Indian Sign Language Recognition

Yashasvi Vaidya, Prachi Vaishya, Sahil Yadav

Thakur College of Engineering and Technology, Mumbai

Abstract:
Hand signs are an effective form of human-to-human communication that has a number of possible applications. Being a natural means of interaction, they are commonly used for communication purposes by speech-impaired people worldwide. In fact, about one percent of the Indian population belongs to this category. The research methodology involves the collection of a substantial dataset comprising diverse ISL gestures performed by native signers. To achieve robust recognition, a multi-modal approach is employed, combining computer vision and deep learning techniques. The proposed model leverages convolutional neural networks (CNNs) to extract spatial features from video frames, alongside recurrent neural networks (RNNs) to capture temporal dependencies within the sign sequences.

Introduction:

The sign language is used widely by people who are deaf-dumb; these are used as a medium for communication. A sign language is nothing but composed of various gestures formed by different shapes of hand, its movements, orientations as well as the facial expressions. These gestures are generally used by deaf-dumb people in order to express their thought. Dumb-deaf persons faces communication barrier in public places while interacting with normal person, such as in bank, hospital and post offices. Sometimes the deaf needs toseek the help of the sign language interpreter so as to translate their thoughts to normal people and vice versa.However, this way turns out to be very costly and does not work throughout the life period of a deaf person. So a system which can automatically recognize the sign language gestures becomes a necessity. Introducing such asystem would lead to minimizing the gap between deaf and normal people in society. The sign language in use at a particular place depends on the culture and spoken language at that place. Indian sign language (ISL) is used by the deaf community in India. ISL is a standard and well-developed way of communication for hearing-impaired people in India and speaking in English. Different symbols are involved for different alphabets in Indian Sign Language. It consists of both word-level gestures and finger spelling. This paper presents a method for the automatic recognition of static gestures in the Indian sign language alphabet. The signs considered for recognition include 17 letters of the English alphabet.In the proposed approach, the main focus is on the classification and recognition of the Indian sign language given by the dumb-deaf user in real-time. Thus, the speed and simplicity of the algorithm is important. The system approach involves segmenting the hand based on the skin color statistics, then converting that segmented
image into binary, apply feature extraction on the binary image

Literature Survey

A real-time recognition approach based on Eigen value-weighted Euclidean distance classification of signs was proposed by J. Singha et al. [1].
In order to classify the signs, P. Kishore et al. [2] devised a system that uses Artificial Neural Networks (ANN) to find active contours from boundary edge maps. Another method for hand gesture identification in a real-time setting used the Viola Jones algorithm with LBP functions [3]. One benefit was that it needed less processing power to identify the movements. In hand processing, segmentation is the first and most crucial step. Otsu's technique often produced a reasonably high accuracy rate [4]. A moving block distance parameterization method was attempted to bypass the segmentation and initialization stages [5].

High-precision static symbols and 33 basic word units were used.

The majority of these studies relied on techniques like feature extraction and pattern recognition [6]. Still, most of the time a system with one functionality is insufficient. Hybrid strategies were thus presented as a solution to this issue. For example, A. Nandy et al. [7] classified gestures from oriented histogram characteristics using hybrid techniques combining K-Nearest Neighbor (KNN) and Euclidean distance. This approach's drawback was its subpar execution of comparable gestures. K Manjushree et al. [8] employed feature matching and a histogram of directed gradients for single-handed sign categorization. Using SVM and PCA features, S. Kanade et al. [9] created a system using a bespoke dataset and achieved high accuracy.

ISL recognition for both single-handed and double-handed character signs was proposed by A. Sahoo [10]. B-spline approximation was utilized by Geetha M et al. [11] to match the shapes of static gestures representing ISL alphabets and numbers. A method for classifying word symbols utilizing the Neuro-Fuzzy methodology and natural language processing (NLP) technology to display the final word was proposed in Ref. [12]. A technique for hand gesture recognition using the AdaBoost algorithm and haar-like features was presented by Q. Chen et al. [13]. To properly understand motions, they also described random grammars that are free of context. The technique based on the condensation algorithm was shown to be inferior to the combination of PCA and the local coordinate system, which provided excellent calculation accuracy [14].

Nonetheless, researchers wanted to find a quicker solution for real-time systems. Deep Learning technology breakthroughs have made it possible to automate the use of different image recognition models for image identification. Convolutional neural networks, for instance, have advanced significantly in the area of deep learning recently [15, 16].

