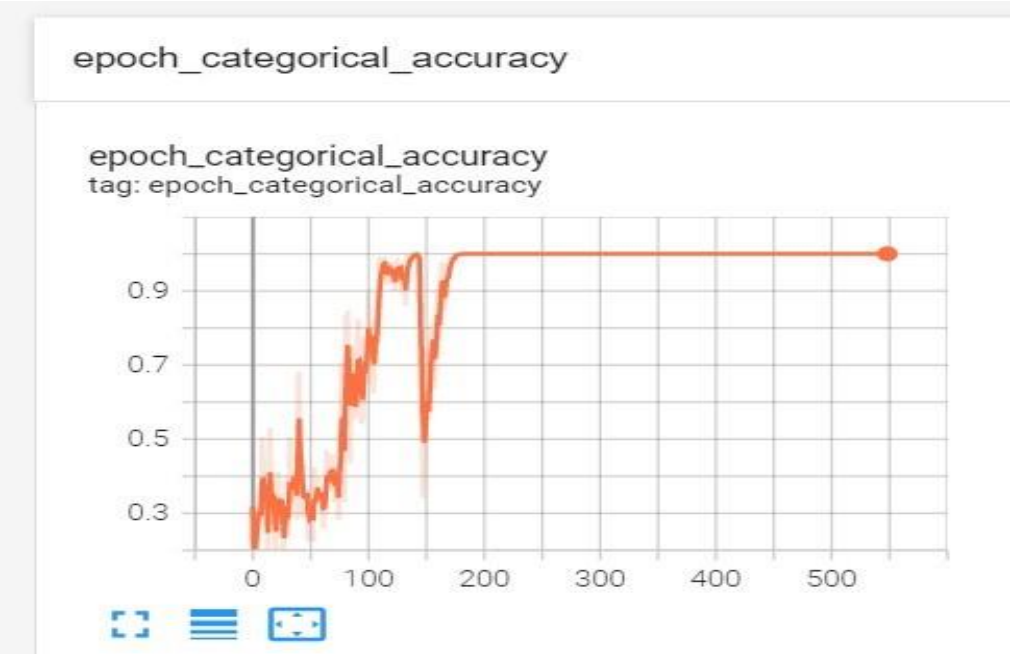

Figure 4: Evaluation using accuracy score of 1.000

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:05/May-2022 Impact Factor- 6.752 www.irjmets.com

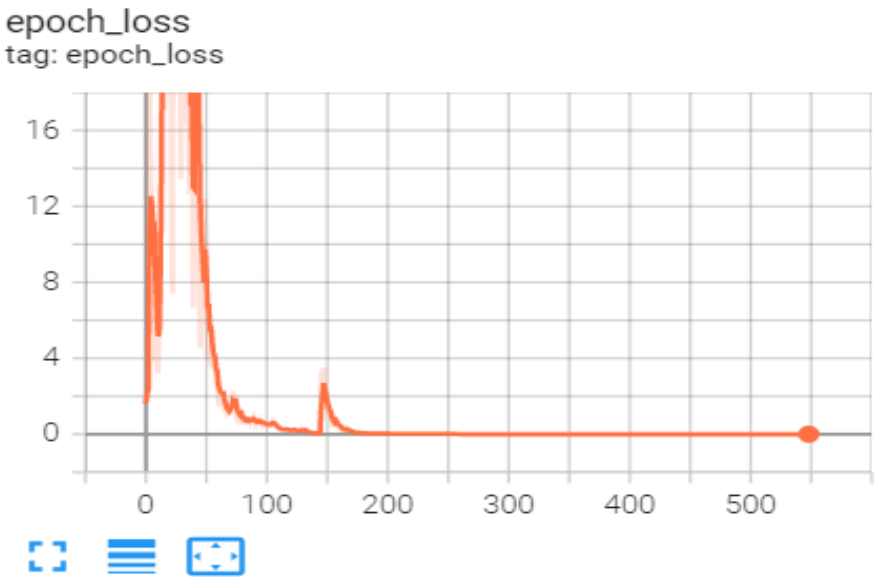

Figure 5: The model is trained to minimize the loss using Adam Optimizer.

