SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V11I2P107
Volume 11 Issue 2, 60-68, February 2024

Original Article

Dynamic Indian Sign Language Recognition Based on Enhanced LSTM with Custom Attention Mechanism

Jay M. Joshi[1], Dhaval U. Patel[2]

[1]*Department of Electronics and Communication Engineering, Government College of Engineering, Gujarat, India.*
[2]*Department of Electronics and Communication Engineering,Gujarat Technical University, Gujarat, India.*

[2]*Corresponding Author : dhaval.patel4@gujgov.edu.in*

Received: 29 October 2023 Revised: 03 December 2023 Accepted: 30 December 2023 Published: 02 February 2024

Abstract - *In this paper, the author developed a system with an enhanced Long-Short-Term Memory (LSTM) model using a custom Attention Mechanism specifically designed for real-time dynamic Indian Sign Language (ISL) recognition. A large custom dataset of 59 signs was employed. For feature extraction, the MediaPipe framework was used. On this dataset, three different models were trained, the Long Short-Term Memory (LSTM) Dense model with Customed Attention, the LSTM-Dense model with a traditional Attention Mechanism, and the LSTM-Dense model without Attention Mechanism. The proposed Long Short-Term Memory (LSTM) dense model with a customed Attention Mechanism (AM) outperformed the other two models. The proposed LSTM-Dense model with a customed Attention Mechanism achieved a maximum accuracy of 96.08% in predicting the sign. When compared with existing Indian Sign Language recognition methods, our proposed model surpassed all others in accuracy, even with a large number of signs. In addition, 5-fold cross-validation of the proposed model confirmed the durability of our results, with a 93% accuracy. The results show that the proposed recognition system can effectively and robustly recognize ISL gestures.*

Keywords - *Indian Sign Language recognition, Computer vision, Gesture recognition, Pattern recognition, Deep Learning.*

1. Introduction

Communication is an essential part of human society. The ability to communicate and connect with others is a fundamental human need. However, People with speech impairments depend on sign language, creating challenges in communicating with those unfamiliar with it [1]. This demands an efficient sign language recognition system, which is essential for converting sign language into text or audio-visuals. However, existing solutions are limited, expensive, and not user-friendly.

With its 17.7% share of the world population, India has done comparatively less research in this field than other countries [2, 3]. In addition, six million individuals are affected by deafness or speech impairments, making the study of Indian Sign Language (ISL) recognition crucial [4]. Research on ISL commenced in 1978. However, it was short-term due to the lack of a standardized sign language. Moreover, the variation in signs across deaf schools varied significantly. In 2003, Indian Sign Language was standardized formally. This development renewed researchers' interest in this field [3].

Indian Sign languages include static and dynamic signs and can also involve single or double-handed signs. A considerable amount of research is done on static signs, but dynamic sign language remains underexplored. The complexity of dynamic signs, attributable to spatial-temporal variations, individual differences, and issues with repeatability, poses significant challenges. Additionally, the absence of a standardized dataset adds to the complexity of research in dynamic ISL. Recently, significant research has started in this field.

Sign Language Recognition (SLR) systems are divided broadly into sensor-based and vision-based approaches [6]. The sensor-based approach involves using gloves or devices like Microsoft Kinect and leap motion to detect and translate signs in electrical signals. Conversely, the vision-based approach uses a webcam to capture images and video. The vision-based approach is more advantageous due to its cost-effectiveness, user-friendliness, and lack of specialized hardware requirements [7, 8].

For effective sign language recognition, extracting and classifying features are very critical. Machine Learning and Deep Learning technology advancements facilitate new approaches to Indian sign language recognition efficiently and accurately. In the proposed work, the author's contributions are as follows. The author developed an extensive, efficient, real-time, customized attention-based LSTM framework. This approach resolves the specific challenges posed by SLR. Focusing on temporal dynamics identifies important frames and improves accuracy for similar signs. In addition, the author developed a large dataset for 59 signs necessary for everyday communication. The performance of the proposed model has been evaluated using a confusion matrix cross-validation and compared with other state-of-the-art techniques. Results demonstrate the potential of this approach in outperforming standard models, promising advancements in real-world applications of ISL recognition.

In the next section of this research paper, we delve deeper into the literature review, followed by the methodology section, which provides an overview of the proposed methods. The next section comprises results and compares our experiments and other state-of-the-art methods. The final section discusses the conclusion and potential future directions.

2. Related Work

Different authors have implemented various algorithms depending on the nature of sign language. Here, the author conducted an extensive literature review on various dynamic sign language recognition systems from 2010 to 2022.

In the early 2010s, in sign language recognition, a vision-based system utilizing the K-nearest neighbor technique was introduced [4], focused on improving sign classification using oriented histogram features. Although it is a novel approach, it encountered challenges in accurately classifying similar signs. In response, 2011 [9] introduced a novel approach with POI tracking integration.

