

# Plagiarism - Report

Originality Assessment

20%



Overall Similarity

**Date:** Nov 16, 2023

**Matches:** 362 / 1834 words

**Sources:** 17

**Remarks:** Moderate similarity detected, you better improve the document (if needed).

**Verify Report:**

# A RNN APPROACH ON COMPUTER VISION <sup>13</sup> FOR SIGN LANGUAGE TRANSLATION

Karpagam V 1 Madhu Maaliga Shree K S2 , Nivethika P3

## Abstract

Sign language serves as a non-verbal means of transferring thoughts and emotions, primarily utilized by individuals who are hearing impaired, mute, or speech-impaired. Unfortunately, the linguistic nuances <sup>14</sup> of sign language often go unnoticed by the general public. To bridge this communication gap, there is a need to develop <sup>6</sup> a Sign Language translation system based on the Indian Sign Language (ISL) Dataset. In this endeavor, the Recurrent Neural Network (RNN) algorithm is employed to recognize and interpret static hand signals representing letters of the <sup>1</sup> Indian Sign Language, ultimately converting them into written text. This technical solution leverages the power of RNNs to capture temporal dependencies in <sup>2</sup> the sign language gestures, facilitating accurate translation and fostering better communication between the hearing impaired and the wider community.

Keywords: ISL,LSTM,RNN , Deep Learning , Natural Language Processing (NLP)

## 1 INTRODUCTION

### Sign language

Sign language is a visual and gestural language used by deaf and hard-of-hearing individuals, as well as those who communicate with them. It <sup>14</sup> is a complete and natural language with its own grammar and vocabulary, distinct from spoken languages. Sign language relies on hand shapes, movements, facial expressions, and body postures to convey meaning. Different countries and regions often <sup>10</sup> have their own sign languages, such as American Sign Language (ASL) in the United States, Indian Sign Language (ISL) in India, and many others around the world. Sign language plays a vital role in facilitating communication for the deaf community and is recognized as an essential means of

expression and linguistic identity.

### LSTM <sup>3</sup> (Long Short-Term Memory)

Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) architecture, designed to handle and model sequences of data. LSTMs are known for their ability to capture long-term dependencies and learn patterns in sequential data. They utilize a more complex memory mechanism compared to traditional RNNs, with components like cell states, input gates, output gates, and forget gates to control and maintain information flow through time steps. LSTMs have become a valuable tool in various applications, including natural language processing, speech recognition, time-series analysis, and more, where modeling and understanding sequences and temporal dependencies are crucial.

Machine learning focuses on the development of algorithms and statistical models that enable computers to improve their performance on a specific task through learning from data, without being explicitly programmed. It allows machines <sup>4</sup> to automatically learn from and make predictions or decisions based on patterns.

<sup>1</sup> Machine learning is applied to hand gesture recognition, a technology that enables computers to interpret and respond to human gestures, including hand movements and gestures. This involves collecting a dataset of hand gestures, preprocessing the data, extracting relevant features, and training machine learning models, <sup>3</sup> such as convolutional neural networks (CNNs) for image-based recognition or recurrent neural networks (RNNs) for sequential data. Trained models are then used for real-time recognition of hand gestures, making them useful in applications like sign language translation, virtual reality interaction, and gaming, where gesture-based input plays a crucial role in enhancing human-computer interaction.

<sup>4</sup> Deep learning is a subfield of machine learning that focuses on using deep neural networks to solve complex tasks by automatically extracting hierarchical features from data. These networks, inspired by the human brain, <sup>15</sup> are trained on extensive datasets and have excelled in areas like computer vision,

natural language processing, and speech recognition.

In hand gesture recognition, deep learning involves using deep neural networks to automatically capture intricate patterns and nuances in hand movements. This technology has significantly improved the accuracy and robustness of hand gesture recognition systems, making them useful in sign language translation, virtual reality interaction, and gesture-based control interfaces.