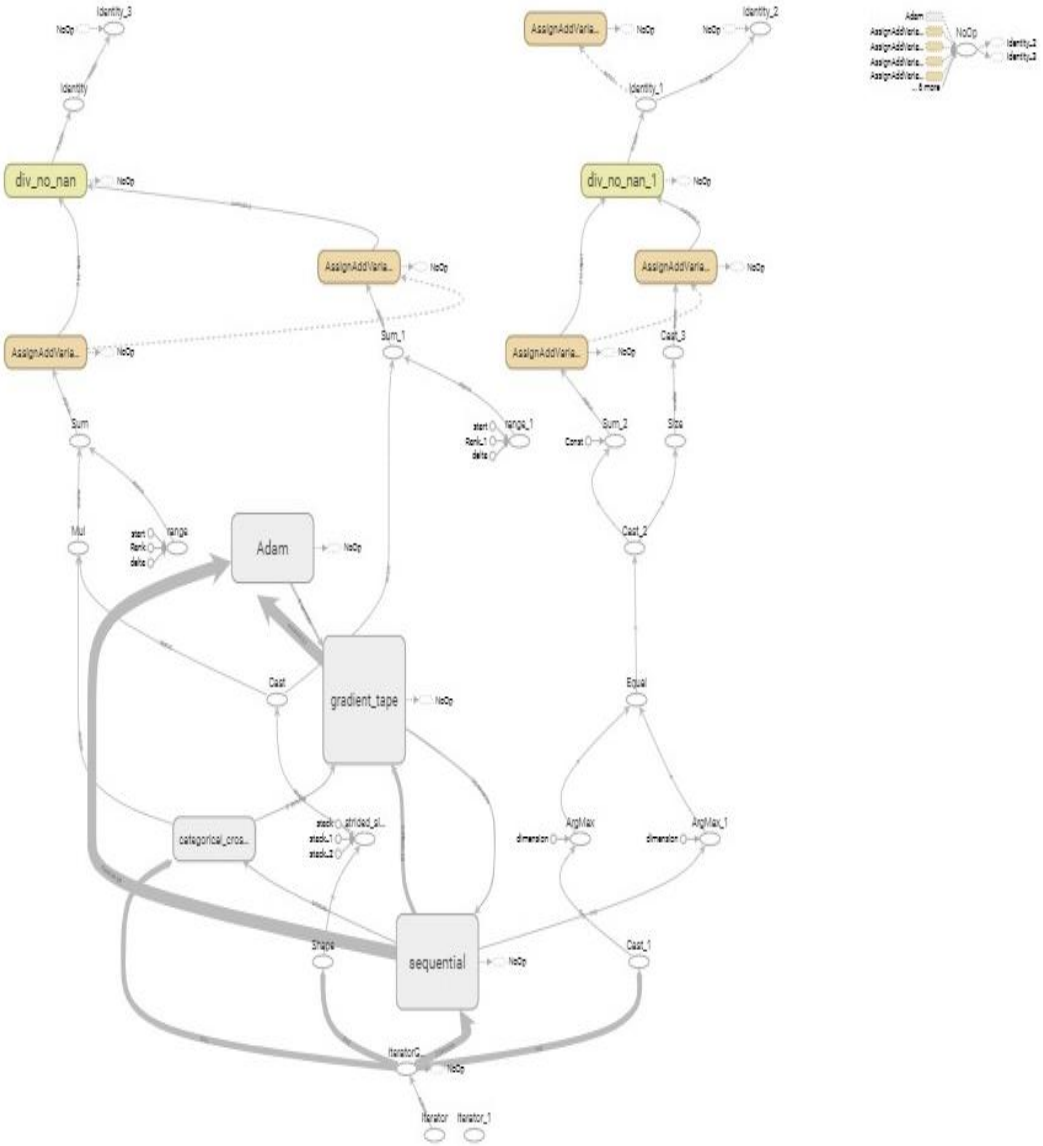

Figure 6: Data flow graph built using TensorFlow for the proposed system.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:05/May-2022 Impact Factor- 6.752 www.irjmets.com

