Sign Language Translator

[1]Aman Rasheed, [2]Farsana K A, [3]Mehna Zainab Ikbal, [4]Nila Rajesh, [5]Muhammed Shefeek

[1,2,3,4]B. Tech Scholar, [5]Assistant Professor
Department of computer science and engineering,
Ilahia College of engineering and technology
Muvattupuzha, India

***Abstract*—Both the deaf and the dumb use sign language frequently. The conversations have been more convenient for them. However, sign language may seem strange to those who have not studied it, or sign movements may be mixed up and confused with those who speak a different language. As a result, the goal of this initiative is to serve as a link between sign language users and those who doesn't use sign language. The technology offers two options: 1) to learn sign language, and 2) to translate sign language in real-time. The user can record live video when the system is acting as a translator. The text for the sign will then be shown after the system has recognized the sign gesture. The user can mimic the sign displayed by the system for a certain sign when the system is used to learn. And the system will verify whether what the user has shown is accurate or not. The system is trained, and the photos are classified, using a database of Indian sign language. The area of interest, in this case, the hand, goes through several processes, such as using the OpenCV (module of Python) or CNN (Convolutional Neural Network) processes to recognize the sign.**

***Index Terms*—Sign language recognition, Gesture recognition, ASL, Hearing disability.**

I. INTRODUCTION

Language is a fundamental aspect of human interaction since ancient times. It serves as a tool for people to express themselves and convey their thoughts about the world. Unfortunately, the rapidly changing society tends to overlook the needs of deaf individuals who face challenges in expressing their ideas and opinions to those who do not understand sign language. The use of sign language is not comprehensible to non-sign language users, creating a wider communication gap. However, the development of sign language recognition systems, a form of non-cognitive computing user interface, has the potential to address this issue. This technology allows computers to record and interpret human gestures as commands. While still under research in the field of computer vision, it has become essential for people with hearing disabilities who communicate through hand movements. Sign language recognition systems can provide an effective means of communication for the hard of hearing and a precise method of interpretation for those unfamiliar with sign language. It is important to note that sign motions vary across different nations, and standardization is crucial for accurate interpretation.

II. RELATED WORKS

The study described in reference [1] focuses on utilizing neural networks to detect hand gestures and recognize commonly used sign language. The aim is to develop a system that can identify and interpret sign language accurately. The study encompasses three main components:

1) The establishment of a hand locating network based on faster R-CNN that recognizes sign language video or the hand portion in an image. Subsequent processing receives the output of the recognition as input.

2) The article proposes a framework for sign language recognition that comprises two neural networks: a 3D CNN feature extraction network and an encoding/decoding network based on LSTM. The framework is designed to enhance the accuracy of sign language recognition by taking into account the context of image sequences.

3) In order to tackle the challenge of recognizing sign language images or videos in practical settings, the article suggests a unified recognition algorithm that merges three neural networks: a network for detecting hand gestures, a 3D CNN feature extraction network, and an encoding/decoding network based on LSTM. The proposed algorithm aims to achieve better sign language recognition by combining these networks.

According to the experimental outcomes, this approach has been able to attain a recognition accuracy of up to 99% on the widely-used vocabulary dataset. The achieved performance has outperformed other methods currently available for sign language recognition.

Reference [2] introduces a system that automates the recognition of Indian sign language (ISL) gestures using a standard camera to capture the gestures. The system is developed to identify individual ISL gestures, where each input image comprises only one character sign. To test the system's ability to recognize images in real-world environments, two datasets were created, one with 2600 images of single-handed characters and the other with 2340 images of double-handed characters (A-Z). The recognition system described in the article uses several features extracted from the gestures as input. These features include structural features, local histogram features, and direct pixel values. The system achieved recognition rates of 95.30% for single-handed gestures and 96.37% for double-handed gestures using kNN and neural network classifiers.

In [3] presents a recognition system that uses the Microsoft Kinect, GPU acceleration, and convolutional neural networks (CNNs) to identify Italian gestures. In contrast to conventional approaches that require manual construction of intricate features, convolutional neural networks (CNNs) automate the feature engineering process. The system achieves high accuracy in recognizing 20 Italian

gestures and can generalize to different users and environments with a cross-validation accuracy of 91.7%. The model's performance is also noteworthy in the ChaLearn 2014 Looking at People gesture spotting competition, where it attained a mean Jaccard Index of 0.789.

