DOI: 10.17148/IARJSET.2023.10448

INDIAN SIGN LANGUAGE RECOGNITION

Mr. Mahesh Rathod[1], Mr. Prajwal Patil[2], Mr. Saurabh Patil[3], Prof. Saurabh Saoji[4]

Student, Computer Engineering, Nutan Maharashtra Institute of Engineering and Technology, Pune, India[1]

Student, Computer Engineering, Nutan Maharashtra Institute of Engineering and Technology, Pune, India[2]

Student, Computer Engineering, Nutan Maharashtra Institute of Engineering and Technology, Pune, India[3]

HOD, Computer Engineering, Nutan Maharashtra Institute of Engineering and Technology, Pune, India[4]

Abstract: People can interact and exchange ideas and sentiments through communication. For the deaf community to engage with the community, there are several hardships. The people converse with one another by using sign language. There is a technology that can translate sign languages into a comprehensible form to enable interaction with everyday people. This project aims to develop a real-time system that can translate Indian Sign Language (ISL) into text. The majority of the work is handcrafted. In this article, we provide a deep learning method that uses a convolutional neural network to classify signs. In the first stage, we create a classifier model that uses the numerical signs and the Python-based Keras convolutional neural network implementation. A second real-time system that used skin segmentation to locate the Region of Interest in the frame that displays the bounding box was used in phase two. To forecast the sign, the segmented region is fed into the classifier model. The system has a 99.56% accuracy rate for the same subject and a 97.26% accuracy rate in low light. With a varied background and image capture angle, the classifier was found to be improving The RGB camera system is the focus of our methodology.

Keywords: Deep Learning, Convolutional Neural Networks, real-time system, Computer Vision, Training, User Interaction, Indian Sign Language

I. INTRODUCTION

When it comes to physically disabled people, both the deaf and dump communities employ various sign languages. In the world, many languages are utilized by people to offer communication. American Sign Language, Chinese Sign Language, Indian Sign Language, and others are among the several sign languages. In each instance, the symbols alter depending on whether motion, single-handed, or double-handed representations are present. In some circumstances, dynamic symbols are utilized for words like "hello," "Hai," etc. instead of static symbols to represent letters. These communities' ability to communicate with one another will be enabled through a real-time system. It can be converted to any language after being transformed using the Computer Vison method. To create an accurate and efficient system, numerous studies have been conducted in this area. The researchers' earlier approach utilized a handcrafted feature, but it was constrained and used under particular circumstances.

The majority of works rely on feature extraction based on HOG, SIFT, LBP, etc., as well as pattern recognition. However, most of the time a system using just one feature is insufficient, hence the hybrid technique was developed to address this issue. However, in a real-time system, we require quicker approaches to problem-solving. Nowadays, we use parallel implementation to increase the processing speed of our computers. Our system uses a single core to solve issues the majority of the time. Parallel computing can be used to solve problems using the GPU system, which has more cores than a CPU system. We can model a self-learning system for our needs using the deep learning methodology. One of the most popular deep learning systems that can handle any computer vision issue is the convolutional neural network. For the real-time implementation of our technique, we used a region of interest-convolutional neural network.

II. PROBLEM STATEMENT

The communication gap between hearing-impaired individuals and the rest of society is a significant challenge that can hinder their ability to access essential services, participate in social interactions, and pursue educational and professional opportunities. Sign language is the primary means of communication for the deaf and hard of hearing, but it requires an interpreter, which can be both costly and time-consuming. Hence, developing a sign language recognition system that can translate sign language into text or speech can bridge this gap and facilitate communication between the deaf and hearing communities. However, sign language recognition poses unique challenges due to variations in signing styles, facial expressions, hand orientation, and lighting conditions. Therefore, the goal of this project is to develop an accurate and robust sign language recognition system that can identify and translate sign language gestures in real-time.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 4, April 2023

DOI: 10.17148/IARJSET.2023.10448

III. PREVIOUS WORK

A One of the most popular trends in technology today is computer vision, which is used in many AI-based systems including robots, cars, markets, etc. The system has a greater influence on object detection and image categorization issues. This technique can be used to implant the sign language system. In the earlier systems, numerous more techniques were employed.