This approach reduced storage requirements, achieved up to 100% accuracy for 24 alphabets using neural networks, and improved the capability to differentiate similar signs to some extent in a controlled environment.

Simultaneously, a sensor-based approach [10] was introduced with an auto-encoder for classifying 10 ASL signs, and it achieved an accuracy of 98.2% under varied lighting conditions. This advancement overcome the inherent limitations of vision-based systems. Highlight the need for more sophisticated feature extraction techniques for efficient vision-based methods.

In 2013, the main objective was to develop a real-time processing system in SLR. The approach of [11] employed Eigen value-weighted Euclidean distance, which was a crucial step in this direction. Subsequently, a 2014 study [12] introduced hierarchical centroid feature extraction with neural network and KNN classification, which achieved an accuracy of 97.10% for ten numeric signs, marking substantial progress in the feature extraction techniques. In 2015, sensor-based approaches marked significant advancement. A critical study [13] leveraging Microsoft Kinect and multi-class SVM achieved an accuracy of 86.16% for 37 signs. Concurrently, a vision-based system [14] using Discrete Wavelet Transform and HMM attained an accuracy of 100% for ten signs.

In 2016, notable advancement was marked in feature extraction techniques. [15] introduced active contours from boundary edge maps using ANN for classification, which represented a significant enhancement. Furthermore, Leap Motion Controller and Multi-Layer Perceptron in a study [16] demonstrate the promising potential of gesture-based controllers in sensor-based sign recognition.

Advancements in sensor-based and vision-based systems continued in 2017, with studies showcasing remarkable accuracy and adaptability. A notable system [17] utilized novel CamShift, P2DHMM, and Haar Cascade-based algorithms and achieved 90% accuracy.

In 2018, the vision-based approach progressed with the study [18]; it introduced a system that improved mobile sign language recognition. Another study using CNN-based spatial embedding [19] also demonstrated high accuracy. Sensor-based approaches also progress with a study [20] utilizing mean and standard deviation feature analysis with HMM and SVM, which achieved an accuracy of 83.77% for 30 signs. In 2020, some hybrid neural approaches were developed. [21] Neuro-Fuzzy with NLP technology and [22] use CNN inception v3 and LSTM models to improve real-time gesture recognition.

In the subsequent years, the vision-based approach evolved with notable developments in advanced deep learning models. The study [23] introduced double-layer CNN in 2021, and [24] employed an LSTM – GRU model in 2022, demonstrating enhanced performance in real-time detection environments.

This survey provides a detailed overview of dynamic Indian sign language recognition systems. In recent years, Researchers' focus has shifted from a sensor to a vision-based approach due to the transition from essential pattern recognition to advanced deep learning models due to their efficient and accurate results. The significant research gaps in the vision-based approach identified are that the standard dataset is not available, most of the research still focuses on alphabetic signs, a small dataset, and performed poorly for similar signs, mostly in dynamic sign language recognition.

Thus, the author developed a large custom dataset in the proposed work and enhanced the Deep Learning-based technique LSTM with a custom Attention Mechanism (AM) optimized explicitly for dynamic Indian Sign Language applications.

3. Proposed Work

Robust and reliable data is essential to design an efficient and real-time Sign Language Recognition (SLR) system. Here, the authors have developed a custom dataset. The different stages of the SLR process, including dataset collection, preprocessing, feature extraction, and sign classification, are illustrated in Figure 1 below.

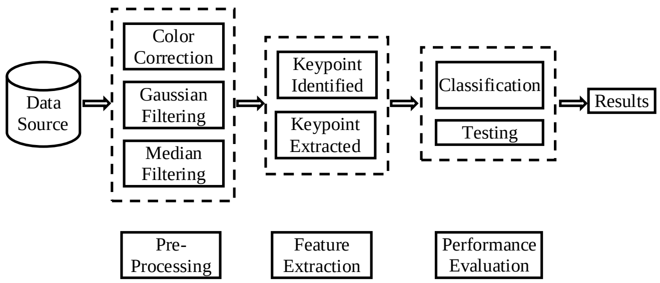

Fig. 1 Block diagram of proposed system

3.1. Dataset Collection

The foundation of any successful sign language recognition system depends on a reliable and comprehensive dataset. For Indian Sign Language Recognition, it becomes even more crucial due to the absence of standard datasets. To overcome this, the author developed a custom dataset. The author considered 59 signs essential for everyday communication, including time, feelings, family, relationships, transportation, and essential communication of 10 individuals.

3.2. Preprocessing and Feature Extraction

In this phase, the first video sequence of 30 frames is captured using a standard 1080p high-resolution webcam. These frames are then converted to RGB color formatting. Following this, frames are normalized and scaled to maintain consistency across the dataset. Additionally, gaussian and median filters were applied to reduce noise and improve image quality, which is essential for accurate feature extraction. The whole process is illustrated in Figure 2.

