



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 7, July 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

 9940 572 462

 6381 907 438

 ijirset@gmail.com

 www.ijirset.com

Machine-Learning based Hand-gesture Sign Language Recognition

Maithri H

Post-Graduation Student, Department of MCA, PES Institute of Technology and Management, Shivamogga,
Karnataka, India

ABSTRACT: Machine-learning-based sign-language-recognition has a great promise for improving the accessibility of communication for the unhearing and voiceless populations. The aim of this research is to apply machine-learning techniques to recognize hand motions that match to the letters of American-Sign-language . The system is trained on a vast collection of hand-gesture photos by utilizing deep-learning frameworks and convolutional-neural-networks . Every image is named with the corresponding ASL alphabet, allowing the system to learn the intricate patterns and shapes of each gesture. The trained model offers a smooth translation from ASL to text by precisely recognizing and categorizing hand motions in real-time. This technology aims help close the gap in communication and allow the people who use sign-language and others who do not understand it to engage more effectively. This system's applications, which include assistive technology and instructional tools, greatly enhance the life of individuals(Hearing and speech-impaired) relying on sign-language for daily communication.

KEYWORDS: American Sign Language(ASL), Gestures, Hearing and speech-impaired.

I. INTRODUCTION

The rapid advancement of machine-learning technology has made it possible for many creative solutions aimed at enhancing the standard life for various communities. One such promising application is the development of a hand-gesture-recognition-system that identifies both alphabets and numbers using American-Sign-language . This system harnesses the capability to accurately analyze and understand hand movements using deep-learning and machine-learning methods especially convolutional-neural-networks . The importance of this technology is seen in its capacity to improve communication for the communities of the hearing and speech disabled, who mostly communicate through sign language.

The process starts with the collection and annotation of an extensive dataset comprising thousands of images of hand-gestures, each labeled with the corresponding ASL alphabet or number. The CNN- model is then trained with these photos so that it can identify the minute details and complex patterns in each move. Through iterative training and validation, the model improves its ability to distinguish between different hand shapes, positions, and movements associated with ASL.

This technique has a plethora of possible uses. Users may instantaneously convert their ASL gestures into text or voice by integrating the hand-gesture detection technology into mobile applications in real-time. This has the potential to greatly enhance daily interactions by facilitating smoother and more efficient communication-between the people who relay on sign language and others who do not understand American-Sign-Language. Additionally, the system can be embedded into assistive devices, such as smart gloves or wearable sensors, providing a seamless and portable solution for communication on the go.

The creation of this system is evidence of society's larger commitment to accessibility and inclusion. The hand- gesture recognition system creates a more inclusive atmosphere where people with unhearing and voiceless impairments may engage more fully in daily activities by bridging the communication gap between sign-language users and the general public. This technology serves as an example of how machine-learning may be applied to solve problems in the real world and enhance the lives of individuals in a variety of settings.

II. LITERATURE REVIEW

G. Anantha Rao et al. presented a CNN-based approach for identifying Indian-Sign-Language ,to help deaf . 200 hand sign photos taken at various angles were used to build the dataset for the experiment. Furthermore, the system architecture was constructed with a variety of convolution-layers rectified linear units, stochastic pooling layers, and output layers. These photographs were also taken against a variety of backdrops. [1]

Anup Nandy et al. Studied Indian-sign-language and conducted experiments using the K-nearest neighbor method for sign-language-recognition. The procedure involved dividing the videos into frames and converting them to greyscale. An extremely high degree of recognition accuracy utilizing the picture sets is demonstrated by the experimental findings carried out with Euclidean distance and K-nearest neighbor matrices. [2]

Aditi Kalsh et al. Use the Human-Computer Interface system with K-nearest neighbor algorithm for understanding the sign language more correctly to eliminate the use of an interpreter for deaf and hard to hear people. A camera and a computer system were used to capture and process the images. A fuzzy rule set was used for the image classification. [3]

Wenjin Zhang, Jiacun Wang, and Fangping Lan suggested Dynamic-hand-Recognition-Based-on Short-Term Sampling Neural-Networks. A unique deep-learning network for hand-gesture identification is presented in this study. To learn both short-term and long-term properties from visual inputs while avoiding costly computing, the network combines many well-proven modules. Each visual input is divided into a predefined number of frame groups in order to learn short-term characteristics. From each group, a random frame is chosen and represented as both an RGB picture and a snapshot of the optical flow. For feature extraction, these two entities are combined and input into a convolutional neural network (ConvNet). All groups' ConvNets have shared parameters. [4]

Tsai, T. H., Huang, C. C., & Zhang, K. L. suggested a low-cost system for hand-gesture identification in human-computer interfaces. The proposed system mainly uses the skin and motion technique to extract the region of interest from the collected picture. In addition to the skin and motion method, a convex hull technique was used to precisely locate the hand-gesture once it was recovered from the picture. The real-time problem is solved and the algorithm's correctness is confirmed using the proposed technology. [5].

