International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Adaptive Sign Language to Bridge Communication Gap for Deaf and Dumb

Dr Chandru A S[1], Kavyashree H S[2], Karthik J[3], Vibha Sree Gowda[4], Spoorthi R[5]

[1] Associate Professor and HOD, Information Science and Engineering, VVIET, VTU Belagavi
email: dr.chandru.vviet@gmail.com
[2] Department of Information Science and Engineering, VVIET, VTU Belagavi
email: kavyashreehs23@gmail.com
[3] Department of Information Science and Engineering, VVIET, VTU Belagavi
email: jkarthik712@gmail.com
[4] Department of Information Science and Engineering, VVIET, VTU Belagavi
email: vibhagowda1624@gmail.com
[5] Department of Information Science and Engineering, VVIET, VTU Belagavi
email: spoorthiiraghu@gmail.com

ABSTRACT –

Additionally, hand signs are a natural way to communicate and have many practical uses. Sign language is frequently used by those who are mute to express themselves. This study introduces a technique for recognizing Indian sign language letters and numerals in live video streams using the Bag of Visual Words concept. The system uses background removal and skin color for segmentation, and it generates predictions for both speech and text. In order to assign the signs to their respective labels, we take SURF features out of the pictures and create rate histograms. In order to classify, support vector machines and convolutional neural networks are used. In order to make things easily accessible, we also created an interactive Graphical User Interface.

Static and dynamic signs, single-handed and double-handed signs, and multiple signs for the same alphabet in various parts of India are all part of Indian Sign Language (ISL). It makes implementing such a plan extremely challenging. Furthermore, there isn't a standard dataset. These all demonstrate how intricate Indian sign language is. Scholars have recently begun to investigate this field. The sensor-based approach and the vision-based approach are the two primary methods that are frequently employed in sign language recognition. In contrast to the vision-based technique, which uses web cameras to record video or images, the sensor-based approach uses gloves or other devices that can identify finger gestures and convert them into equivalent electrical impulses for sign determination.

KEYWORDS - Adaptive Sign Language, Deaf and Dumb Communication, SVM, CNN, SURF, Gesture Recognition, Feature Extraction, Pre-Processing, Reverse Recognition, Gaussian Filtering, Skin Segmentation, Canny Edge Detection.

I. INTRODUCTION

Hand signals are a natural way for people who are mute to interact with one another. With output predictions in both text and speech formats, this research study presents the Bag of Visual Words technique for identifying Indian sign language letters and numbers in live video feeds. We employ background subtraction and skin color techniques to divide the movie into segments. To map the signs to their associated labels, we create histograms using the SURF features that we extract from the photos. Lastly, for classification, Convolutional Neural Networks and Support Vector Machines are used. For convenience, we also created an intuitive Graphical User Interface. Indian Sign Language (ISL) uses both single-handed and double-handed signs, as well as static and dynamic signs. The same alphabet has many signs in different parts of India. The methodology presented in this paper aims to develop a broad, diverse, and reliable real-time alphabet (A-Z) and digit (0-9) identification system for Indian Sign Language. Here, we have identified indications based on photos (viewed through a webcam). Bridging the communication gap requires real-time, accurate, and efficient judgment on ISL sign recognition.

II. LITERATURE REVIEW

Deep learning algorithms have been a major factor in recent advances in sign language recognition (SLR) due to their high accuracy and versatility across a variety of input formats. Deep neural networks, including Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Gated Recurrent Units (GRU), have been used in numerous studies to efficiently recognize sign language from visual and sensor data. These models have shown their capacity to process and categorize a broad range of indications when applied to multimedia inputs such as sequential movies, skeletal data, thermal scans, and static photographs. CNNs are the most widely used architecture among them because of their exceptional feature extraction capabilities, particularly when dealing with huge datasets and environmental fluctuations.

