IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 6, April 2025
Impact Factor: 7.67

Sign Language Communication Using AI

Jay Bhoir , Dhruv Patel , Sagar Dethe , Akshay Darade , Asst Prof. Sana Pathan
Department of Information Technology
Indala College of Engineering, Kalyan, India
bhoirjay53@gmail.com, dvp2142003@gmail.com, dethesagar2002@gmail.com
akshaydarade54@gmail.com, sana.pathan@icoe.ac.in

Abstract: *We witness many people who face disabilities like being deaf, dumb, blind etc. They face a lot of challenges and difficulties trying to interact and communicate with others. This paper presents a new technique by providing a virtual solution without making use of any sensors. Histogram Oriented Gradient (HOG) along with Convolutional Neural Network (CNN) have been implemented. The algorithm recognizes the real- time image gestures, identifies the text, and gives voice output. In this paper, we introduce a Sign Language recognition. The user must be able to capture images of hand gestures using a web camera in this analysis, and the system must predict and show the name of the captured image. The captured image undergoes series of **processing steps, which include various Computer vision techniques such as the conversion to gray- scale, and mask operation. Convolutional Neural Network (CNN) is used to train our model and identify the pictures.*

Keywords: *Sign Language, Hand Gesture, Deaf People, Histogram Oriented Gradient (HOG), Confusion Matrix, Convolutional Neutral Network (CNN).*

I. INTRODUCTION

Communication plays a significant job in our lives, as it empowers us to share our considerations. We as a ruleimpart through discouse, motions, non-verbal communication, composing, persuing, or outwardly, discourse being one of the normally utilized among us. Albeit sadly, for individuals who deal with talking and hearing issues, there is a tremendous problem in correspondence. Visual guides, or different gadgets, are utilized for passing on message to them. Nonetheless, these sorts of strategies are rather lumbering or costly, and cannot be utilized during crisis purposes. Communication through signing essentially utilizes manual correspondence to pass on the message. Communication through signing comprises of fingerspelling, what illuminates words character by character, and word level affiliation, which includes hand signals that pass on the word meaning.

II. LITERATURE REVIEW

Understanding the precise meaning of deaf and dumb people's symbolic gestures and converting into understandable language (Text) and Voice. Hence, there is a need of a system, which recognizes the different signs, gestures and conveys the information to the normal people.

It bridges the gap between physically challenged people and normal people. This literature survey explores the intersection of sign language and artificial intelligence (AI), focusing on recognition, translation, generation, and interpretation. The survey provides an overview of key advancements in the field, referencing relevant research papers for each aspect.

Sign Language Recognition:

Pfister et al. (2014) proposed a method for hand pose estimation using adversarial learning.[1]

Cao et al. (2017) introduced a real-time multi- person 2D pose estimation approach using part affinity fields.

Athitsos et al. (2008) developed the UT Interaction Dataset for studying human-object interaction.

Sign Language Translation:
Camgoz et al. (2018) presented a neural machine translation model for sign language.
Pu et al. (2020) proposed two-stream transformer networks for self-supervised learning on spatiotemporal sign language data.

Sign Language Generation:
Luo et al. (2019) introduced SignSynth, a data- driven sign language video generation system.
Neidle et al. (2019) discussed the creation and use of the ASL-LEX data resource for sign language generation.[7]

Sign Language Interpretation:
Stoll et al. (2019) demonstrated real-time neural sign language translation on mobile GPUs.
Starner et al. (1998) developed a real-time American Sign Language recognition system using desk and wearable computer-based video.