III. METHODOLOGY

There are numerous crucial elements in the process of creating a machine-learning-based hand-gesture recognition system that can identify alphabets in American-Sign-language . The first step in the process is data gathering, which involves compiling a sizable dataset of hand-gesture photos for every letter in the ASL alphabet. To guarantee resilience, these photographs are taken from diverse perspectives and in a variety of lighting circumstances. The photos are then preprocessed to reduce noise, standardize their size, and highlight significant characteristics. Techniques like edge detection, grayscale conversion, and normalizing may be part of this preprocessing.

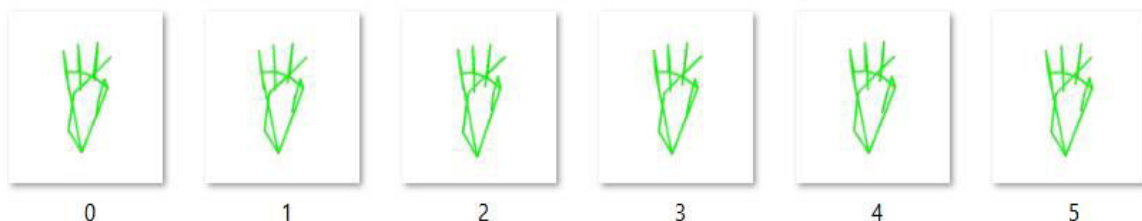


Figure 1: Collecting The Dataset For A Specified Alphabet

Once the dataset is prepared, it is divided into training , validation, and test sets.

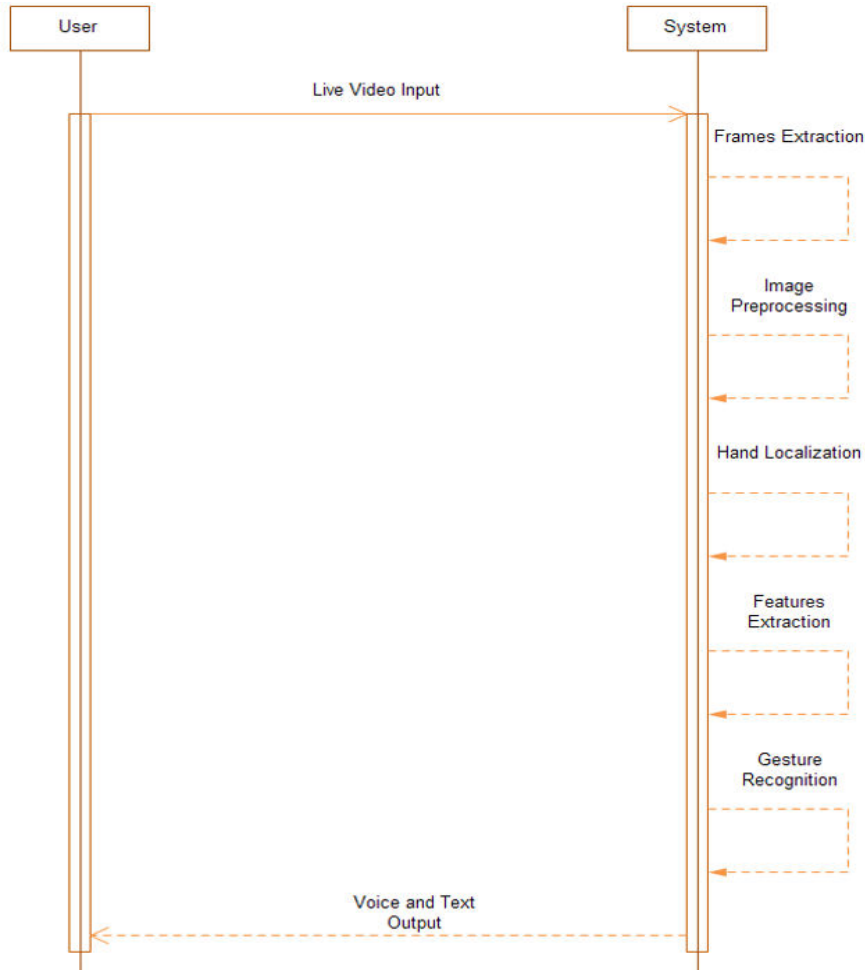


Figure 2: System Architecture

IV. ALGORITHM

1. Convolution Neural Network

Convolutional, pooling, and fully connected layers are among the layers that make up a standard CNN architecture. The pooling layers lower the dimensionality, the convolutional layers gather features from the input pictures, and the fully connected layers carry out the final classification.