A substantial dataset is required to achieve accurate results to recognize a wide range of signs. However, as the dataset increases, it corresponds with an increased requirement for large computational resources. To mitigate this, it is crucial to eliminate unwanted background data and concentrate on essential vital points. Here, the author employed a media pipe framework for feature extraction.

This advanced framework is designed to extract detailed spatial information from each frame. A total of 1662 key points were identified, including 1404 facial features, 132 body posture key points, and 126 hand positioning key points. The coordinates of these key points give a detailed representation of each sign in multi-frame analysis using these key points' temporal information determined by tracking their movement over time. This tracking illustrates a pattern of motion and alteration between key points across subsequent frames, revealing material information.

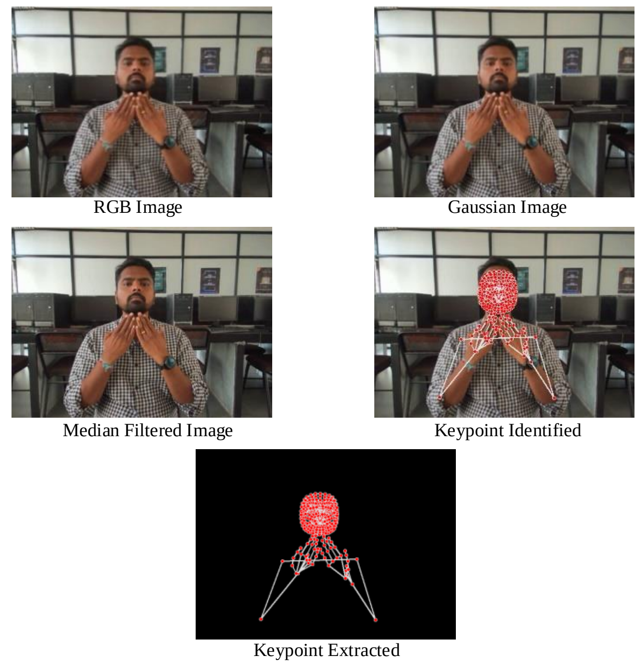

Fig. 2 Preprocessing and feature extractions steps

Combining spatial and temporal data from critical points enables the interpretation of complex hand signs that include dynamic gestures, providing a comprehensive knowledge of nonverbal communication signs.

3.3. Classification

Once dataset collection is complete, the next critical step involves initiating the training process for classification purposes. For dynamic sign language recognition, spatial and temporal information is crucial [25]. Thus, LSTM networks emerge as a natural choice. Long Short-Term Memory (LSTM) network is most relevant because it can process sequential data effectively [26, 27].

LSTM, a specialized form of Recurrent Neural Networks (RNNs), is designed to capture long-term dependencies and mitigate the vanishing gradient problem inherent in traditional RNNs [28]. This is achieved through their unique architecture, which incorporates memory cells and multiple gates – input, forget, and output gates. The core equations governing the LSTM's operation involve these gates and the cell states, ensuring the selective retention and failing of information.

For instance, the cell state update is governed by,

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{1}$$

Where f_t, i_t, and $\tilde{C}_t$ are the outputs of the forget gate, input gate, and cell candidate, respectively. "The integrated functioning of these components in the LSTM regulates the information flow, making LSTMs a key aspect in understanding time-dependent data. However, extended sequences can pose challenges in retaining prior sequence information due to their recurrent nature.

The Attention Mechanism mitigates this by offering a dynamic method to weigh the significance of each input contextually. In other words, not all frames are equally important while extracting temporal information. Some frames capture crucial parts of a sign, while others might be transitional or less informative. Attention Mechanisms can solve this problem by allowing the network to refer back to specific parts of the input sequence [29, 30]. This enhances the model's capability to capture long-term dependencies in longer sequences.

3.3.1. Attention Mechanism

The Attention Mechanism, commonly used in sequence-to-sequence tasks such as machine translation, functions on the principle of sequence alignment. In mathematical terms, the fundamental elements of this mechanism include:

Score Calculation

$$\mathrm{score}(s_t, h_i) = v_a^T \tanh(W_a s_t + U_a h_i) \quad (2)$$

Here, s_t represents the decoder's state, h_i is the encoder's output at timestep i, and v_a, W_a, and U_a are trainable parameters.

Softmax Application

$$\alpha_{ti} = \frac{\exp(\mathrm{score}(s_t,h_i))}{\sum_j \exp(\mathrm{score}(s_t,h_j))} \quad (3)$$

This produces attention weights by normalizing the scores.

Context Vector Formation

$$c_t = \sum_i \alpha_{ti} h_i \quad (4)$$

The context vector is a weighted sum of the encoder outputs, which feeds into the decoder.[29]

3.3.2. Proposed Customization

Although the typical Attention Mechanism is effective for sequence alignment, its complexity might not be advantageous for SLR.