G. Jayadeep et al. [17] classified these movements and converted them into text using an LSTM (Long Short Term Memory) after using a CNN (Convolutional Neural Network) to extract features from the images. The InceptionV3 model was presented by Bin et al. [18] to identify static indications using depth sensors. The gesture segmentation and feature extraction processes were removed. Vivek Bheda et al. presented an approach in Ref. [19] for classifying images for each digit (0–9) and American Sign Language letter using deep convolutional neural networks and a mini-batch supervised learning method of stochastic gradient descent.

After reading through these studies, the authors were inspired to develop a new dataset and an algorithm that would function flawlessly on it without compromising the video detection accuracy. To make the system usable outside of controlled situations, the paper's authors have also tackled the issue of background reliance.

Proposed Work:

Dataset:

Models suitable:

Convolutional Neural Networks (CNNs):

Application: CNNs are effective for image and video processing tasks. They can be employed to extract spatial features from video frames of sign gestures in ISL. Transfer learning with pre-trained CNNs (e.g., ResNet, VGG) can also enhance performance when data is limited.

Recurrent Neural Networks (RNNs):

Application: RNNs are well-suited for sequential data. In the context of ISL, RNNs can capture the temporal dependencies inherent in sign language gestures. Long Short-Term Memory Networks (LSTMs) and Gated Recurrent Units (GRUs) are popular RNN architectures that can be applied to sequence modeling.

Long Short-Term Memory Networks (LSTMs):

Application: LSTMs are a type of RNN designed to handle long-term dependencies. They can be employed to model sequential aspects of sign language gestures, making them suitable for ISL recognition tasks.

Gated Recurrent Units (GRUs):

Application: Similar to LSTMs, GRUs are another variant of RNNs. They are known for their simpler architecture while still being effective in capturing sequential patterns. GRUs can be used for temporal modeling in ISL recognition.

3D Convolutional Neural Networks (3D CNNs):

Application: As sign language involves dynamic movements over time, 3D CNNs can be employed to capture both spatial and temporal information simultaneously. This makes them suitable for processing video sequences of ISL gestures.

Gesture Transformers:

Application: Transformer architectures, originally designed for natural language processing, have shown success in various computer vision tasks. They can be adapted to handle sequential data in ISL videos, capturing both spatial and temporal dependencies effectively.

Joint Models (CNN-RNN Fusion):

Application: Combining the strengths of both CNNs and RNNs, a joint model can be constructed for ISL recognition. CNNs can handle spatial features, while RNNs capture temporal dependencies, providing a comprehensive understanding of sign gestures.

Preprocessing:

The image is made ready for feature detection and extraction in this phase. To preserve uniformity of scale, the dimensions of all the images are kept the sIn the context of Indian Sign Language (ISL) translation, data preprocessing plays a pivotal role in preparing a comprehensive and representative dataset for model training. The collected dataset encompasses a diverse range of hand gestures, reflective of the intricacies inherent in ISL communication. Following meticulous data labeling, the cleaning process involves eliminating noise and irrelevant frames, ensuring dataset integrity. Resizing and cropping images focus on standardizing resolution and isolating the region of interest containing the hand gestures. The normalization of pixel values facilitates stable model training by bringing uniformity to the data. Given the temporal nature of sign language, particularly in videos, sequence formation involves organizing frames into meaningful temporal sequences, aligning with the sequential aspects of sign gestures. Augmentation techniques introduce variability, aiding the model in learning robust features and improving generalization. Attention to class balance and encoding categorical labels ensures the model's ability to effectively distinguish between various signs within the rich tapestry of Indian Sign Language. Overall, this meticulous data preprocessing pipeline lays the foundation for robust and accurate ISL translation models.ame.

RESULT AND DISCUSSION:

An essential aspect of the research focused on the cultural adaptation of the ISL translation system. Findings highlighted the significance of considering regional variations and cultural nuances within ISL, emphasizing the need for continuous updates and community collaboration. The system demonstrated adaptability to diverse expressions and signs, showcasing its potential to cater to the rich and varied nature of ISL.While the ISL translation system showed commendable accuracy, challenges in achieving real-time processing were identified. Processing delays were observed, particularly in dynamic environments with rapid hand movements. This limitation calls for further optimization of the algorithm and exploration of advanced real-time processing techniques to enhance the system's responsiveness.Feedback from user testing highlighted the importance of a user-friendly interface for both deaf and non-deaf users. Customization options for the interface, including font size, color schemes, and gesture display preferences, were well-received. However, challenges in ensuring accessibility for users with diverse needs were identified, suggesting the need for additional features and adjustments to enhance inclusivity.The research findings pave the way for several future directions. Improvements in gesture recognition algorithms, addressing real-time processing challenges, and deepening cultural adaptation are key areas for further investigation. Collaboration with the ISL community for ongoing feedback and updates remains crucial for the sustained relevance and effectiveness of the translation system.

