JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

A REVIEW ON INDIAN SIGN LANGUAGE LEXICAL INTERPRETER

[1] Spoorti S Aldi, [2] Omkar Belamkar, [3] Vishal Nagmule, [4] Suraj Bhajantri, [5] Pushpalatha S Nikkam

[1]B.E Student, [2]B.E Student, [3]B.E Student, [4]B.E Student, [5]Assistant Professor
Department of Information Science and Engineering,
SDM College of Engineering and Technology, Dharwad, India

***Abstract* -** Establishing and maintaining communication among individuals is crucial. Specially Abled individuals – people with speech and hearing problems always depend upon communication through sight. This group comprises around 1% of the Indian community. This serves as the primary defense on why adding a framework that could comprehend Indian Sign Language would be helpful to these people. Here, the study suggests a continuous Indian Sign Language gesture detection system that allows motions to be performed with either one or both hands. It understands different motions and stances and outputs the relevant information. The goal to improve the effectiveness of interaction for the public and those with disabilities by presenting a novel approach to real-time hand gesture interpretation. The suggested solution shows promise for improving accessibility and human-computer interaction, with implications for a range of uses beyond helping people with special needs. Using Long Short-Term Memory and MediaPipe, the developed translator was trained to recognize actions and forecast the sign.

***Keywords* -** Indian Sign Language, Long Short-Term Memory, OpenCV, MediaPipe, Computer Vision, TensorFlow, Recurrent Neural Network **.**

1. INTRODUCTION

India is a country with a rich cultural variety and a diverse range of languages spoken there. Even though spoken languages are the primary means of communication, a sizeable segment of the populace uses Indian Sign Language (ISL), a vital but frequently overlooked mode of expression. Unlike a single unified sign language, India boasts a diverse landscape of regional variations within ISL. This unique characteristic reflects the nation's linguistic and cultural heterogeneity. Understanding these variations is crucial for effective communication and language development. Communication hurdles between the hearing and deaf communities still exist, nevertheless, and they provide serious obstacles in many facets of daily life.

Humans ability to communicate verbally is one among the unusual characteristics in the Human Kingdom. Unfortunately, though, public with verbal or hearing impairments are incapable of carrying out this fundamental human communication. They become socially isolated due to this communication barrier, which also makes it more difficult to communicate effectively. We need a way to break down this barrier to communication since it is impractical to presume that people interacting with these impaired people understands sign language. Communication barriers between the deaf and hearing communities pose significant challenges in areas like education, employment, healthcare, and social interaction. These barriers arise from a lack of awareness and understanding of ISL amongst the hearing population and the absence of readily available resources for interpretation.

More than 18 million people use ISL as their primary medium of communication, making up a sizeable portion of the deaf community in India. However, because of barriers to communication, individuals frequently experience social stigma, have less access to educational opportunities, and have trouble finding work. Beyond the initial communication barrier, these folks face a variety of difficulties. These include being socially isolated, having restricted access to details, and having trouble getting necessary services because there aren't enough qualified sign language interpreters.

In the view of urgent need for increased involvement and communication, this study suggests an "Indian Sign Language Lexical Interpreter" as a possible remedy. This technologically driven strategy makes use of advances in artificial intelligence and language processing to try and close the distance in interaction. This paper discusses research on utilizing a Long Short-Term Memory (LSTM) model to recognize gestures live and provide findings in both text and speech. The technique achieves great accuracy in hand gesture detection by using both vision-based technology and an LSTM model.

The influence of complex backgrounds and light, coupled with the need for huge number of image sequences in the language of signs, have posed considerable challenges to the precision and reliability of hand location in sign recognition. Sign recognition requires not just the spatial domain but also motion information from multiple consecutive video frames, since the language of signs consists of continuous movements. Concurrently, research has long focused on the creation of a compatible and efficient model for sign recognition.

