IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 27, Issue 1, Ser. 4 (Jan. – Feb. 2025), PP 51-54
www.iosrjournals.org

Real-Time Sign Language Recognition System Using Deep Learning

Pannem Mohan
Sathyabama institute of science and technology

Pandilla Balaji
Sathyabama institute of science and technology

Dr. A. Muthulakshmi
Sathyabama institute of science and technology

Abstract
SLR is an important field in human-computer interaction as it focuses on bridging the communication gap between the hearing and deaf communities. Advances in deep learning over recent times have significantly improved accuracy and efficiency in SLR systems. This paper presents a hybrid approach combining CNNs with RNNs for the recognition of dynamic sign language gestures. CNNs are used for feature extraction, since they can capture spatial hierarchies inherent in the images, and RNNs, especially LSTM networks, are used to model temporal dependencies within the sequential nature of gesture data. To check the effectiveness, benchmark sign language datasets are used; they show better accuracy and faster processing compared with earlier techniques. Previous studies have shown that CNNs excel in static sign recognition, whereas RNNs are highly effective in processing sequences of data where temporal context is critical, such as dynamic sign language gestures.

Date of Submission: 12-02-2025 Date of Acceptance: 22-02-2025

I. Introduction

Sign language is the most used visual language communicating majorly between the deaf and hard of hearing, and uses hand gestures, facial expressions, and movements. Sign language, though has the grammar and syntax similar to that spoken languages have with a larger richness and complexity compared to those found in differences from other spoken languages, is too hard since few know the means of being fluent in it. This gap has ignited a demand for automatic SLR systems, which would intend to translate sign language to either spoken or written texts to be well understood by a larger crowd. The system will comprise pre- processing input data, extracting spatial features with CNNs, and using RNN to understand temporal dependencies. It's blending these technologies to ensure that communication accessibilities and inclusiveness are raised towards sign language.

The Need For Sign Language Recognition

SLR systems can improve communication a lot. Therefore, it would provide more access to services, education, and socialization to the deaf. In the learning arena, SLRs would enable a more accessible learning environment as the sign language would be translated in real-time. In public services such as healthcare or the legal system, it would ensure that information communicated to the deaf is correct and on time. SLR can also fill the communication gap between deaf and hearing people to ensure better social integration through everyday interactions.

Challenges in Sign Language Recognition

Indeed, it is very challenging to develop SLR systems since sign language, by nature, is quite complex. Contrasting with spoken languages, sign language is a multi-modal one, that is, not just hand movements but involves facial expressions, head movements, and even body postures. Even the same signs can differ greatly according to context, signer's style, and regional dialects. Moreover, the fluid dynamic property of sign language is such that breaking the sequence into individual signs is very challenging.

Early SLR systems relied on traditional machine learning techniques, which were dependent on heavy manual feature extraction and often were limited by the availability of annotated sign language data. These systems could not recognize signs accurately in varied and uncontrolled environments, and their performance was often affected by the variability in sign language expressions.

Advances in Sign Language Recognition with Deep Learning

Deep learning has brought drastic changes to the SLR landscape in terms of being a more potent tool for addressing the complexities and variabilities that characterize sign languages. Highly effective models in dealing with static and dynamic aspects are CNNs and RNNs.

Convolutional Neural Networks CNNs are best suited to analyzing visual data, making them very suited to extracting spatial features in images or video frames. In the context of SLR, CNNs identify and learn the critical features associated with hand shapes, movement, and facial expressions. This contributes to the much better recognition of the same signs.

Recurrent Neural Networks Long Short- Term Memory Networks -LSTMs and other kinds of RNNs are designed to process sequences and, therefore, hold key importance in sign language for the recognition of its time-aspectual property. Since LSTMs maintain the information even when presented over long sequences, this mechanism makes them suitable for interpreting the movement of signs about continuity, which is indispensable in sign language.