## 7 Recurrent Neural Network (RNN)

A Recurrent Neural Network (RNN) is a specialized type of artificial neural network designed to process sequences of data by maintaining internal memory and context information through looping connections. This attribute makes RNNs particularly suited for tasks where the order and context of information are crucial, such as time-series data, natural language processing, and sequential data analysis. RNNs are commonly employed in applications like text generation, speech recognition, and language translation, as they can effectively capture dependencies and patterns within data over time. In the context of hand gesture recognition, RNNs serve as valuable tools for processing and interpreting sequences of data, such as the temporal information associated with hand movements and gestures. Their recurrent connections enable the preservation of context and memory, making them adept at recognizing and categorizing different hand gestures based on sequential data. This capability is particularly advantageous in applications like sign language recognition and gesture-based human-computer interaction, where understanding the temporal aspect of hand gestures is critical for accurately interpreting the intended meanings.

1 In the context of hand gesture recognition, Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) architecture that is employed to process and analyze sequences of data related to hand movements and gestures. LSTMs are well-suited for capturing the temporal dynamics and dependencies in hand gestures over time. By utilizing their memory cells and gating mechanisms, LSTMs can effectively store and update information about hand movements, allowing them to recognize and classify different hand gestures based on the sequential data provided. This

makes LSTMs valuable for applications involving gesture-based interfaces, **sign language recognition**, and other tasks where the temporal aspect of hand gestures **is crucial for** accurate interpretation and classification.

### Data Collection

During this phase, the necessary data was gathered, consisting **8** of sign language gestures performed by participants. The primary data source was a webcam, which captured real-time video of participants performing **sign language gestures**. These gestures represent the core input to **6 the sign language recognition** model.

To capture the hand movements and body postures relevant to sign language, the MediaPipe Holistic model was utilized. This model enabled the **8** detection and tracking of key points, including facial landmarks, body posture, and hand positions, in real-time.

The detected key points were visualized using MediaPipe's drawing utilities to ensure the quality and accuracy of data collection. These visualizations helped confirm **1 that the model was** correctly tracking relevant body and hand movements. Captured key point data, representing **8 the sign language gestures**, was stored and organized in specific folders for further processing. The data was categorized by the type **of sign language** gesture to ensure easy access during the subsequent phases.

### Data Preprocessing

The collected data was meticulously examined to identify and address any missing or inconsistent values. In cases where data was incomplete or erroneous, data imputation and removal techniques were applied to enhance data quality.

To facilitate model convergence and consistent performance, the pixel **1 values in the** images were normalized. This step ensured that all input data had consistent scaling.

Data augmentation techniques were applied to improve the model's generalization and robustness.

These included random rotations and flips **to increase the** diversity **of the training** dataset.

To evaluate the model's performance, the preprocessed data was divided into **training, validation, and** testing sets. A standard data splitting approach was employed to maintain balanced data distributions among different sign language gestures.

The collected images and their corresponding labels were converted into suitable formats, typically arrays or tensors, to prepare the data for **3 input to the deep learning model.**

## Model Training

**1 The model training** phase is at the core of the **sign language recognition** system. During this phase, a **Long Short-Term Memory (LSTM)** neural network was built and trained using the preprocessed data. The LSTM-based **2 neural network architecture** was designed, consisting of multiple LSTM layers followed by dense layers. The LSTM layers were configured **to capture temporal dependencies within the data,** while the dense layers performed final classifications.

**1 The model was trained** over a specific number of epochs, with TensorBoard **used as a** callback for logging and monitoring training progress. The model's predictions were compared with the ground truth labels to assess its accuracy in recognizing sign language gestures. Metrics such as accuracy, precision, and recall were computed to evaluate its performance.

Model testing provided insights into the model's ability **1 to generalize and** accurately recognize sign language **gestures in real-time**

## Model Evaluation

During model evaluation and deployment phase, the performance is assessed **and it is** prepared for real-world use.

## Confusion Matrix

A multilabel confusion matrix was employed to analyze the model's performance in recognizing different sign language gestures. This matrix offered a comprehensive view of the model's strengths and weaknesses. The accuracy score was calculated **1 to provide a** quantitative **measure of the** model's overall performance in correctly classifying sign language gestures.