2. LITERATURE REVIEW

The method uses an LSTM network to convert actions into text and speech, with the goal of improving interaction between other people and individuals with impairments. The paper proposes an efficient deep CNN method for sign recognition that leverages transfer learning to handle the lack of a large, labeled action dataset. For gesture analysis and hand identification, the suggested system makes use of TensorFlow object detectors, OpenCV, vision-based technologies, and a histogram clustering technique. With an 98% category accuracy, the results demonstrate that the LSTM network can achieve in identifying actions or gestures in sign language from video sequences. [1].

The article details the approach involves preprocessing, feature extraction, training, and testing stages, employing LSTM and GRU neural network layers to capture semantic relationships effectively. For model evaluation, the source data is split into training as well as validation subsets, and dropout is used to prevent overfitting. The SoftMax function is used to generate the ultimate forecasts. Additionally, a custom dataset for linguistic sign recognition is introduced, and the experimental findings are evaluated, in this article. The main objective is to recognize sign language offsite without the use of third-party tools like digital gloves or devices. [2].

Image processing and machine learning have been widely employed in research and development to produce detection of signs system. A detailed summary of these advancements is provided in this paper. In these systems, the writers emphasize the key technologies and approaches used, including deep neural systems (CNN, Inception model, LSTM), linear discriminant analysis, eigenvalue and eigenvector feature extraction techniques, and skin segmentation and morphological operations preprocessing techniques. The authors also reviewed the difficulties that include the need for bigger and varied datasets, the potential of sensor-based methods, etc. [3].

The paper highlights the relevance of active hand motions in sign language comprehension and includes a survey of studies on fingerspelling. The authors examine the difficulties encountered in vision-based hand motion identification systems and emphasize the necessity for additional research towards a workable vision-based detection of gestures system. The research also fills a gap in the present literature by investigating the evolution of vision-based recognition of hand movements systems and discussing potential future possibilities. [4].

Using Open Pose for hand detection and improving body posture estimation, the paper develops an action recognition system. To improve hand and body feature identification, it trains two 3D neural networks with an emphasis on knowledge transfer and architectural optimization. It also presents a new Open Pose based hand segmentation technique. All things considered, it offers a novel system that has enhanced feature learning and hand segmentation. [5].

The research provides a means of interaction interpretation tool for ISL that employs image processing and machine learning approaches. To extract features, the method preprocesses gesture images with Haar, Cosine, and hybrid wavelet Transforms. The KNN technique is used to categorize these features, which has an 85% accuracy. The method uses pre-processing, feature extraction, and classification to accurately understand sign motions, facilitating communication between ISL users and people unfamiliar with it. [6]

This work offers an on-demand sign language identification tool for ISL based on the fuzzy c-means clustering machine learning technique. The system can recognize 40 ISL words of 75% accuracy for live data. Its possible applications include gesture-controlled robotics, automated houses, game control, Human-Computer Interaction (HCI), and sign language interpretation. [7].

The paper discusses the development of a gesture conversion mechanism using deep learning techniques. This includes the establishment of a hand-locating network based on Faster R-CNN for recognizing sign words in videos and images, as well as the 3D CNN feature extraction network and an LSTM encoding and decoding network to improve recognition accuracy by learning the context of sign vocabulary sequence. Experimental results show a recognition rate high in common vocabulary datasets, outperforming other known methods. [8].

The study offers a software prototype that uses image-based action recognition methods to identify signs . Three processes make up the image-based gesture recognition system: tracking, recognition, and picture preprocessing. The suggested approach includes data collection via camera interface, color filtering, skin segmentation, erosion and dilation noise removal, thresholding, image analysis using blob detection, and contour recognition using convexity hull algorithm. [9].

The research describes a method for live translation of sign vocabulary movements into text and speech that employs technologies such as motion capture, dactyl-based modeling alphabet, and LSTM networks to achieve high gesture recognition accuracy. The research explores the automation of Arabic sign conversion and statistical models for converting speech to sign language, showcasing the system's potential impact on facilitating interaction for the hearing-impaired community across different languages. [10].