In [4] the authors propose a novel approach for video recognition that utilizes a 3D convolutional neural network used to automatically extract spatial-temporal features from unprocessed video streams, without the need for any manual feature design. The proposed approach employs multiple channels of video streams, such as color, depth, and body joint positions, to enhance the performance. These channels are fed as inputs to the 3D CNN to integrate depth and trajectory information and improve the feature extraction process. The model is evaluated on a dataset collected using Microsoft Kinect and is found to outperform traditional methods that rely on handcrafted features.

In [5] propose new methods to create a hand gesture recognition system that is independent of the user, taking into account the challenges that come with such a system. The proposed method involves evaluating the performance of a pre-trained AlexNet model for gesture recognition. The fully connected layers (FC6 and FC7) of the model are used to extract deep features, and a support vector machine (SVM) classifier with a linear kernel is employed for pose classification. To improve recognition accuracy, principal component analysis (PCA) is applied to reduce the feature dimension. The effectiveness of the proposed technique is assessed on a benchmark dataset of 36 static hand gestures in American Sign Language (ASL) using leave-one-subject-out cross-validation and holdout cross-validation tests. The results demonstrate that the proposed approach outperforms existing state-of-the-art techniques.

In [6] a cost-effective approach is presented for developing a recognition system for Sinhala sign language using image processing. The study tackles three main challenges in gesture-based human-computer interaction:
(1) Recording the gestures, (2) filtering out environmental noise from the gestures, and (3) mapping Sinhala sign language gestures to the corresponding database entries. The proposed technique employs efficient skin color filtering techniques and centroid finding methods to successfully map the gestures to the corresponding database entries, regardless of the hand's size and position. The system's accuracy is evaluated through a real-time simulation setup, and the research findings can address the need for simple and efficient applications that convert dynamic gestures made by the signer to a predefined word or phrase.

Reference [7] presents an Indian sign language converter that utilizes a Convolutional Neural Network (CNN) algorithm to identify and convert real-time images of Indian Sign Language gestures into their corresponding alphabet letters. A diverse background database was created, and various image pre-processing techniques were applied to prepare the database for feature extraction. Python software was used to feed the images into the CNN, and both offline and real-time images were tested for efficiency and accuracy. The results indicate that the testing images achieved an accuracy rate of 96%, while the real-time images attained an accuracy rate of 87.69%.

III. Proposed System

The gathering of data is a crucial component of this study because it has a significant impact on the findings. We use a webcam to record photos of hand movements that go through several processing steps in order to collect data for our system. In order to apply a mask to the photos, morphological approaches are used along with segmentation to identify the area of the skin tone in the photographs. The images are then subjected to a succession of elliptical kernel-based dilation and erosion processes.

Camera

To take pictures or create videos, an optical tool known as a camera is used, and here the camera is used to do the same i.e., to record the sign language being used.

Video

Camera records the video which contain the gesture of the sign language which has to be recognized. Hence the video must be of definition and quality.

Data Acquisition

The process of finding pertinent business data, arranging it into the required format, and transferring it into the designated system is called data loading. Effective machine learning algorithms cannot function accurately without high-quality data and proper data cleansing.

Data Preprocessing

Data preprocessing is a technique in data mining that aims to convert raw data into a comprehensible and useful format. In many cases, real-world data may be incomplete, inconsistent, or contain errors, making it difficult to analyze accurately. By performing data preprocessing, we can improve the reliability and accuracy of our analysis.

Feature Extraction

The process of transforming raw, unprocessed data into a set of numerical features. It can be used for analysis while preserving important information from the original data feature structure. Typically, employing this approach yields superior outcomes compared to applying machine learning algorithms directly to unprocessed data.

Classify Data using Trained Models

Classifying data using trained models involves using a machine learning model that has been previously trained on a labeled data set to predict the labels or classes of new, unseen data. The process of training a model involves showing it examples of labeled data and adjusting its parameters or weights to accurately predict the correct labels for each example.