Thus, the author proposed an updated Attention Mechanism that focuses on the temporal dynamic characteristic in SLR.

Score Calculation

$$\mathrm{score}(h_i) = W.h_i \quad (5)$$

Here, each timestep in the sequence gets a score based solely on its content, with h_i being the output at timestep i and W being a trainable weight matrix.

Softmax Application

$$\alpha_i = \frac{\exp(\mathrm{score}(h_i))}{\sum_j \exp(\mathrm{score}(h_j))} \quad (6)$$

This produces weights for each timestep, highlighting the importance of each frame.

Weighted Sum

$$c = \sum_i \alpha_i h_i \quad (7)$$

Unlike original attention, which computes a context for each decoder timestep, this mechanism reduces the entire sequence to a single representative vector. This customization Attention Mechanism responds to the specific challenges posed by SLR. The model emphasizes important gestures or frames by focusing on temporal dynamics, thereby improving recognition accuracy. This guarantees computational efficiency, preserving the core of the Attention Mechanism, and results in a robust and understandable model.

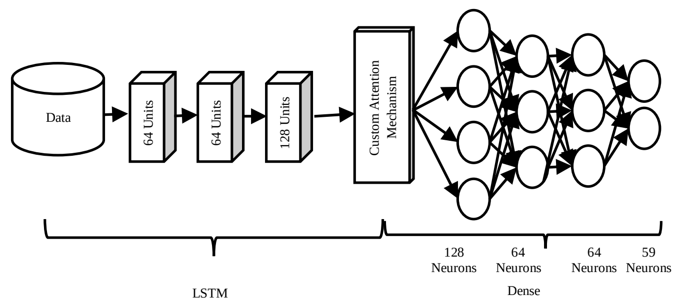

Fig. 3 Proposed LSTM model architecture

The author optimizes the model using a grid search for the LSTM and dense layers. At the end of the grid search, the optimal configuration was determined to include three LSTM layers with units of 64-64-128, one Attention Mechanism, and three dense layers of size 128-64-64. The diagram of the proposed LSTM dense model with customed Attention Mechanism is described in Figure 3.

3.4. Real Time Gesture Recognition

The real-time gesture recognition system, a GUI, is developed for ease. First, apply input video-sign via webcam to the proposed optimized model. A prediction confidence level greater than 80% would be considered to minimize incorrect predictions. Snapshots of the results are shown in Figure 4. The proposed methodology contributes to sign language recognition regarding technology and assistive devices for deaf people. This approach presents a potential solution to improve communication between deaf or dumb people, enhancing their ability to engage with others in a broader societal context.

Fig. 4 Result snapshots

4. Results and Comparison

Our research involved an exhaustive examination of over 100 model architectures using a grid search approach to identify our dataset's optimal baseline LSTM model, which was essential for guiding further enhancements. Model accuracy, loss, confusion matrix, and recognition rate are illustrated below in Figures 5,6, and 7.