Indian Sign Language (ISL) has been the topic of particular research in recent years. A CNN-based system was shown in one study. It was trained on a customized dataset of 35,000 grayscale photos of 100 static ISL signs in a variety of environmental settings. When Stochastic Gradient Descent (SGD) was used for optimization, the architecture obtained exceptional training and validation accuracies of up to 99.90% and 98.70%, respectively. Using the Bag of Visual Words (BoVW) model in conjunction with SURF feature extraction and segmentation based on background subtraction and skin color, as well as SVM and CNN classifiers, another method achieved 99% recognition accuracy. Additionally, these systems included user-friendly graphical interfaces and real-time feedback via text and audio outputs.By improving real-time sign language translation, making it more resilient to rotation and background changes, and expanding recognition capabilities to include entire words and expressions, these efforts aim to make sign language recognition more widely available and useful for daily use.

III. METHODOLOGY

3.1 Introduction

Creating an SURF with SVM and CNN system for the recognition of Indian Sign Language. To create a highly accurate system that would benefit real-time users, sign language recognition needs reliable and efficient data. Here, the sign recognition and classification challenge was resolved using the specially created dataset. For sign language recognition, the data flows through several steps, including dataset, image acquisition, pre-processing, feature extraction, and sign classification. Figure 3.1 displays the design flow diagram.

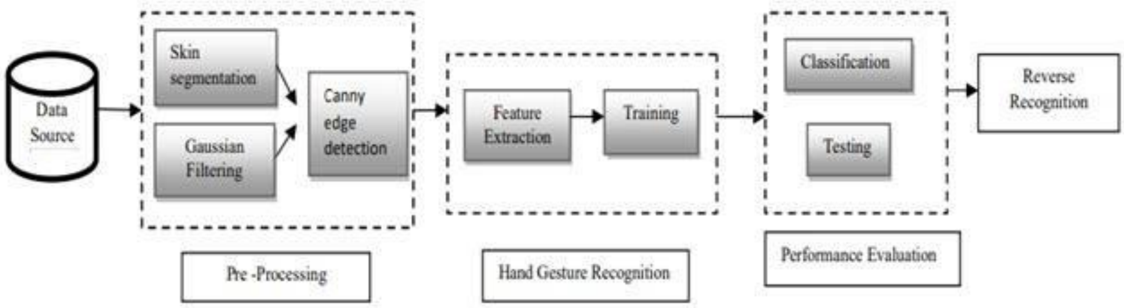


Figure 1. Flow Diagram

1.Data Collection: Gathering a dataset of Indian Sign Language motions is the initial step. Images or videos of sign language movements made by various people should be included in the collection. These should include a variety of lighting, background, and hand orientation variations.

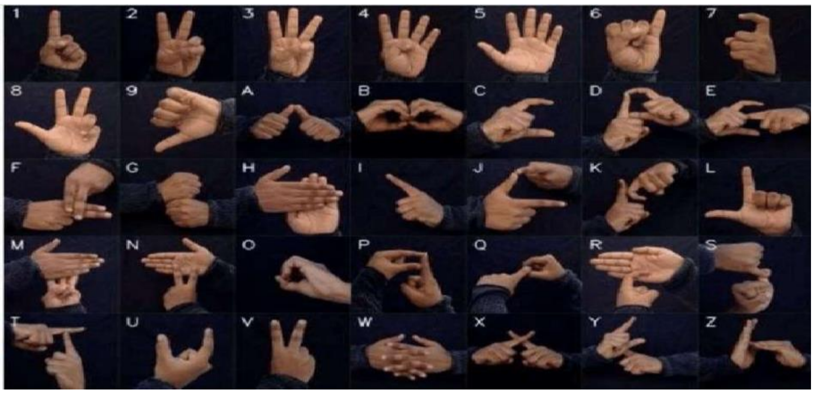


Figure 2. Dataset

2.Preprocessing: To eliminate noise or artifacts from the photos, the gathered dataset is pre-processed in this stage. To separate the hand region from the backdrop, this may entail methods like segmentation, thresholding, and picture filtering.

RGB Image

Binary Image

Mask

Canny Edge Detection

Figure 3. Image pre-processing steps.

3.Feature Extraction: To capture the most discriminative aspects of the sign language gestures, feature extraction is carried out after the photos have undergone pre-processing. The Speeded-Up Robust Feature (SURF) algorithm is a well-liked feature extraction technique that identifies and characterizes visual interest points.