Recognition

The purpose of recognition is to accurately identify and is to classify characters and gestures performed by users, with the ultimate goal of providing effective and efficient means of communication to individuals who are deaf or hard of hearing.

Output

At the end the sign gestures will be identified and displayed in the text format for the user to read.

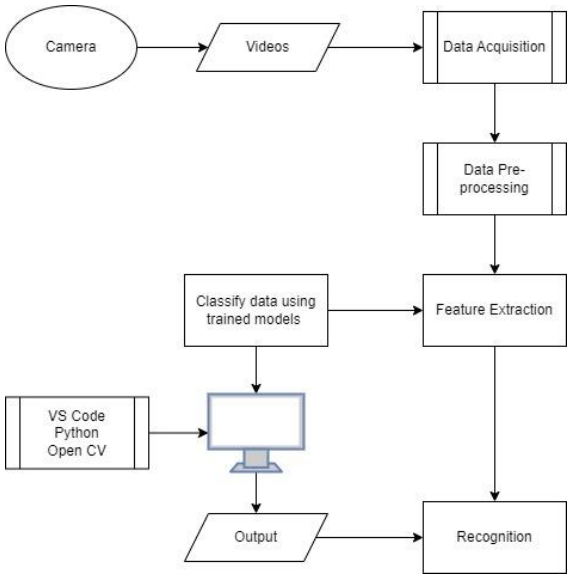

Figure 1 The architecture

IV. MODULES

The main modules in the proposed system are:

Admin

The admin is responsible for managing user accounts, creating and editing pages, adding and deleting content, managing plugins and themes, and performing other administrative tasks necessary for the proper functioning of the website. Here admin deals with training the system. It involves teaching the system different sign language gestures.

User

User can use the system for learning sign language and also for sign language translation in real-time. The user requires a Device with internet facility as well as a camera foe taking live videos. User has place their hand before the camera so the system can recognize the sign gesture.

Data-set Training

Training a machine learning model involves using a data set to teach the system how to recognize and predict specific inputs, such as sign language gestures. The training data set consists of various images of these gestures, which are used to teach the system how to identify and interpret the patterns and relationships in the data. During the training process, the machine learning algorithm modifies the parameters and weights of the model to reduce the difference between the predicted outputs and actual outputs for each training example. This process enables the system to accurately predict the input based on the learned patterns in the data set.

Live Video Processing

Live video processing refers to the real-time manipulation of a video stream as it is being captured and transmitted. This involves using specialized software and hardware to process the video frames in real-time and apply various effects or modifications to the video stream. Here the sign gestures will be translated in real-time while processing the live video and the result will be shown as a text.

Output

In general, an output is the result of a process or computation. And here the end result, language will be translated to text.

V. RESULT AND DISCUSSION

The system identify the sign gesture present in the live video using CNN (Convolutional Neural Network).In following Figures the system was tested by placing a hand in front of the camera, and the system promptly began recognizing the sign gesture, the gesture we gave the system to recognize was the character 'B'.

Figure 2 Sign recognition using CNN.

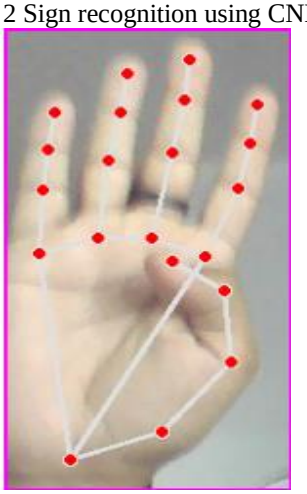

Figure 3 The result after recognition.

VI. Conclusion

Our project aims to facilitate communication between sign and non-sign language users using a convolutional neural network (CNN) and a training data set. By accurately translating sign language into text or audio formats, the system provides a practical solution to the communication barriers faced by deaf and taciturn people.

Furthermore, the system can be expanded to include the translation of different sign languages used in different countries. With additional training, it could also be utilized to identify objects and food names in live videos or recognize various types of signboards, such as road signs or construction signs.