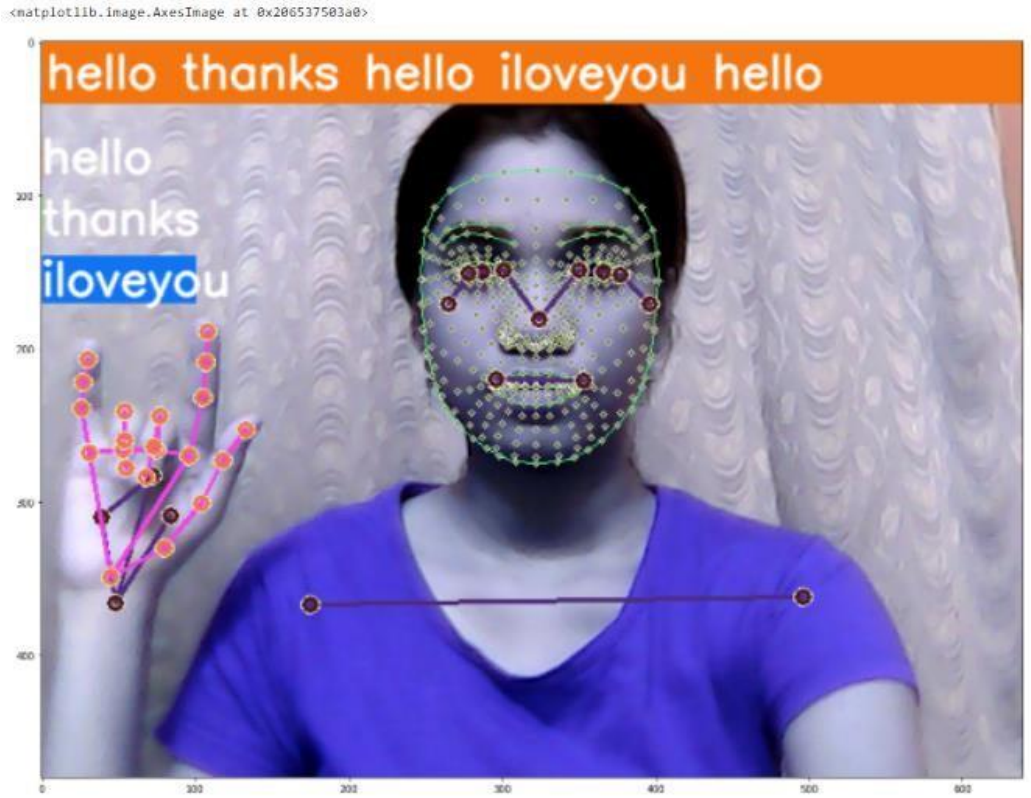

Figure 7: Performing real-time sign language recognition using OpenCV

V. FUTURE ENHANCEMENT

The future scope of this work can be extended:

- To implement other sign languages such as American Sign Language and Indian Sign Language.
- To further train the LSTM model to recognize alphabets and symbols.
- To enhance the model to detect facial expressions.
- To display sentences instead of words which could be a more appropriate sign language translation which also increases readability.
- To add a greater number of training data which results in a higher accuracy score.
- To convert signs to speech.
- To develop a proficient system completely that could help the deaf and dumb people.

VI. CONCLUSION

From classifying signs and numbers, the Sign Language Recognition System can be progressed to a system that can recognize dynamic movements in continuous sequences of images. Nowadays, both researchers and developers are focusing their efforts on developing a wide vocabulary for sign language recognition systems. They differ in their classification methods and the model being trained for detecting sign language as each one of them uses their customized working model. Because of the differences in sign language between countries and the conditions set, fair comparisons between various models are limited. The majority of the country's sign language variations are dependent on their grammar and how they portray each word.

No matter what whichever model being used can break the bridge gap between the people with hearing impairments and the general public from which many can be benefited during real-time activities such as education, global meetings, and medical and legal appointments.

VII. REFERENCES

[1] Aditya Das, Shantanu Gawde, Khyati Suratwala, Dr. Dhananjay Kalbande (2018, February). Facial expression recognition from video sequences: temporal and static modelling. Computer Vision and Image Undertaking 91.