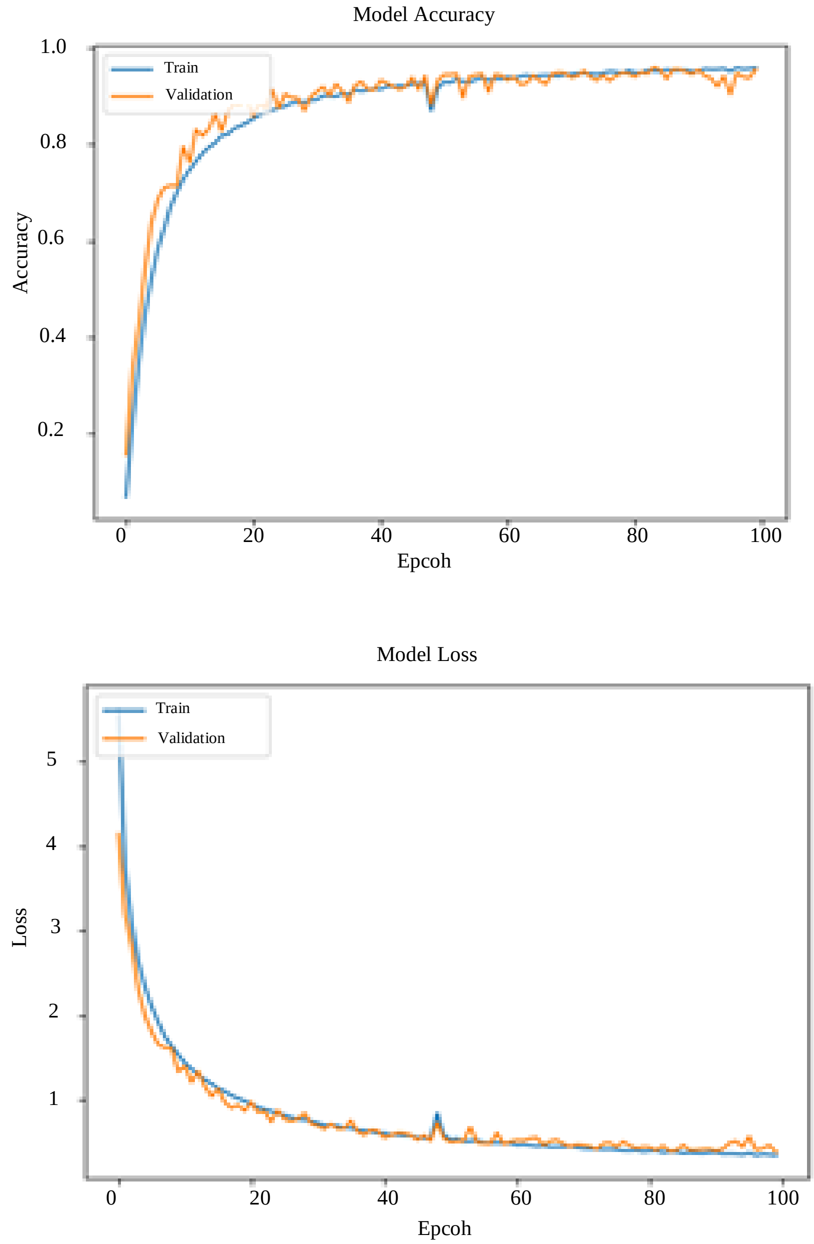

Fig. 5 Model accuracy and loss

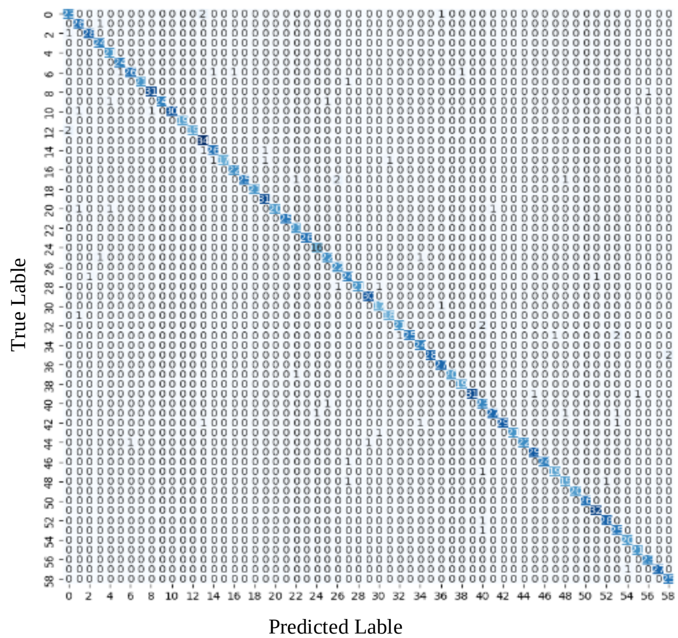

Fig. 6 Confusion matrix

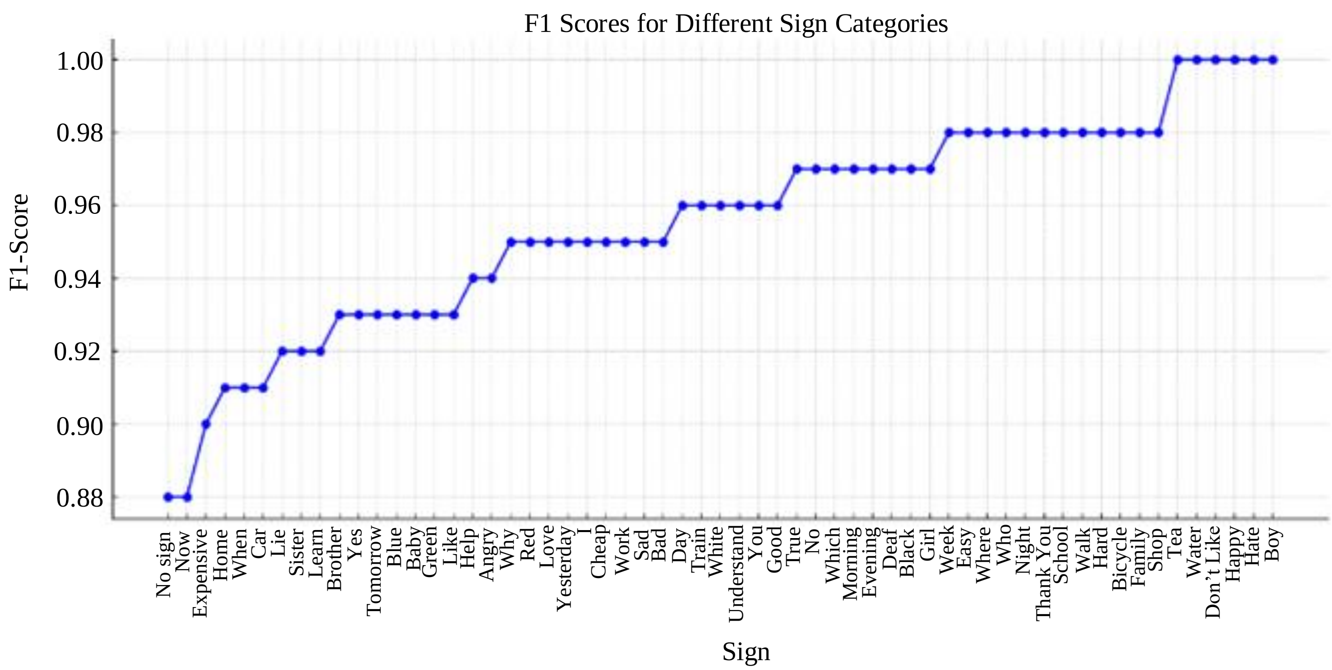

Fig. 7 Recognition rate of different signs

4.1. Enhancement with A Custom Attention Mechanism

To Enhance the baseline LSTM model, we have employed a custom Attention Mechanism developed explicitly for sign language recognition.

This mechanism assigns higher weightage to frames containing more significant information, as discussed earlier. In Table 1 below, we present the results of a comparative analysis between the proposed LSTM model with a custom Attention Mechanism, the proposed LSTM model with traditional Attention Mechanisms, and the proposed LSTM model with no Attention Mechanism.

These results provide a clear perspective of these different Attention Mechanism implementations on model performance.

Table 1. Custom AM comparison

	Precision	Recall	F1-Score	Accuracy
Proposed LSTM model with custom AM	96%	96%	96%	96.08%
Proposed LSTM model with traditional AM	94%	94%	94%	93.50%
Proposed LSTM model with no AM	95%	94%	94%	94.01%