The study offers a vision-based system that uses MediaPipe to recognize words in Indian Sign Language (ISL). It uses the method called Random Forest to classify and Google's MediaPipe as a feature extractor, it obtains a high accuracy rate. The suggested process is utilizing a webcam to take pictures of signs, detecting landmarks with MediaPipe, creating a dataset, training unique models, and assessing the models. The method outperforms remaining machine learning algorithms in the study. It also showed how to recognize ISL signs in live. [11].

The creation of a system for recognizing Indian Sign Language using MediaPipe and LSTM models. Three LSTM models—Simple, Bidirectional, and Stacked—were employed, and their performance measures were compared. The outcomes demonstrated that the Bidirectional and Simple LSTMs performed better than the Stacked LSTM. In addition to the application of deep neural networks and multimodal isolated sign language recognition, it emphasizes the value of vision-based and glove-based techniques. [12].

This study addresses advanced strategies for character and word recognition in ISL. They use Meta-Learning to optimize word-level recognition, saving both training time and memory usage while keeping accuracy at 84.2%. Additionally, the study can concentrate on improving the word-level ISL recognition training procedure. This can entail investigating cutting-edge ways of data augmentation for the aim to raise the caliber and volume of the training dataset. [13].

To enhance the recognition, the study presents a data augmentation technique that creates new "compound words" by merging existing signs. This method, tested with an attention-based neural network, achieved a promising 18.85% error rate and improves both accuracy and the variety of recognizable sentences. The paper dives into the model architecture, extraction of features, and evaluation, demonstrating the efficacy of linear interpolation in relation to alternative approaches. [14].

The method for employing deep learning techniques to recognize 26 ISL words in real-time is shown in this study. Two methods are suggested: one utilizes Microsoft Kinect depth + RGB data, while the other acquires of semantic segmentation and a standard RGB camera. High classification accuracies were attained by the authors by training LSTM models and CNNs which has static as well as live actions. They also achieved remarkable IOU and F1 scores by using U-Net with ResNet 101 as the foundation for semantic segmentation. The models underwent real-time testing. [15].

The paper says about a novel CNN-based dynamic hand gesture-based sign word recognition system is presented by converting sign signals into text. There are numerous critical processes in the proposed methodology, including a range of preprocessing methods such as YCbCr conversion, grayscale picture selection, binarization, erosion, and hole filling to achieve hand motion segmentation. After that, a CNN with feature fusion is applied for extraction of features and classification, and a SoftMax classifier is applied for final classification. A dataset of 15 distinct gestures was used for experimental evaluation, and the results showed an amazing 96% average recognition accuracy. [16].

The creation of a lightweight application that uses the MobileNet Classification Model and CNN to recognize ISL is covered in this study. The aim of the application is to help people who are speech- and hearing-impaired communicate using gestures. Using a dataset of binary pictures from a real-time capture stream, the authors used transfer learning to alter the MobileNet V2 model for high accuracy in identifying ISL motions. The effective recognition scale of the model was achieved. Additionally, the report examines possible future advancements like hardware implementation and real-time phrase creation and contrasts the suggested methodology with other studies. [17].

The paper proposes a real-time sign language identification by applying the LSTM model. The methodology involves capturing and analyzing hand movements to recognize words with the help of MediaPipe platform and OpenCV. Landmarks are identified in the images for the face and hands, and a model with LSTM layers is generated for action detection. The Parrot framework is then used to turn the forecasted words into coherent sentences, and the gTTS library is used to further convert the text into voice. The proposed system demonstrates an accuracy of 80% in recognizing hand motions in real-time. [18].