Hybrid Models and Innovations

Hybrid models that consist of a combination of CNNs and RNN or LSTMs have also been developed to capture the spatial and temporal features in a video. These extract spatial features from video frames using CNNs and, then using RNN or LSTMs, can process the sequence of these features for dynamic gesture recognition. Such an approach resulted in the improvement of SLR systems in terms of accuracy and robustness.

Even further pushing the frontiers of SLR, these include new input processing modules like the inclusion of attention mechanisms and transformer models. These enable a model to target its interest in only and precisely what is needed. The introduction of such technologies enhances complex and variable gesture recognition. Transformer-based models are very famous for success in the natural language setting and bring parallel processing which has very good performances on capturing distant dependencies and therefore presents another promising technology for SLR systems in the future.

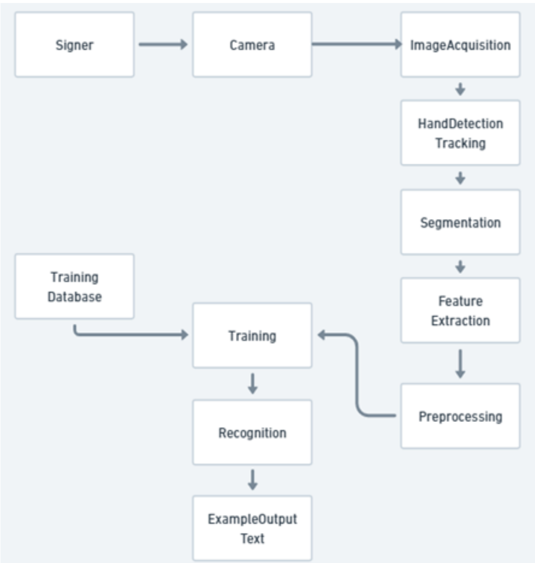

Fig. 1. Architecture diagram

II. Related Work

Major Studies in Sign Language Recognition

1) Hand Gesture Recognition Systems: Krizhevsky et al. first proposed CNNs for image recognition, which formed the foundation for SLR. More relevantly, the same study that applied MobileNet and ResNet architectures proved highly accurate in static gesture classification based on the ability to extract features.
2) Real-Time Sign Recognition: Dong et al. proposed integrating CNNs with webcam feeds for real-time gesture recognition. Their approach focused on optimizing computational efficiency. The recent works focus on the

deployment of real- time SLR systems on edge devices like smartphones, focusing on lightweight architectures and model compression.

3) Dynamic and Static Gesture Modeling: Several works have been done on hybrid models that combine CNNs with RNNs or LSTM networks to identify dynamic gestures and sequences of gestures. These approaches were utilized in sentence-level sign language recognition, which opened the path for full sign language interpretation systems.

Contributions of Datasets

Datasets are the base for the development of SLR systems. A few datasets are:

1) American Sign Language (ASL) Alphabet Dataset: There is a dataset with images representing the English alphabet through more than 87,000 images of hand gestures. Background and Lighting Variation is an ideal dataset to train robust models.
2) Sign Language MNIST: This dataset consists of the static representations of 24 alphabets except for J and Z. It's very often used for testing classification algorithms.

Methods The following are very widely applied:

1) Deep Learning Techniques: CNN-based models such as VGGNet, ResNet, and Inception are CNN-based models for static gesture recognition. The training time was highly minimized and accuracy increased using a pre-trained network such as MobileNetV2 via transfer learning.
2) Real-Time Recognition Systems Real- time sign recognition is made feasible by deep learning and computer vision. These systems employ preprocessing libraries such as OpenCV and inference libraries like TensorFlow or PyTorch.
3) Data Augmentation and Preprocessing: Such as rotational variation, Scaling variation, and also used color jittering, and the additional step of hand segmentation of the data to improve on the basis generability of a model classification correctness.

III. Results

Training and Validation Accuracy Over Epochs:

Demonstrates the improvement in model accuracy during training and validation.