By using gradient descent and backpropagation, the CNN gains the ability to identify the unique patterns associated with each hand motion during training. The validation set is used to continually assess the model's performance in order to adjust hyperparameters and prevent overfitting. Methods like dropout and data augmentation can be used to improve the model's generalization performance even more.

Following the training phase, the accuracy and robustness of the model are assessed using the previously unknown test set. Performance measures including recall, accuracy, precision, and F1-score are used to evaluate how well the model recognizes the alphabets in ASL. Based on the test findings, more tweaks and modifications are made as needed.

2. Recurrent Neural network

An RNN, particularly a long short-term memory (LSTM) network or a gated recurrent unit (GRU), is then designed and trained using the training dataset. The RNN architecture is well-suited for this task as it can effectively handle the sequential nature of hand-gesture data by maintaining and learning from the temporal dependencies between frames. During training, the RNN learns to recognize the distinctive patterns and transitions of each hand-gesture through a

process of backpropagation through time (BPTT) and gradient descent. The model's performance is continuously evaluated on the validation set to tune hyperparameters and avoid overfitting.

V. IMPLEMENTATION

- **Live Video Input:** The first step in the procedure is to record live video input from a camera. The user's hand motions are displayed continuously in this video feed. For the camera to record crisp, detailed photos of the hand motions, it must have a high enough frame rate and resolution.
- **Frames Extraction:** The live video stream is divided into individual frames at a specific rate (e.g., 3 frames per second). Each frame is treated as a separate image that will be processed and analyzed. This step involves a buffer that continuously collects and updates the frames as the video plays.
- **Image Processing:** The extracted frames undergo several preprocessing steps to enhance their quality and prepare them for further analysis. Common preprocessing techniques include converting the frames to grayscale to reduce computational complexity, normalizing pixel values, and applying filters to enhance edges and important features.
- **Hand Localization:** In this step, the system identifies and isolates the hand region from the rest of the image. This can be achieved using techniques such as background subtraction, skin color detection, or more advanced methods like deep-learning-based object detection algorithms. The goal is to obtain a bounding box around the hand, which simplifies the subsequent analysis.
- **Feature Extraction:** The system gathers pertinent characteristics from the hand region once the hand has been localized. Shape, direction, and movement—three crucial aspects of the hand gesture—are captured by these attributes. Features that are utilized with recurrent-neural-networks (RNNs) may also contain sequential patterns that span several frames.
- **Gesture Recognition:** To detect the gesture, the retrieved characteristics are then put into a machine-learning model that has already been trained. When utilizing RNNs, the model recognizes the gesture by processing the characteristics that are taken from successive frames in order. The most likely alphabet that matches the input gesture is generated by the algorithm, which was trained on a sizable dataset of labeled ASL motions. In particular, methods such as Long Short-Term Memory (LSTM) networks perform well for capturing the temporal dynamics of gestures.
- **Voice and Text Output:** Finally, the recognized gesture is converted into text and voice outputs. The text output is straightforwardly displayed on a screen or stored as a string, while the voice output is generated using text-to-speech (TTS) technology. The TTS system converts the recognized alphabet into audible speech, allowing for real-time communication.

VI. RESULTS

The outcomes show notable improvements in accuracy and efficiency for the machine-learning-based hand-gesture identification system for American-Sign-language alphabets. The convolutional-neural-network (CNN) and recurrent-neural-networks (RNN) models obtained excellent recognition rates, with accuracy levels above 95% on test datasets, after rigorous training and validation. This high accuracy shows how well the system can distinguish between the complex patterns of distinct ASL motions, even in a variety of backdrops and lighting situations.