[1] utilized a foundation for the ISL Recognition system in literature. Colour segmentation is done using a glove, and PCA (Principal Component Analysis) is employed for recognition. Every 20th frame, real-time data frames are used as input to perform recognition. This method had issues with the sign that had both overlapping and motion. PCA and the fingertip algorithm are both utilized for recognition. Recent studies have concentrated on static indicators of ISL [2] from photo or video sequences that were captured using data glove or coloured glove under controlled circumstances such a single background and specialized gear. In the system, the light and position are more significant. To work under these circumstances, the signer needs to be knowledgeable of the system.

Pre-processing Otsu's thresholding can be done in a variety of ways, including by considering skin tone, motion-based segmentation, and backdrop subtraction [3–5].Scale-invariant feature transform, Fourier descriptors, and wavelet decomposition are used in the feature extraction phase. K Nearest Neighbour (KNN), Hidden Markov Models (HMM), Multiclass Support Vector Machines (SVM)[6], Fuzzy systems, Artificial Neural Networks (ANN), and many other classifiers are used to categorize signs.

implemented an edge detection approach for hand gesture identification in another study [7]. Edge detection and sorting characteristics in the database are used to retrieve the frame features. Utilize template matching to forecast the gesture using the newly built database. The smallest distance is used in this case to match templates. Both static symbols and dynamic gestures can be recognized by the system. A fuzzy membership function is used by the system to extract the spatial properties of signs utilizing a fuzzy [8] based approach. The Nearest Neighbour classifier is paired with a suitable symbolic similarity measure.

Using the Microsoft Kinect sensor device, Raheja [9] et al. devised a gesture detection system for Indian sign language. They conducted experiments using Kinect photos in RGB and Depth. According to the research, employing RGB-D photos improves the system's accuracy. The HU-Moments, which are moments that are angle, position, and shape invariant, are extracted by the model and fed to the SVM classifier as features. Indian sign language has an android app-based system designed by Pranjali Loke[10]. all. Images are collected by the Android system and sent to the server. The server system sends these photos to the MATLAB application, where the system is trained using a neural network and features are extracted using the Sobel operator. The system analyses the photos using pattern recognition and classification to produce text as the result. A system to recognize American Sign Language (ASL) using the depth images captured by the Kinect sensor was created by Beena M.V. et A total of 1000 photographs of each numerical sign were used to train the system. The approach produced a 99.46% accuracy for the depth pictures after extracting features from the block-processed images and training an artificial neural network (ANN).On the, the system has been taught for quicker execution. Convolutional Neural Network [2017-2] (CNN) with SoftMax classification is used as an extension of the work for 33 static symbols of Kinect depth images. The implementation demonstrates that the handcrafted features become insufficient for classification purposes as the number of classes rises. The CNN structure will perform better in terms of accuracy compared to other conventional methods because it can learn from the provided training data.

IV. PROPOSED SYSTEM ARCHITECTURE

Indian sign language is a sophisticated technique that uses both hands. Convolutional neural networks are applied to the image in an effective method to improve classification accuracy and for practical use. The suggested system's fundamental processes are illustrated below.