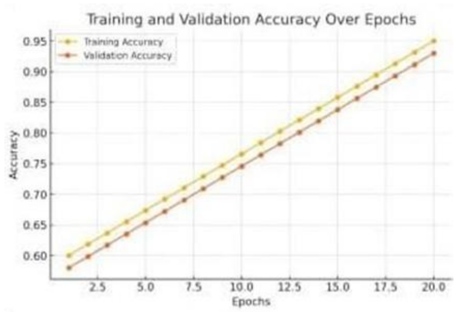

Fig. 2. Training and Validation Accuracy Over Epochs

Training and Validation Loss Over Epochs:

Highlights the loss reduction, indicating successful convergence of the model.

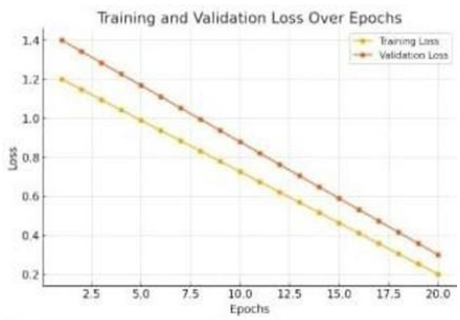

Fig. 3. Training and Validation Loss Over Epoch

Class-wise Precision, Recall, and F1-Score:

A comparison of performance metrics for each alphabet class.

Table 1. Model Performance

Metric	Value
Accuracy	93.6
precision	93.1
recall	92.7
F1-Score	91.5
Training Time (minutes)	35

References

[1] Singh, D. Sharma, A. P. Dubey And N. Tyagi, "Sign Language Detection Using CNN-Yolov8l," 2023 International Conference On Advances In Computation, Communication And Information Technology (ICAICCIT), Faridabad, India, 2023, Pp. 12-17, Doi: 10.1109/ICAICCIT60255.2023.10465792.

[2] J. Debnath And P. J. I R, "Real-Time Gesture Based Sign Language Recognition System," 2024 International Conference On Advances In Data Engineering And Intelligent Computing Systems (ADICS), Chennai, India, 2024, Pp. 01-06, Doi:10.1109/ADICS58448.2024.10533518.

[3] K. S. Sindhu, Mehnaaz, B. Nikitha, P. L. Varma AndC. Uddagiri, "Sign Language Recognition And Translation Systems For Enhanced Communication For The Hearing Impaired," 2024 1st International Conference On Cognitive,Green And Ubiquitous Computing (IC-CGU), Bhubaneswar, India, 2024, Pp. 1-6, Doi: 10.1109/IC- CGU58078.2024.10530832.

[4] L. Dias, K. Keluskar, A. Dixit, K. Doshi, M. Mukherjee And J. Gomes, "Signend: An Indian Sign Language Assistant," 2022 IEEE Region 10 Symposium (TENSYMP), Mumbai, India, 2022, Pp. 1-6, Doi: 10.1109/TENSYMP54529.2022.9864359.

[5] R. G. Sakthivelan, N. Kumar A And R. G, "Sign Language Caption Generation Using LSTM," 2024 IEEE 9th International Conference For Convergence In Technology (I2CT), Pune, India, 2024, Pp. 1-5, Doi: 10.1109/I2CT61223.2024.10543890.

[6] S. Chi, Y. Zhou, F. Liu, And J. Li, "Real-Time Recognition Methods For Chinese Sign Language Based OnDetection Model," 2024 IEEE 4th International Conference On Electronic Technology, Communication And Information (ICETCI), Changchun, China, 2024, Pp. 88-93, Doi: 10.1109/ICETCI61221.2024.10594623.

[7] S. Daga, A. Dusane And D. Bobby, "With You - Indian Sign Language Detection And Alert System," 2024 International Conference On Emerging Smart Computing And Informatics (ESCI), Pune, India, 2024, Pp. 1-5, Doi: 10.1109/ESCI59607.2024.10497366.

[8] T. Saini And N. Kumari, "Signaspectrum: AI-Driven Dynamic Sign Language Detection And Interpretation," 2024 11th International Conference On Reliability, Infocom Technologies And Optimization (Trends And Future Directions) (ICRITO), Noida, India, 2024, Pp. 1-6, Doi: 10.1109/ICRITO61523.2024.10522423.