CONCLUSION:

In conclusion, the use of machine learning for identifying high-risk crime hotspots is a valuable tool for law enforcement. It empowers proactive policing, efficient resource allocation, and long-term crime prevention, ultimately enhancing public safety and community well-being. Careful data management and ethical considerations are essential for the responsible deployment of this technology.

REFERENCES:

[1] Singha J, Das K. Recognition of Indian sign language in live video. Int J Comput Appl 2013;70(19):17–22.

[2] Kishore PVV, Kumar DA. Optical flow hand tracking and active contour hand shape features for continuous sign language recognition with artificial neural networks. In: IEEE 6th international conference on advanced Computing; 2016.

[3] Swamy Shanmukha, Chethan MP, Gatwadi Mahantesh. Indian sign language interpreter with android implementation. Int J Comput Appl 2014:975–8887.

[4] Agrawal SC, Jalal AS, Bhatnagar C, Ieee. Recognition of Indian sign language using feature Fusion. 2012.

[5] Aviles-Arriaga HH, Sucar-Succar LE, Mendoza-Duran CE, Pineda-Cortes LA. A comparison of dynamic naive bayesian classifiers and hidden markov models for gesture recognition. J Appl Res Technol 2011;9:81–102.

[6] Rokade Yogeshwar I, , et alJadav Prashant M. Indian sign language recognition system. In: International Journal of Engineering and Technology July; 2017.

[7] Nandy Anup, Prasad Jay Shankar, Mondal Soumik, Chakraborty Pavan, Nandi Gora Chand. Recognition of isolated Indian sign language gestures in real time. In:

International Conference on Business Administration and Information Processing;
2010.
[8] Manjushree K, Divyashree. Gesture recognition for Indian sign language using HOG
and SVM. International Research Journal of Engineering and Technology 2019;6
(7).
[9] Kanade Sudhir S, Deshpande Padmanabh D. Recognition of Indian sign language
using SVM classifier. International Journal of Scientific Research and Development
2018;2(3).
[10] Sahoo Ashok Kumar, Kumar Ravulakollu Kiran. Vision based Indian sign language
character recognition. J Theor Appl Inf Technol 2014;67(3).
[11] Geetha M, Manjusha UC. A vision based Recognition of Indian sign language
Alphabets and Numerals Using B-Spline Approximation International Journal on
Computer Science and Engineering (IJCSE). 2012.
[12] Bhavsar Hemina, Trivedi Jeegar. Indian sign language recognition using
framework of skin color detection, Viola- Jones algorithm, correlation-coefficient
technique and distance based neuro-fuzzy classification approach. Emerging
Technology Trends in Electronics, Communication and Networking 2020;1214:
235–43.
[13] Chen Q, Georganas ND, Petriu EM. Hand gesture recognition using Haar-like
features and a stochastic context-free grammar. IEEE Trans Instrum Meas 2008;57
(8):1562–71.
https://doi.org/10.1109/TIM.2008.922070.
[14] Dan L, Ohya J. Study of recognizing multiple persons' complicated hand gestures
from the video sequence acquired by a moving camera. In: Rogowitz BE,
Pappas TN, editors. Human Vision and Electronic Imaging Xv, vol. 7527; 2010.
[15] Sahoo Ashok K, Mishra Gouri Sankar, Kumar Ravulakollu Kiran. sign language
recognition: state of the art. In: ARPN Journal of Engineering and Applied Sciences;
2014.
[16] Bachani Shailesh, Dixit Shubham, Chadha Rohin, Bagul Prof Avinash. sign
language recognition using neural network. International Research Journal of
Engineering and Technology (IRJET) 2020;7(4).
[17] Jayadeep G, Vishnupriya NV, Venugopal V, Vishnu S, Geetha M. Mudra:
convolutional neural network based Indian sign language translator for banks. In:
4th International Conference on Intelligent Computing and Control Systems
(ICICCS), 2020; 2020. p. 1228–32.
[18] Xie B, He Xy, Li Y. RGB-D static gesture recognition based on convolutional neural
network. J Eng 2018;2018(16):1515–20.
[19] Vivek Bheda and N. Dianna Radpour. Using Deep Convolutional Net-works for
Gesture Recognition in American sign language. In Department of Computer
Science Department of Linguistics State University of New York at Buffalo.