Steps
1: Enter the video frame or picture.
2. Track down the handicraft.
3. Take the feature out.
4. Sorting and forecasting.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 4, April 2023
DOI: 10.17148/IARJSET.2023.10448

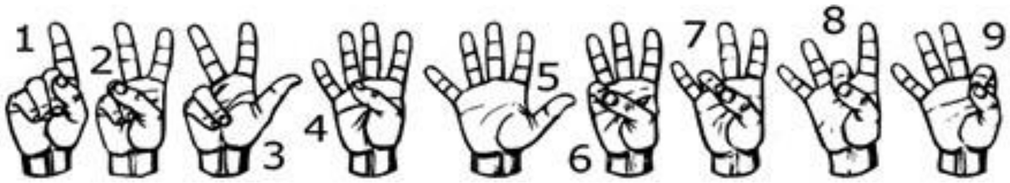

Fig. 1 signs

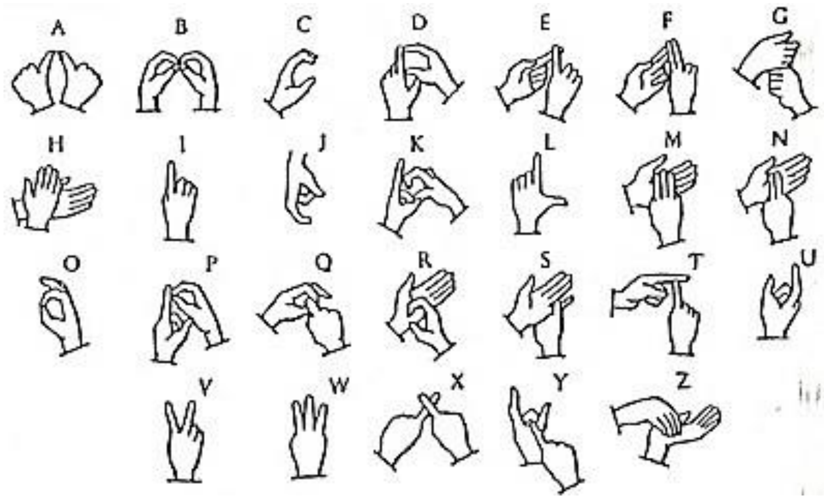

Fig. 2 signs

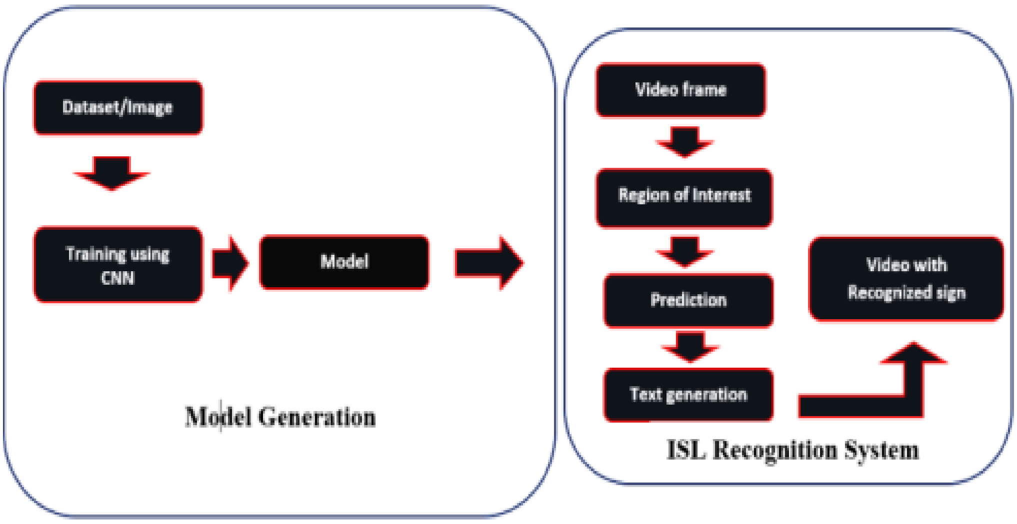

Fig. 3 System Architecture

The majority of the object detection issues use an image data set and bounding box mapping to train the model. It costs money to label the bounding box for each image. Additionally, using skin segmentation, we suggested a region of interest predictor. We crop the image from the segmented, constrained region and feed it to the classifier for prediction.