4.2. Validation through Cross-Verification

To validate the robustness of our model, we conducted a comprehensive 5-fold cross-validation. This process resulted in consistently high performance, with a validation accuracy of 93.24% and a training accuracy of 92.99%. Moreover, the model demonstrated a validation loss of 0.48 and a training loss of 0.47. These results highlight the model's reliability and stable performance across varied testing conditions. Results are illustrated in Figure 8.

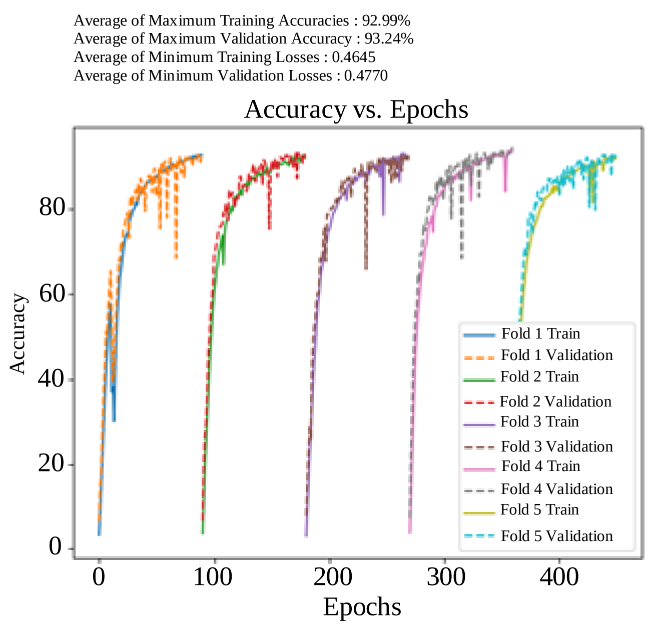

Fig. 8 5-Fold cross validation

4.3. Comparative Analysis

In the field of SLR, the lack of a standardized dataset necessitates researchers using various techniques and custom datasets. Consequently, comparisons between the latest deep learning and traditional machine learning models are irrelevant. For this reason, our comparison is confined to leading-edge deep learning-based models. Additionally, to showcase the robustness of our model, we included comparisons based on the number of signs recognized. Results are shown below in Figure 9 and Table 2.