[2] http://ai.googleblog.com/2019/08/on-device-real-time-hand-tracking-with.html

[3] https://mediapipe.dev/

[4] https://ieeexplore.ieee.org/document/8537248

[5] A Practical Introduction to Computer Vision with OpenCV-WILEY Continuous dynamic Indian Sign Language gesture recognition with invariant backgrounds by Kumud Tripathi, Neha Baranwal, G. C. Nandi at 2015 Conference on Advances in Computing, Communications and Informatics (ICACCI).

[6] https://www.researchgate.net/publication/342331104_Sign_Language_Recognition_Using_Deep_Learning_and_Computer_Vision

[7] https://towardsdatascience.com/lstm-recurrent-neural-networks-how-to-teach-a-network-to-remember- the-past-55e54c

International Research Journal Of Modernization in Engineering Technology and Science

(Peer-Reviewed, Open Access, Fully Refereed International Journal)

e-ISSN: 2582-5208

Ref: IRJMETS/Certificate/Volume 4/Issue 05/40500037818

Date: 07/05/2022

Certificate of Publication

*This is to certify that author "**Arugula Sismai**" with paper ID "**IRJMETS40500037818**" has published a paper entitled "SIGN LANGUAGE RECOGNITION" **in International Research Journal Of Modernization In Engineering Technology And Science (IRJMETS), Volume 4, Issue 05, May 2022***

Editor in Chief

IRJMETS Impact Factor 6.752

We Wish For Your Better Future

www.irjmets.com

International Research Journal Of Modernization in Engineering Technology and Science

(Peer-Reviewed, Open Access, Fully Refereed International Journal)

e-ISSN: 2582-5208

Ref: IRJMETS/Certificate/Volume 4/Issue 05/40500037818

Date: 07/05/2022

Certificate of Publication

*This is to certify that author "**Asritha Diddi**" with paper ID "**IRJMETS40500037818**" has published a paper entitled "SIGN LANGUAGE RECOGNITION" **in International Research Journal Of Modernization In Engineering Technology And Science (IRJMETS), Volume 4, Issue 05, May 2022***

Editor in Chief

IRJMETS Impact Factor 6.752

We Wish For Your Better Future

www.irjmets.com

International Research Journal Of Modernization in Engineering Technology and Science

(Peer-Reviewed, Open Access, Fully Refereed International Journal)

e-ISSN: 2582-5208

Ref: IRJMETS/Certificate/Volume 4/Issue 05/40500037818

Date: 07/05/2022

Certificate of Publication

*This is to certify that author "**Chavali Anusha**" with paper ID "**IRJMETS40500037818**" has published a paper entitled "SIGN LANGUAGE RECOGNITION" **in International Research Journal Of Modernization In Engineering Technology And Science (IRJMETS), Volume 4, Issue 05, May 2022***

Editor in Chief

IRJMETS Impact Factor 6.752

We Wish For Your Better Future

www.irjmets.com

International Research Journal Of Modernization in Engineering Technology and Science

(Peer-Reviewed, Open Access, Fully Refereed International Journal)

e-ISSN: 2582-5208

Ref: IRJMETS/Certificate/Volume 4/Issue 05/40500037818

Date: 07/05/2022

Certificate of Publication

*This is to certify that author "**Chinta Sundara Sreya**" with paper ID "**IRJMETS40500037818**" has published a paper entitled "SIGN LANGUAGE RECOGNITION" **in International Research Journal Of Modernization In Engineering Technology And Science (IRJMETS), Volume 4, Issue 05, May 2022***

Editor in Chief

IRJMETS Impact Factor 6.752

We Wish For Your Better Future

www.irjmets.com

International Research Journal Of Modernization in Engineering Technology and Science

(Peer-Reviewed, Open Access, Fully Refereed International Journal)

e-ISSN: 2582-5208

Ref: IRJMETS/Certificate/Volume 4/Issue 05/40500037818

Date: 07/05/2022

Certificate of Publication

*This is to certify that author "**Prof. B. Prajna**" with paper ID "**IRJMETS40500037818**" has published a paper entitled "SIGN LANGUAGE RECOGNITION" **in International Research Journal Of Modernization In Engineering Technology And Science (IRJMETS), Volume 4, Issue 05, May 2022***

Editor in Chief

IRJMETS Impact Factor 6.752

We Wish For Your Better Future

www.irjmets.com