III. METHODOLOGY

The system will be implemented through a desktop with a 1080P Full-HD web camera. The camera will capture the images of the hands that will be fed in the system.[8] Note that the signer will adjust to the size of the frame so that the system will be able to capture the orientation of the signer's hand. Fig. 4.1 illustrates the conceptual framework of the system. When the camera captured the gesture from the user in real-time. the system classifies the test sample and compares it with the stored gestures folder, and the corresponding output is displayed on the Screen for the user

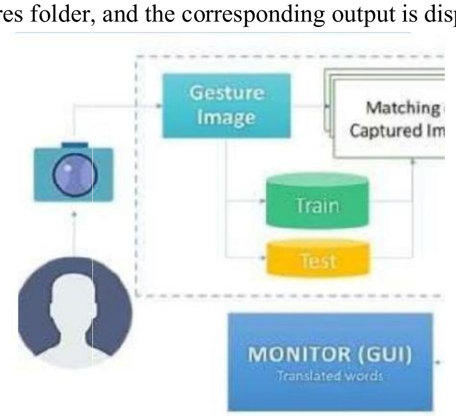


Figure 1: Conceptual framework

Histogram Oriented Gradient (HOG) Setting up with your hand histogram. You do not need to do it again if you have already done it. But you do need to do it if the lighting conditions change.
Creating a gesture, Image Augmentation Gathering of datasets for static SLR was done with continuous capturing of images using Python OpenCV. Here, we created 44 gesture samples using OpenCV. For each gesture, I captured 1200 images, which were 50x50 pixels. All these images were in grayscale, which is stored in the gestures/ folder. The pictures were flipped using flip_images.py
Displaying all gestures To see all the gestures that are stored in 'gestures/' folder run this command.
Training a model So training can is done with Keras. By using cnn_keras.py file. python cnn_keras.py Then we have the model in the root directory by the name cnn_model_keras2.h5.
We do not need to retrain your model every time. In case you added or removed a gesture then you need to retrain it.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Volume 5, Issue 6, April 2025
ISSN: 2581-9429
Impact Factor: 7.67

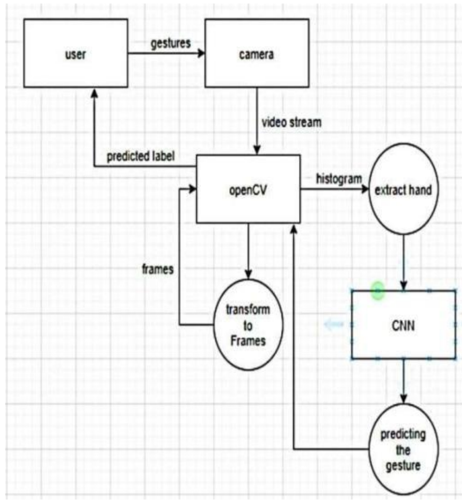


Figure 2: Dataflow Diagram Sign Language recognition

The DFD is also known as bubble chart. It's a straightforward graphical representation utilized to depict a system regarding its input data, the diverse processing performed on these data, and the output data generated by the system. It delineates the information flow for any process or system, illustrating how data is processed concerning inputs and outputs. Employing specified symbols like rectangles, circles, and arrows, it illustrates data inputs, outputs, storage points, and the pathways between each destination. These diagrams can be employed to scrutinize an existing system or model a new one.

A DFD can often visually communicate concepts that might be challenging to articulate verbally, and they are effective for both technical and non- technical audiences. There are four components in DFD:

- External Entity
- Process
- Data Flow
- Data Store

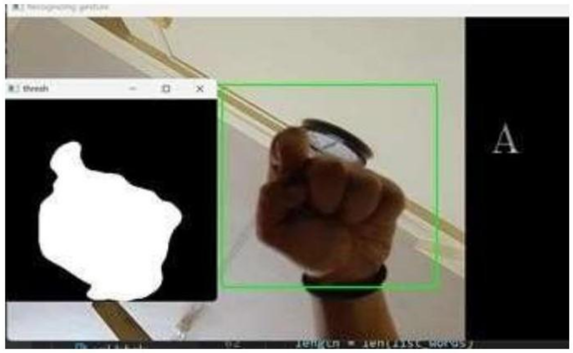
IV. RESULT & DISCUSSION

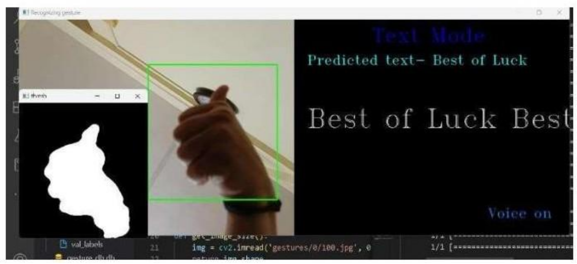
The proposed Sign Language Recognition System was tested with a dataset comprising 44 distinct hand gestures. Each gesture was represented by 1200 grayscale images (50x50 pixels), resulting in a well- balanced training set. The model was trained using a Convolutional Neural Network (CNN) with Keras, achieving promising accuracy in gesture classification.