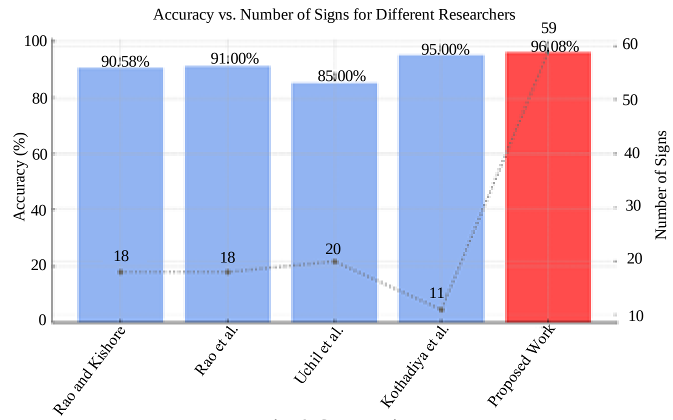

Fig. 9 Comparison

Table 2. Comparative analysis

Year	Researchers	Signs	Dataset Size	Technique	Accuracy
2018	Rao and Kishore [18]	18	--	ANN and MDC	90.58%
2018	Rao et al. [31]	18	180	ANN	91%
2020	Uchil et al. [5]	20	1,120 Videos of 20 Signs	CNN	85%
2022	Kothadiya et al. [24]	11	1100	LSTM and GRU	95%
	Proposed Work	59	36,875 samples	LSTM with customed AM	96.08%

5. Conclusion

A versatile SLR system is essential today, motivating the author to work on a dynamic Indian sign language recognition system. The designed system precisely predicts the sign language feed through the webcam as an input. The achieved results showed that the proposed LSTM-Dense model with a customed Attention Mechanism has maximum accuracy of sign prediction, i.e., 96.08%, compared to the 94.01% accuracy of the LSTM-Dense model without an Attention Mechanism and 93.50% accuracy of LSTM-Dense model with a traditional Attention Mechanism.

Our model is trained on the most extensive dataset of Indian Sign Language (ISL) signs thus far, setting new benchmarks in the number of signs recognized and the accuracy achieved. The development of a real-time detection system lays the foundation for practical applications, making ISL interpretation more accessible and efficient.

5.1. Future Scope

We could extend the dimension of the dataset in future studies. Although the current dataset is extensive, future work could explore more refined signs of regional variations. The current real-time detection system does not offer real-time feedback. We may incorporate it. This would improve the user experience and provide valuable data for future model training. We could optimize the model for mobile devices to make it more accessible.

Acknowledgments

We thank everyone who helped and supported this research. Our heartfelt thanks go to our advisors, colleagues, and peers, who provided invaluable insights and comments throughout this research. In addition, we thank the participants who invested their time and effort to help us obtain the data set required for our study. Their contributions were crucial to the development of the research and its results.

References

[1] Bransford J.D., Brown A.L., and Cocking R.R., "How People Learn: Brain, Mind, Experience, and School," *National Academy Press*, pp. 1-2, 1999. [Google Scholar] [Publisher Link]

[2] Kshitij Bantupalli, and Ying Xie, "*American Sign Language Recognition Using Machine Learning and Computer Vision*," Master of Science in Computer Science Theses, pp. 1-45, 2019. [Google Scholar] [Publisher Link]

[3] Shivashankara S., and S. Srinath, "A Comparative Study of Various Techniques and Outcomes of Recognizing American Sign Language: A Review" *International Journal of Scientific Research Engineering & Technology (IJSRET)* vol. 6, no. 9, pp. 1013-1023, 2017. [Google Scholar]

[4] Anup Nandy et al., "Recognition of Isolated Indian Sign Language Gestures in Real-Time," *International Conference on Recent Trends in Business Administration and Information Processing: Information Processing and Management*, Kerala, India, pp. 102-107, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[5] Aditya P. Uchil, Smriti Jha, and B.G. Sudha, "Vision Based Deep Learning Approach for Dynamic Indian Sign Language Recognition in Healthcare," *International Conference on Computational Vision and Bio Inspired Computing*, vol. 1108, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[6] Anuja V. Nair, and V. Bindu, "A Review on Indian Sign Language Recognition," *International Journal of Computer Applications*, vol. 73, no. 22, pp. 33-38, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[7] P.K. Athira, C.J. Sruthi, and A. Lijiya, "A Signer Independent Sign Language Recognition with Co-Articulation Elimination from Live Videos: An Indian Scenario," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 3, pp. 771-781, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[8] G.R. Sinha, "Indian Sign Language (ISL) Biometrics for Hearing and Speech Impaired Persons: Review and Recommendation," *International Journal of Information Technology*, vol. 9, no. 4, pp. 425-430, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[9] Priyanka Mekala et al., "Real-Time Sign Language Recognition Based on Neural Network Architecture," *43rd Southeastern Symposium on System Theory*, pp. 195-199, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[10] Justin K. Chen, Debabrata Sengupta, and Rukmani Ravi Sundaram, "*Sign Language Recognition with Unsupervised Feature Learning*," CS229 Project Final Report, Stanford University, CA, USA, 2011. [Google Scholar] [Publisher Link]