In the first section, we input the video frame and use CLACHE (Contrast Limited Adaptive Histogram Equalization) to adjust the image's lightness using the LAB colour system. Apply blurring using Gaussian blurring to the original image in the following step. We do a thresholding operation using the HSV color space to generate the skin. We can change the

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 4, April 2023

DOI: 10.17148/IARJSET.2023.10448

threshold values while running in some circumstances where the light variation is high. Finding the biggest contours in the segmented photos and drawing a rectangular box around the area that displays the output categorized result as text are the final steps. The convolutional neural network model is fed the bounding box in order to produce the sign as text.

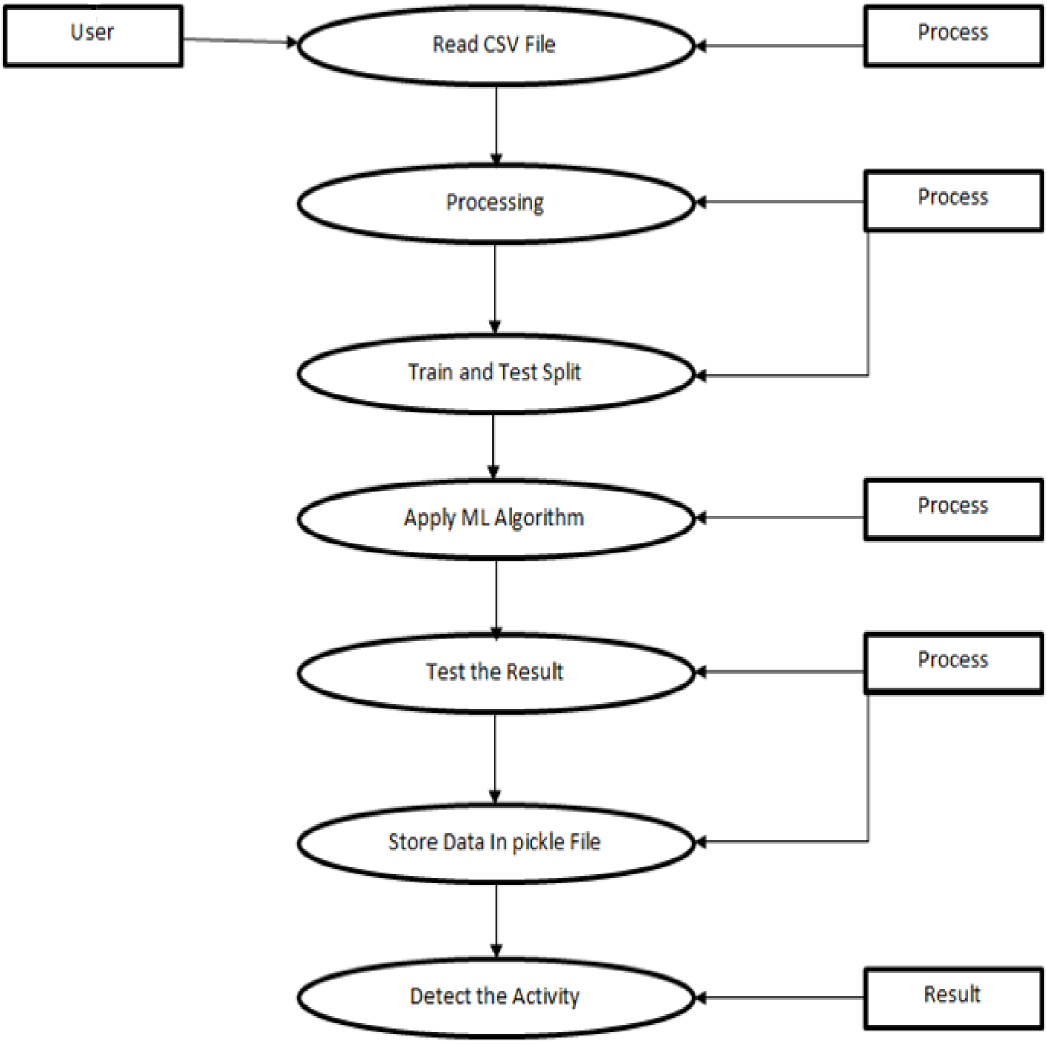

Fig. 4 Data flow diagram

V. CONCLUSION

The real-time system has been built for numeral signals from 0-9. This is the first step towards the recognition of Indian Sign Language. The 3000 static symbols of RGB images from regular camera images were used to teach the system. For testing, the system used 100 photos for each symbol. The model was developed through the effective use of a region-based convolutional neural network in a deep learning system. For the same subject, the system achieved an accuracy of 99.56% during testing, but in low light, the accuracy dropped to 97.26%. Add more symbols from the alphabets of the Indian sign language's static symbols in the future, including the double hand notation. The dataset must be expanded in order to address the low light issues.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 4, April 2023

DOI: 10.17148/IARJSET.2023.10448

REFERENCES

[1] Divya Deora, Nikesh Bajaj, Indian Sign Language Recognition, 2012 1st International Conference on Emerging Technology Trends in Electronics, Communication and Networking, IEEE 2012-978-1-4673-1627-9/12.
[2] Anuja V. Nair, Bindu V., A Review on Indian Sign Language Recognition, International Journal of Computer Applications (0975 – 8887) July 2013, Volume 73– No.22.