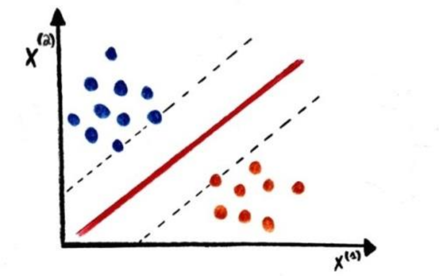
Image Input

Extracted Features

Figure 4. SURF Feature Extraction

4.Classification: Following their extraction, the pertinent characteristics are sent into CNN and SVM, two classification algorithms.

SVM: The Support Vector Machine (SVM) is a supervised model that may resolve regression and classification issues in both linear and non-linear fashions. It functions using the concept of decision planes, which define the parameters for choices. We utilized SVM with a linear kernel for this classification. For the purpose of classifying and identifying ISL signs, we have fed the SVM the visual word histograms as feature vectors. A total of 28,800 photos are used during the training process. Following training, the classifier's performance is assessed using a testing set consisting of 7236 images. The performance is measured using a number of criteria, including accuracy, precision, recall, and so on.



SVM Model

CNN: Modeled after the human visual cortex, CNNs uncover patterns at particular spots by swiping filter maps across local patches to extract picture information. They perform better on image classification tasks than conventional neural networks. The suggested CNN design uses three convolutional blocks to handle 100x100 input images: two 3x3 convolutional layers and 32 filters in the first block, two 64-filter layers in the second, and two more 64-filter layers in the third. A max pooling and dropout layer comes after each block to lower dimensionality and avoid overfitting. Lastly, the network has a SoftMax output layer with 36 neurons that each represent an ISL sign category and a dense layer with 512 ReLU-activated neurons.To capitalize on their complementing advantages, SVM and CNN can be employed alone for classification or in combination in an ensemble architecture.

IV. RESULTS AND DISCUSSION

Two separate sets of the dataset were created, with 80% of the data going toward training and the remaining 20% going toward testing. Notably, the images demonstrated remarkable accuracy for both the SVM and CNN classifiers. Nevertheless, the CNN fared better than the SVM, even though it used a smaller feature set. In particular, the system was made to identify 36 signs, which included 10 digits and 26 alphabets. much if the results are encouraging now, there is a chance that they may be much better with a few improvements.

SVM Performance

SVM's accuracy on the test data was 99.14%. SVM-classified alphabets and digits have an overall accuracy of 99% according to the computed precision and recall values.

CNN's Performance

We found that CNN achieved an overall accuracy of 94% on the training set during the most recent epoch, while testing accuracy exceeded 99%. There are 50 epochs in total.

Table 1
Class wise accuracy table.

Label	SVM (%)	CNN (%)	Label	SVM (%)	CNN (%)	Label	SVM (%)	CNN (%)
0	100	100	**C**	100	100	**O**	99	99
1	99	100	**D**	100	100	**P**	100	100
2	98	100	**E**	96	97	**Q**	100	100
3	96	98	**F**	95	100	**R**	98	98
4	100	100	**G**	98	100	**S**	100	100
5	100	99	**H**	100	100	**T**	99	100
6	100	100	**I**	98	100	**U**	99	100
7	100	100	**J**	100	100	**V**	100	100
8	98	100	**K**	100	100	**W**	99	100
9	98	100	**L**	99	100	**X**	100	99
A	100	100	**M**	100	99	**Y**	99	100
B	100	100	**N**	99	100	**Z**	100	100

Table 2
Accuracy table.

SVM	CNN
99.17%	99.64%

Table 3
Performance metrics table.

Measure	SVM	CNN
Precision	99.09	99.57
Recall	99.02	99.57
F1 Score	99.09	99.57

User Interface

As seen in Figures 5 and 6, the system was developed with an intuitive GUI to improve user engagement. Tkinter is used to effortlessly integrate a full sign-in and sign-up system into the GUI. "Predict Sign" is a button that allows users to easily anticipate signs based on the model learned using our dataset. Additionally, the "Create Signs" button offers a fascinating feature that lets users build their own sign database. Additionally, there is a speech-to-sign conversion option in the GUI. Screenshots of real-time video testing are shown in Figure 7. For further adaptability, users are given two options for collecting sign input images: one with a simple background and one without.