[11] Joyeeta Singha, and Karen Das, "Recognition of Indian Sign Language in Live Video," *International Journal of Computer Applications*, vol. 70, no. 19, pp. 17-22, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[12] Madhuri Sharma, Ranjna Pal, and Ashok Kumar Sahoo, "Indian Sign Language Recognition Using Neural Networks and KNN Classifiers," *ARPN Journal of Engineering and Applied Sciences*, vol. 9, no. 8, pp. 1255-1259, 2014. [Google Scholar] [Publisher Link]

[13] Kapil Mehrotra, Atul Godbole, and Swapnil Belhe, "Indian Sign Language Recognition Using Kinect Sensor," *Image Analysis and Recognition: Image Analysis and Recognition, ICIAR 2015*, Canada, vol. 9164, pp. 528-535, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[14] Kumud Tripathi, Neha Baranwal, and G.C. Nandi, "Continuous Dynamic Indian Sign Language Gesture Recognition with Invariant Backgrounds," *International Conference on Advances in Computing, Communications, and Informatics (ICACCI)*, pp. 2211-2216, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[15] P.V.V. Kishore et al., "Optical Flow Hand Tracking and Active Contour Hand Shape Features for Continuous Sign Language Recognition with Artificial Neural Networks," *6th International Conference on Advanced Computing (IACC)*, pp. 346-351, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[16] Deepali Naglot, and Milind Kulkarni, "ANN-Based Indian Sign Language Numerals Recognition Using the Leap Motion Controller," *International Conference on Inventive Computation Technologies (ICICT)*, vol. 3, pp. 1-6, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[17] Pradeep Kumar et al., "Coupled HMM-Based Multi-Sensor Data Fusion for Sign Language Recognition," *Pattern Recognition Letters*, vol. 86, pp. 1-8, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[18] G. Ananth Rao, and P.V.V. Kishore, "Selfie Video Based Continuous Indian Sign Language Recognition System," *Ain Shams Engineering Journal*, vol. 9, no. 4, pp. 1929-1939, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[19] Necati Cihan Camgoz et al., "Neural Sign Language Translation," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7784-7793, 2018. [Google Scholar] [Publisher Link]

[20] Pradeep Kumar et al., "A Position and Rotation Invariant Framework for Sign Language Recognition (SLR) Using Kinect," *Multimedia Tools and Applications*, vol. 77, pp. 8823-8846, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[21] Hemina Bhavsar, and Jeegar Trivedi, "Indian Sign Language Recognition Using Framework of Skin Color Detection, Viola-Jones Algorithm, Correlation-Coefficient Technique and Distance Based Neuro-Fuzzy Classification Approach," *International Conference on Emerging Technology Trends in Electronics Communication and Networking*, Surat, India, pp. 235-243, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[22] Bin Xie, Xiaoyu He, and Yi Li, "RGB-D Static Gesture Recognition Based on Convolutional Neural Network," *The 2nd Asian Conference on Artificial Intelligence Technology (ACAIT)*, vol. 16, pp. 1515-1520, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[23] Vanita Jain et al., "American Sign Language Recognition Using Support Vector Machine and Convolutional Neural Network," *International Journal of Information Technology*, vol. 13, no. 3, pp. 1193-1200, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[24] Deep Kothadiya et al., "Deepsign: Sign Language Detection and Recognition Using Deep Learning," *Electronics*, vol. 11, no. 11, pp. 1-12, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[25] Adam A.Q. Mohammed et al., "Multi-Model Ensemble Gesture Recognition Network for High-Accuracy Dynamic Hand Gesture Recognition," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 6, pp. 6829-6842, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[26] Anshul Mittal et al., "A Modified LSTM Model for Continuous Sign Language Recognition Using Leap Motion," *IEEE Sensors Journal*, vol. 19, no. 16, pp. 7056-7063, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[27] Chenguang Song et al., "A Multimodal Fake News Detection Model Based on Crossmodal Attention Residual and Multichannel Convolutional Neural Networks," *Information Processing & Management*, vol. 58, no. 1, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[28] Yuhuang Hu et al., "Overcoming the Vanishing Gradient Problem in Plain Recurrent Networks," *Neural and Evolutionary Computing*, pp. 1-20, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[29] Wei Pan, Xiongquan Zhang, and Zhongfu Ye, "Attention-Based Sign Language Recognition Network Utilizing Keyframe Sampling and Skeletal Features," vol. 8, pp. 215592-215602, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[30] Sarthak Yadav et al., "Comparing Biosignal and Acoustic Feature Representation for Continuous Emotion Recognition," *Proceedings of the 3rd International on Multimodal Sentiment Analysis Workshop and Challenge*, pp. 37-45, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[31] G. Anantha Rao et al., "Selfie Continuous Sign Language Recognition with Neural Network Classifier," *Proceedings of 2nd International Conference on Micro-Electronics, Electromagnetics and Telecommunications: ICMEET 2016*, vol. 434, pp. 31-40, 2018. [CrossRef] [Google Scholar] [Publisher Link]