The work presents a sight-based system that utilizes YOLOv4 for words recognition. With a mean Average Precision (mAP) of 92% and real-time translation capabilities, the system is suitable for practical uses. The dataset used here is from ISLCSLTR which is expanded with images from additional signers to improve accuracy, pre-trained network fine-tuning and data augmentation are used. When compared to current techniques, this methodology achieves the greatest results for word-level recognition at real-time speed and accuracy. [19].

The suggested methodology uses a camera-based method for feature extraction and hand recognition, then RNN is used for classification. To balance the dataset and improve training data, the authors record additional films and use the Word-Level dataset. Frame duplication, video cropping, resizing, and the use of a data generator for augmentation are all list of the preprocessing and data preparation. The video frames are subjected to hand detection algorithms, and features are retrieved with the application of techniques like CNN transfer learning and HOG. After being retrieved, an LSTM network is employed to make predictions using the features. [20].

The suggested gesture recognition system has results of 98% accuracy on the whole testing set, which was acquired in a range of indoor circumstances. The combination of CNN + SoftMax gives a better result of 96%. The algorithms with RNN and YOLOv4 also gives better result of 90% and 92% respectively.

Table I. comparison and accuracy of different algorithms

Paper Referred	No. of Words /Alphabets Used	Accuracy of the Gesture/word Recognition						
		Fuzzy c-means	KNN	LSTM+GRU	CNN + SoftMax	RNN	YOLOv4	LSTM + TensorFlow + Feature Extraction
[1]	10	-	-	-	-	-	-	98%
[2]	16	-	-	89%	-	-	-	-
[6]	26	-	85%	-	85%	-	-	-
[7]	40	75%	-	-	-	-	-	-
[26]	15	-	-	-	96%	-	-	-
[19]	24	-	-	-	-	-	92%	-
[20]	20	-	-	-	-	90%	-	-

3. RESEARCH GAP

Although Fuzzy C-Means (FCM) is a useful tool for managing noisy data in images, there is still a research void for recognition ISL gestures. The research gap states its inability to handle intricate hand poses and different signing styles. Because FCM relies on simple distance metrics, it may misclassify because of subtleties of finger positioning and bends. Lastly, FCM prioritizes pixel intensities over the important spatial relationships between fingers that are required to understand some ISL indications. Research depicts that deep learning methods or 3D hand-tracking models should be utilized in addition to FCM to address these issues and raise the precision of ISL detection.

The research gap in adopting K-Nearest Neighbor (KNN) approach for an ISL word interpreter stem from a number of issues. First and foremost, present research typically misses the importance of feature selection and parameter optimization, which are specifically tailored to the nuances of ISL motions. Furthermore, more research is needed on dynamic k selection techniques or adaptive algorithms that lead to complexities of ISL gestures to find the right k value for closest neighbor comparisons, which is an important factor influencing accuracy. The study gap is worsened by the lack of exploration into KNN adaptation strategies to handle the unique characteristics of ISL gestures, such as elaborate hand movements and disparities in signing styles.

The use of LSTM and GRU techniques for an ISL interpreter raises several key difficulties. The fundamental difficulty is that elements peculiar to Indian Sign Language may not be fully exploited. This means that the system may not be capable of detecting minor differences in finger movements, hand orientation changes, or unique finger combinations that distinguish signs. To enhance the efficacy of these recurrent neural network-based models for ISL interpretation, systematic study on suitable architecture configurations, sequence lengths, and training protocols is needed.

While CNNs and SoftMax excel in image classification, their implementation in the ISL Lexical Interpreter has a research gap. The current technique may struggle to capture the dynamic aspect of sign language, possibly missing the critical sequence of hand motions inside a sign. This can lead to misinterpretations since static hand postures may not be sufficient for correct recognition. The development of a precise and flexible ISL interpreter has been hampered by the lack of focus on improving CNN parameters and introducing efficient SoftMax classifiers tailored to ISL gestures.