## Conclusion

**16 The Sign Language to Text Conversion** exemplifies the transformative potential of technology in breaking communication barriers for the deaf and hard-of-hearing community. By harnessing

computer vision <sup>13</sup> and deep learning, the project enables real-time sign language recognition and translation into text, fostering effective communication and inclusivity. With future applications spanning mobile technology, education, and accessibility in public spaces, the project highlights the evolving role of AI in enhancing lives and connecting communities. Looking ahead, the commitment to cultural sensitivity and user feedback underscores the dedication to building a more inclusive and accessible world, where communication knows no bounds.

#### Future Work

Implement cloud-based text-to-speech and speech-to-sign conversion for accessibility on various devices. Integrate natural language processing to improve the system's accuracy and understanding of user input. Develop a user-friendly mobile application for seamless two-way communication <sup>1</sup> without the need for specialized hardware. Explore augmented reality (AR) and virtual reality (VR) for immersive sign language interaction.

#### References

- [1] Zhibo Wang , Tengda Zhao , Jinxin Ma , Hongkai Chen , Kaixin Liu , Huajie Shao , Qian Wang, Ju Ren (2020), "Hear Sign Language: "A Real-time End-to-End Sign Language Recognition System", IEEE Transactions On Mobile Computing.
- [2] Jinsu Kunjumon , Dr. Rajesh Kannan Megalingam(2019) , "Hand Gesture <sup>6</sup> Recognition System For Translating Indian Sign Language Into Text And Speech", IEEE Second International Conference on Smart Systems and Inventive Technology (ICSSIT)
- [3] Aishwarya Sharma , Dr. Siba Panda , Prof. Saurav Verma (2020). "Sign Language to Speech Translation", IEEE
- [4] Zhongxu Hu, <sup>17</sup> Youmin Hu, Bo Wu, Jie Liu (2017) , " Hand pose estimation with CNN-RNN", European Conference on Electrical Engineering and Computer Science
- [5] Kenneth Lai and Svetlana N. Yanushkevich(2018), " <sup>11</sup> CNN+RNN Depth and Skeleton based Dynamic Hand Gesture Recognition", 24th International Conference on Pattern Recognition (ICPR)
- [6] Klaus Greff, Rupesh K. Srivastava, Jan Koutník, Bas R. Steunebrink, and Jürgen

Schmidhuber(2016),"LSTM: A Search Space Odyssey", IEEE Transactions On Neural Networks And Learning Systems

[7] Yu Wang(2017) ,"A New Concept using LSTM Neural Networks for Dynamic System Identification", American Control Conference Sheraton Seattle Hotel

[8] 5 Zhibo Wang, Tengda Zhao, Jinxin Ma, Hongkai Chen, Kaixin Liu, Huajie Shao, Qian Wang,Ju Ren(2020),"Hear Sign Language: A Real-time End-to-End Sign Language Recognition System",IEEE Transactions On Mobile Computing

[9] Aakash Deep; Aashutosh Litoriya, Akshay Ingole, Vaibhav Asare, 9 Shubham M Bhole, Shantanu Pathak(2022),"Realtime Sign Language Detection and Recognition", 2022 2nd Asian Conference on Innovation in Technology (ASIANCON)

[10] Subhangi Adhikary; Anjan Kumar Talukdar; Kandarpa Kumar Sarma,"A 12 Vision-based System for Recognition of Words used in Indian Sign Language Using MediaPipe",2021 Sixth International Conference on Image Information Processing (ICIIP)