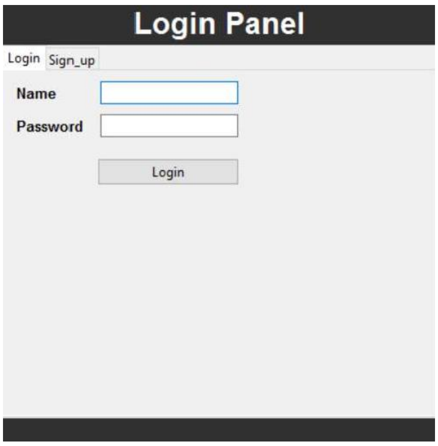


Figure 5. System GUI

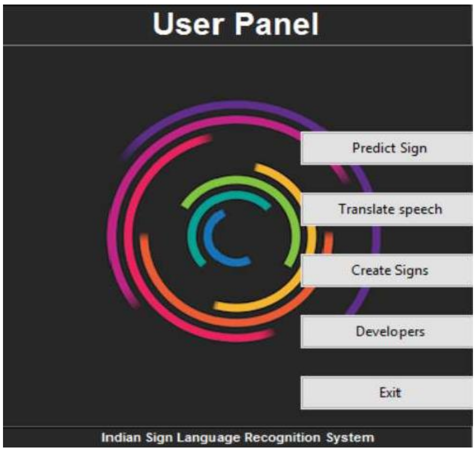


Figure 6. User Panel

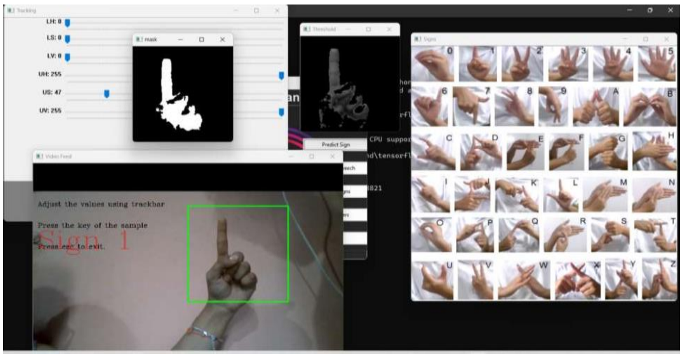


Figure 7. System Testing prediction of sign I

V. CONCLUSION

The development of a real-time recognition utility that is applicable in a variety of settings is the main goal of our work. By creating a unique dataset that tackles issues like background dependency and rotation invariance, this goal is fully achieved. Remarkably, the system successfully learns all 36 ISL static alphabets and numbers, achieving a remarkable 99% accuracy rate.

A more complete framework for real-time applications can be achieved in future undertakings by adding more indicators from various languages spoken in various nations to the dataset. The suggested approach can also be expanded to recognize basic words and expressions, meeting the needs of both isolated and continuous recognition tasks.Improving the system's response time is essential to enabling true real-time applications.

VI. REFERENCES

[1] Shagun Katoch, Varsha Singh, Uma Shanker Tiwary - Indian Sign Language recognition system using SURF with SVM and CNN- 29 March 2022.

[2] Shravani K, Sree Lakshmi A, Sri Geethika M, Dr. Sapna B Kulkarni- Indian Sign Language Character Recognition- 08-05-2020.

[3] J. L. Rahejaa, A. Mishrab, and A. Chaudharyc- Indian Sign Language Recognition Using SVM -September 1, 2015.

[4] Manjushree K, Divyashree. International Research Journal of Engineering and Technology- Gesture recognition for Indian sign language using HOG and SVM 2019.

[5] Kanade Sudhir S, Deshpande Padmanabh D. International Journal of Scientific Research and Development - Recognition of Indian sign language using SVM classifier -2018.

[6] Nair Anuja V, Bindu V. Int J Comput Appl - A review on Indian sign language recognition- 2013.

[7] Bantupalli Kshitij, Xie Ying. Master of Science in Computer Science -American sign language recognition using machine learning and computer vision- 2019.

[8] Bhattacharya Abhiruchi, Zope Vidya, Kumbhar Kasturi, Borwankar Padmaja, Mendes.