The research gap in using Recurrent Neural Networks (RNN) for an Indian Sign Language (ISL) Words Interpreter reveals some significant problems. The gradient-vanishing problem, which restricts RNNs' ability to accurately interpret signs with lengthy hand movement sequences, is the core challenge. This might lead to misinterpretations as the model tries to grasp long-term dependencies inside the sign. The adaptation of RNN architectures to ISL's distinct properties, such as complex hand movements and signing variants, is an underexplored area. Furthermore, to increase the model's ability to translate ISL gestures accurately, more study is required to optimize RNN parameters including training protocols and sequence lengths.

The current approach to YOLOv4 technique has a significant lack of comprehensive datasets that appropriately capture the linguistic and regional diversity inherent in Indian Sign Language, limiting the model's generalizability across varied signing styles. It may not be specifically designed for the complexities of ISL signals. This means YOLOv4 may struggle to recognize tiny changes in hand postures and finger placements that distinguish between signs, resulting in mistakes. Addressing this gap by creating domain-specific object detection models trained on large ISL gesture datasets is critical for the system's accuracy and reliability in sign detection.

5. CONCLUSION

The research is made on developing real-time interpreter and translator for ISL using OpenCV, LSTM network and MediaPipe Holistic The introduction highlights the significance of ISL as a mode of communication for a substantial portion of the population, emphasizing the need for understanding and addressing communication barriers between the deaf and hearing communities. ISL translation promotes the perspective or idea of specially abled people interacting with others using machine learning and computer vision technologies. A broad dataset that contains a variety of ISL behaviors and perspectives ensures representation. The literature review section thoroughly investigates the many strategies and technologies used in ISL recognition systems, spanning from deep learning to vision-based approaches. Each study mentioned in the literature review offers unique perspectives on the challenges and accomplishments in ISL recognition, emphasizing the necessity of ongoing research and development in this subject. The identified research gaps highlight the necessity of extra inquiry and progress in ISL identification systems. These gaps reflect the limits of existing methodologies, such as shallow machine learning algorithms like KNN and FCM, which struggle with the intricacies of ISL motions. While RNNs and LSTMs have shown promise, further research is needed on their architecture, training, and adaption to ISL. CNNs with SoftMax activation may not represent the dynamic character of sign language. Finally, YOLOv4's usefulness is hindered by a lack of domain-specific training information. Furthermore, the scarcity of comprehensive datasets reflecting the linguistic and regional peculiarities of ISL presents a substantial difficulty to the research. By tailoring novel approaches to the unique complexities of ISL, the proposed system seeks to overcome the constraints of existing procedures and pave the way for more accessibility and inclusivity.

ACKNOWLEDGEMENT

The completion of “A Review on Indian Sign Language Lexical Interpreter” would be impossible without the dedicated support of Dr. Pushpalatha S. Nikkam, whose guidance and expertise were invaluable throughout. We extend our sincere appreciation to Department of Information science and Engineering for their assistance. Special thanks to all participants for their crucial contributions. We are also grateful to our families and friends for their unwavering encouragement.

REFERENCES

[1] Rohini K Katti, Sujatha C, Padmashri Desai, Shankar G, “Character and Word Level Gesture Recognition of Indian Sign Language”, IEEE 8th International Conference for Convergence in Technology (I2CT), April 2023.

[2] Anuj Chavan, Jayesh Bane, Vishrut Chokshi, Dayanand Ambawade, “Indian Sign Language Recognition Using MobileNet”, IEEE Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI), May 2023.

[3] Sakshi Makar, Kanishka Mohapatra, Ashwin Avate, Mansi Talavadekar, Prof. Surendra Sutar, “Realtime Hand Gesture Recognition using LSTM model and Conversion into Speech”, International Journal of Innovative Research in Technology, March 2022.

[4] Kothadiya, D., Bhatt, C., Sapariya, K., Patel, K., Gil-González, A., Corchado, J.M., “Deepsign: Sign Language Detection and Recognition Using Deep Learning.” Research Gate under Electronics, June 2022.