[3] Jorge Badenas, Josee Miguel Sanchiz, Filiberto Pla, Motion-based Segmentation and Region Tracking in Image Sequences, Pattern recognition 2001, 34, pp. 661-670.
[4] Ping-Sung Liao, Tse-Sheng Chen, Pau-Choo Chung, 2001, A Fast Algorithm for Multilevel Thresholding, Journal of Information Science and Engineering 17, pp. 713-727
[5] Dr. Alan M McIvor, Background subtraction techniques, Image and Vision Computing, Newz Zealand 2000 (IVCNZ00).
[6] Aseema Sultana, T. Rajapushpa, Vision Based Gesture Recognition for Alphabetical Hand gestures Using the SVM Classifier, International Journal of Computer Science and Engineering Technology, Volume 3, No. 7, 2012.
[7] Purva A. Nanivadekar, Dr. Vaishali Kulkarni, Indian Sign Language Recognition: Database Creation, Hand Tracking and Segmentation, International Conference on Circuits, Systems, Communication and Information Technology Applications, IEEE 2014,978-1-4799-2494-3/14.
[8] Nagendraswamy H S, Chethana Kumara B M, Lekha Chinmayi R, Indian Sign Language Recognition: An Approach Based on Fuzzy-Symbolic Data, 2016 Intl. Conference on Advances in Computing, Communications and Informatics (ICACCI), Sept. 21-24, 2016, 978-1-5090-2029-4/16.
[9] J. L. Raheja , A. Mishra, A. Chaudhary, Indian Sign Language Recognition Using SVM, Pattern Recognition and Image Analysis, 2016, Vol. 26, No. 2, pp. 434–441.
[10] Pranali Loke, Juilee Paranjpe, Sayli Bhabal, Ketan Kanere, Indian Sign Language Converter System Using An Android App. ,International Conference on Electronics, Communication and Aerospace Technology, 2017 IEEE ,978- 1-5090-5686-6/17.
[11] M.V. Beena and M.N. Agnisarman Namboodiri, ASL Numerals Recognition from Depth Maps Using Artificial Neural Networks, Middle-East Journal of Scientific Research 25 (7): 1407-1413, 2017,ISSN 1990-9233.
[12] Beena M.V., Dr. M.N. Agnisarman Namboodiri, Automatic Sign Language Finger Spelling Using Convolution Neural Network: Analysis, International Journal of Pure and Applied Mathematics, Volume 117 No. 20 2017, 9-15.