## Sources

1	<a href="https://www.mdpi.com/2076-3417/12/19/9700">https://www.mdpi.com/2076-3417/12/19/9700</a> INTERNET 4%
2	<a href="https://www.interviewkickstart.com/learn/recurrent-neural-networks-sequential-data-processing">https://www.interviewkickstart.com/learn/recurrent-neural-networks-sequential-data-processing</a> INTERNET 3%
3	<a href="https://www.geeksforgeeks.org/deep-learning-introduction-to-long-short-term-memory/">https://www.geeksforgeeks.org/deep-learning-introduction-to-long-short-term-memory/</a> INTERNET 2%
4	<a href="https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-ml/">https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-ml/</a> INTERNET 1%
5	<a href="https://researchr.org/publication/WangZMCLSWR22">https://researchr.org/publication/WangZMCLSWR22</a> INTERNET 1%
6	<a href="https://www.mdpi.com/1999-5903/14/9/253">https://www.mdpi.com/1999-5903/14/9/253</a> INTERNET 1%
7	<a href="https://deeptai.org/machine-learning-glossary-and-terms/recurrent-neural-network#:~:text=A Recurrent Neural Network (RNN) is a class,internal state (memory) to process sequences of inputs.">https://deeptai.org/machine-learning-glossary-and-terms/recurrent-neural-network#:~:text=A Recurrent Neural Network (RNN) is a class,internal state (memory) to process sequences of inputs.</a> INTERNET 1%
8	<a href="https://www.semanticscholar.org/paper/Realtime-Sign-Language-Detection-and-Recognition-Deep-Litoriya/e2eae213180fc1e955371d1c4b6fb22488a89fce">https://www.semanticscholar.org/paper/Realtime-Sign-Language-Detection-and-Recognition-Deep-Litoriya/e2eae213180fc1e955371d1c4b6fb22488a89fce</a> INTERNET 1%
9	<a href="https://www.semanticscholar.org/paper/An-approach-to-Generation-of-sentences-using-Sign-Vikash-Jayakrishnan/bb642c40a6d50541f4f72a574b78026881e05991">https://www.semanticscholar.org/paper/An-approach-to-Generation-of-sentences-using-Sign-Vikash-Jayakrishnan/bb642c40a6d50541f4f72a574b78026881e05991</a> INTERNET 1%
10	<a href="https://education.nationalgeographic.org/resource/sign-language/">https://education.nationalgeographic.org/resource/sign-language/</a> INTERNET 1%
11	<a href="https://www.researchgate.net/publication/329318028_CNNRNN_Depth_and_Skeleton_based_Dynamic_Hand_Gesture_Recognition">https://www.researchgate.net/publication/329318028_CNNRNN_Depth_and_Skeleton_based_Dynamic_Hand_Gesture_Recognition</a> INTERNET 1%
12	<a href="https://ieeexplore.ieee.org/document/9702551/">https://ieeexplore.ieee.org/document/9702551/</a> INTERNET 1%
13	<a href="https://link.springer.com/article/10.1007/s00521-019-04691-y">https://link.springer.com/article/10.1007/s00521-019-04691-y</a> INTERNET 1%
14	<a href="https://www.nidcd.nih.gov/health/american-sign-language">https://www.nidcd.nih.gov/health/american-sign-language</a> INTERNET 1%

15

<https://medium.com/analytics-vidhya/part-1-introduction-to-natural-language-processing-nlp-a66ad8773b3>  
INTERNET  
<1%

---

16

[https://www.researchgate.net/publication/335433017\\_Conversion\\_of\\_Sign\\_Language\\_To\\_Text\\_And\\_Speech\\_Using\\_Machine\\_Learning\\_Techniques](https://www.researchgate.net/publication/335433017_Conversion_of_Sign_Language_To_Text_And_Speech_Using_Machine_Learning_Techniques)  
INTERNET  
<1%

---

17

[https://www.researchgate.net/publication/371701102\\_Accurate\\_Hand\\_Gesture\\_Recognition\\_using\\_CNN\\_and\\_RNN\\_Approaches](https://www.researchgate.net/publication/371701102_Accurate_Hand_Gesture_Recognition_using_CNN_and_RNN_Approaches)  
INTERNET  
<1%

---

EXCLUDE CUSTOM MATCHES OFF

EXCLUDE QUOTES OFF

EXCLUDE BIBLIOGRAPHY OFF