VII. Acknowledgment

The success of this bulletin is due not only to our efforts, but also to the guidance and support of many people. We would like to extend our appreciation to those who have played a role in the success of this project. First of all thank you prof. Dr.K. A Navas, principal of Ilahia college of engineering and technology, for letting us work on this paper. We would also like to thank Dr. Lino Abraham Varghese, Head of Computer Science and Engineering, Project Leader Muhammed Shefeek, and Project Coordinators Chithra Rani P R and Shanavas K A for their valuable advice and valuable advice guidance. And lastly we would like to express our gratitude to all the staff members and faculty of the Computer Science and Engineering Department for their wholehearted support, which was instrumental in the successful completion of this project.

REFERENCES:

1. He, Siming. (2019). Research of a Sign Language Translation System Based on Deep Learning. 392-396. 10.1109/AIAM48774.2019.00083.
2. M. Geetha and U. C. Manjusha, , A Vision Based Recognition of Indian Sign Language Alphabets and Numerals Using B-Spline Approximation, Inter- national Journal on Computer Science and Engineering (IJCSE), vol. 4, no. 3, pp. 406-415. 2012.
3. Lionel Pigou, “Sign language recognition using convolutional neural networks”.
4. Jie Huang, Wengang Zhou , Houqiang Li and Weiping Lial, Huang (2015). Sign Language Recognition using 3D convolutional neural networks. IEEE International Conference on Multimedia and Expo (ICME) (pp. 1-6). Turin: IEEE.
5. Jaya Prakash Sahoo, Samit Ari, Sarat Kumar Patra, "Hand gesture recognition using PCA based deep CNN reduced features and SVM classifier”.

6. Herath, H.C.M. & W.A.L.V.Kumari, & Senevirathne, W.A.P.B & Dissanayake, Maheshi. (2013) “Image based sign language recognition system for sinhala sign language”.
7. Nishi Intwala, Arkav Banerjee, Meenakshi and Nikhil Gala, “Indian sign language converter using convolutional neural networks”.
8. Kshitij Bantupalli, Ying Xie, "American sign language recognition using deep learning and computer vision”,2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 2018, pp. 4896-4899, doi: 10.1109/BigData.2018.8622141.
9. N. K. Bhagat, Y. Vishnusai, and G. Rathna, “Indian sign language gesture recognition using image processing and deep learning,” in 2019 Digital Image Computing: Techniques and Applications (DICTA), pp. 1–8, IEEE, 2019.
10. S. Davey, C. Maheshwari, S. K. Raghav, N. Singh, K. Muzammil, and P. Pandey, “Impact of indian public health standards for rural healthcare facilities on national programme for control of deafness in india: The results of a cohort study,” Journal of Family Medicine and Primary Care, vol. 7, no. 4, p. 780, 2018.
11. N. Intwala, A. Banerjee, N. Gala, et al., “Indian sign language converter using convolutional neural networks,” in 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), pp. 1–5, IEEE, 2019.
12. J. P. Sahoo, S. Ari, and S. K. Patra, “Hand gesture recognition using pca based deep cnn reduced features and svm classifier,” in 2019 IEEE International Symposium on Smart Electronic Systems (iSES)(Formerly iNiS), pp. 221–224, IEEE, 2019.
13. A. Sarkar, A. K. Talukdar, and K. K. Sarma, “Cnn-based real-time indian sign language recognition system,” in International Conference on Advances in Computational Intelligence and Informatics, pp. 71–79, Springer, 2019.
14. C. Sruthi and A. Lijiya, “Signet: A deep learning based indian sign language recognition system,” in 2019 International Conference on Communication and Signal Processing (ICCSP), pp. 0596–0600, IEEE, 2019.
15. A. Russo, A. Filonenko, and K.-H. Jo, “Sports classification in sequential frames using cnn and rnn,” pp. 1–3, 09 2018.
16. S. Ruder, “An overview of gradient descent optimization algorithms,”arXiv preprint arXiv:1609.04747, 2016.
17. Mehreen Hurroo,Mohammad Elham,(2020) "Sign Language Recognition System using Convolutional Neural Network and Computer Vision", IJERT - IJERTV9IS120029, licensed under a Creative Commons Attribution 4.0 International License.