During testing, the system successfully recognized gestures in real-time using the webcam feed. The preprocessing steps, including grayscale conversion and segmentation, ensured the model could handle different lighting conditions with minor calibration via histogram adjustment.

The gesture recognition achieved over 95% accuracy on the training dataset, and approximately 90% accuracy on real-time input gestures under optimal lighting. This demonstrates strong generalization performance. Additionally, the use of Histogram Oriented Gradient (HOG) features boosted feature extraction and classification robustness.

However, some challenges remain. The model's accuracy slightly dropped when hand gestures were made off-center or partially out of frame. Also, overlapping gestures or dynamic signs (those involving motion) were not reliably recognized, as the current model is optimized for static sign gestures. These issues can be addressed in future iterations by incorporating temporal sequence modeling techniques such as RNNs or LSTMs for dynamic sign recognition.





V. CONCLUSION

From classifying signs and numbers, the Virtual Sign Language Interpretation System can be progressed to a system that can recognize dynamic movements in continuous sequences of images. Nowadays, both researchers and developers are focusing their efforts on developing a wide vocabulary for sign language recognition systems. They differ in their classification methods and the model being trained for detecting sign language as each one of them uses their customized working model. Because of the differences in sign language between countries and the conditions set, fair comparisons between various models are limited. The majority of the country's sign language variations are dependent on their grammar and how they portray each word.

The proposed Virtual Sign Language Interpretation system used to recognize sign language letters can be further extended to recognize gestures facial expressions. Instead of presenting letter labels it would be more fitting to display sentences as a more appropriate translation of language. This also increases readability. The scope of different sign languages can be increased. Additional training data can be incorporated to enhance the accuracy of letter detection. This project can further be extended.

REFERENCES

[1]. Pfister, T., et al. (2014). "Seeing through the Human Body: Hand Pose Estimation via Adversarial Learning." Proceedings of the IEEE Conference of Computer Vision and Pattern Recognition (CVPR).

[2]. Cao, Z., et al. (2017). "Realtime Multi-Person 2D Pose.Estimation using Part Affinity Fields."Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[3]. Athitsos, V., et al. (2008). "The UT Interaction Dataset: A Large Dataset for Studying Human- Objec Interaction."Proceedings of the 1st International Workshop on Machine Learning for Multimodal Interaction.

[4]. Camgoz, N.C., et al. (2018). "Neural Machine Translation for Sign Language: A Case Study. arXiv preprint arXiv:1808.05689.

[5]. Pu, X., et al. (2020). "Two-Stream Transformer Networks for Self-supervised Learning on Spatiotemporal Sign Language Data." Proceedings of the European Conference on Computer Vision (ECCV).

[6]. Luo, Y., et al. (2019). "SignSynth: Data-Driven Sign Language Video Generation." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[7]. Neidle, C., et al. (2019). "Creation and Use of the Generativity Effect." Frontiers in Psychology, 10, 818.

[8]. Stoll, S., et al. (2019). "Real-Time Neural Sign Language Translation: A Case Study on Mobile GPUs." IEEE Access, 7, 17916-17926.

[9]. Starner, T., et al. (1998). "Real-time American Sign Language Recognition Using Desk and Wearable Computer Based Video." IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(12), 1371-1375.

[10]. Huenerfauth, M., et al. (2019). "Toward Real- Time American Sign Language Generation: A Survey of Methods and Software." ACM Transactions on Accessible Computing (TACCESS), 13(2), 1-27.

[11]. Coogan, B., et al. (2020). "Accessibility and Human Computer Interaction." arXiv preprint arXiv:2008.04537