[5] Vedant Puranik, Varun Gawande, Jash Gujarathi, Ayushi Patani, Tushar Rane, “Video-based Sign Language Recognition using Recurrent Neural Networks”, IEEE 2nd Asian Conference on Innovation in Technology (ASIANCON), August 2022.

[6] Kunal Sharma, Kunwar Abhayjeet Aaryan, Urvashi Dhangar, Richa Sharma, Shweta Taneja, "Automated Indian Sign Language Recognition System Using LSTM models", IEEE International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), November 2022.

[7] Pragati Goel, Ashutosh Sharma, Vikas Goel, Vikas Jain, "Real-Time Sign Language to Text and Speech Translation and Hand Gesture Recognition using the LSTM Model", IEEE 3rd International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), November 2022.

[8] Sneha Sharma, R Sreemathy, Mousami Turuk, Jayashree Jagdale, Soumya Khurana, "Real-Time Word Level Sign Language Recognition Using YOLOv4", IEEE International Conference on Futuristic Technologies (INCOFT), November 2022.

[9] Pratik Likhar, Neel Kamal Bhagat, Rathna G N, "Deep Learning Methods for Indian Sign Language Recognition", IEEE 10th International Conference on Consumer Electronics (ICCE-Berlin), February 2021.

[10] Shruty M. Tomar, Dr. Narendra M. Patel, Dr. Darshak G. T., "A Survey on Sign Language Recognition Systems," International Journal of Creative research Thoughts (IJCRT) 2021, Volume 9, Issue 3 March 2021.

[11] Shreyas Viswanathan, Saurabh Pandey, Kartik Sharma, Dr P Vijayakumar, "Sign Language to Text and Speech Conversion Using CNN", International Research Journal of Modernization in Engineering Technology and Science, May 2021.

[12] Noraini Mohamed, Mumtaz Begum Mustafa, Nazean Jomhari, "A Review of the Hand Gesture Recognition System: Current Progress and Future Directions", IEEE Access Volume 9., November 2021.

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

[14] Md Abdur Rahim, Jungpil Shin, Md Rashedul Islam, "Dynamic Hand Gesture Based Sign Word Recognition Using Convolutional Neural Network with Feature Fusion", IEEE 2nd International Conference on Knowledge Innovation and Invention (ICKII), March 2020.

[15] Shubhendu Apoorv, Sudharshan Kumar Bhowmick, and R Sakthi Prabha – "Indian sign language interpreter using image processing and machine learning", IOP Conference Series: Materials Science and Engineering 872:012026, June 2020.

[16] Natsuki Takayama, Hiroki Takahashi, "Data Augmentation Using Feature Interpolation of Individual Words for Compound Word Recognition of Sign Language", IEEE International Conference on Cyberworlds (CW), October 2020.

[17] Muneer al-Hammadi, Ahulam Muhammad, Wadood Abdul, Mansour Alsulaiman, Mohammed a. Bencherif Tareq s. Alrayes, Hassan Mathkour, and Mohamed Amine Mekhtiche, "Deep Learning-Based Approach for Sign Language Gesture Recognition with Efficient Hand Gesture Representation", IEEE Access Volume 8, November 2020.

[18] Muthu Mariappan h, Dr Gomathi V, "Real-Time Recognition of Indian Sign Language", IEEE International Conference on Computational Intelligence in Data Science (ICCIDS), February 2019.

[19] Hema B N., Sania Anjum, Umme Hani, Vanaja P., Akshatha M., "Sign Language and Gesture Recognition for Deaf and Dumb People", International Research Journal of Engineering and Technology (IRJET), e-ISSN: 2395-0056 Volume: 06 Issue: 03, Mar 2019.

[20] He Siming, "Research of a Sign Language Translation System Based on Deep Learning", IEEE International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM), November 2019.