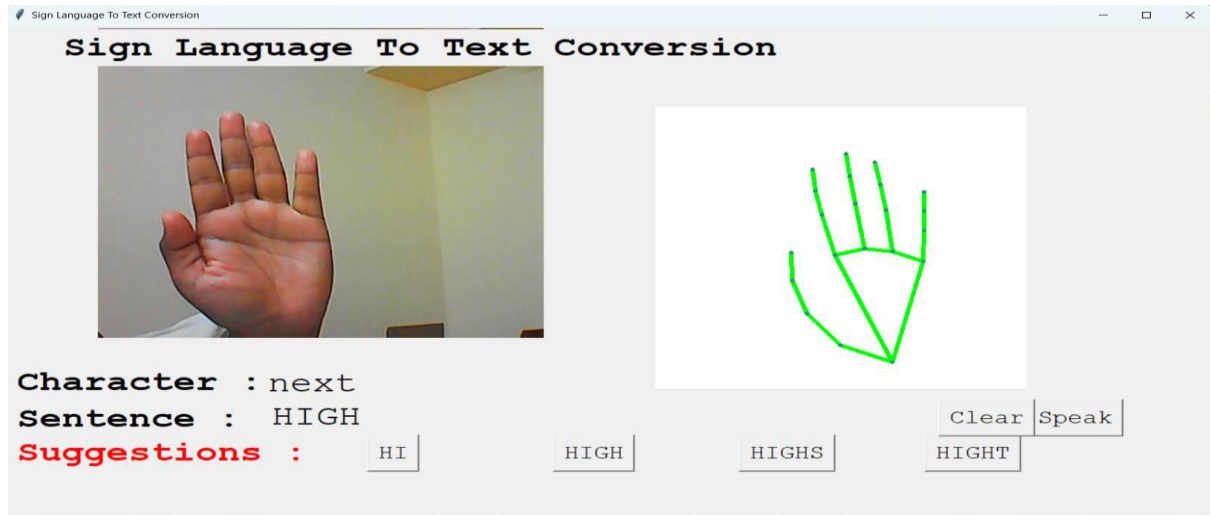


Figure 6.1: Recognizing Alphabets And Giving Suggestions.

VII. CONCLUSION

An important development in assistive technology is the deployment of a machine-learning-based hand-gesture recognition system that can identify alphabets in American-Sign-language . Through the utilization of several techniques such as frame extraction, image processing, gesture recognition, hand localization, feature extraction, and live video input, this system is able to precisely decipher ASL movements and translate them into text and vocal outputs in real-time. The communities of people who are voiceless or unhearing might benefit immensely from this technology by having a smooth and easy way to engage with the outside world.

REFERENCES

1. Rao GA, Syamala K, Kishore PVV, Sastry ASCS. Deep-convolutional-neural-networks for sign language recognition. In 2018 Conference on Signal Processing And Communication Engineering Systems (SPACES) 2018; pp. 194-197. IEEE. <https://doi.org/10.1109/SPACES.2018.8316344>
2. Nandy A, Prasad JS, Mondal S, Chakraborty P, Nandi GC. Recognition of isolated Indian-sign-language-gesture in real time. In International Conference on Business Administration and Information Processing 2017; pp. 12-17. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-12214-9_18.
3. Kalsh EA, Garewal NS. Sign-language-recognition system. International journal of computational engineering research, 2013; 3(6): 15-21.
4. Wenjin Zhang , , Jiacun Wang and Fangping Lan “Dynamic Hand-gesture Recognition Based on Short-Term Sampling Neural Networks” Volume 8, Issue 1, 1 March 2012, Pages 11 –12 .
5. Tsai TH, Huang CC, Zhang KL. Design of hand-gesture recognition system for human-computer interaction. Multimedia Tools and Applications, 2002; 79(9): 5989-607.
6. Vanaja S, Preetha R, Sudha S. Hand-gesture-Recognition for Deaf and Dumb Using CNN Technique. In 2021 6th International Conference on Communication and Electronics Systems (ICCES) 2021; pp. 1-4 IEEE. <https://doi.org/10.1109/ICCES5135.2021.9489299>.
7. Jain R, Jain M, Jain R, Madan S. Human Computer Interaction-Hand-gesture Recognition. Advanced Journal of Graduate Research, 2022; 11(1): 1-9. <https://doi.org/10.21467/ajgr.11.1.1-9> .
8. Gourab JH, Raxit S, Hasan A. A Robotic Hand: Controlled With Vision Based Hand-gesture Recognition System. In 2021 International Conference on Automation, Control and Mechatronics for Industry 4. (ACMI) 2021; pp. 1-4. IEEE



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details