

# Machine Learning driven Sign Language Recognition with Optimized Analysis and Multilingual Speech Generation

<sup>1</sup>Manpreet Kaur

*Assistant Professor*

*Lovely Professional University*

Jalandhar, Punjab, India

mann.bhullar9799@gmail.com

<sup>2</sup>Pourush Sharma

*Student*

*Lovely Professional University*

Jalandhar, Punjab, India

pourushs9928@gmail.com

<sup>3</sup>Sonu Bhardwaj

*Student*

*Lovely Professional University*

Jalandhar, Punjab, India

sbhardwaj1315@gmail.com

<sup>4</sup>Sachin Kumar

*Student*

*Lovely Professional University*

Jalandhar, Punjab, India

sachinjaiswalihcas@gmail.com

<sup>5</sup>Abhishek Pathak

*Student*

*Lovely Professional University*

Jalandhar, Punjab, India

abhishepathak9135@gmail.com

**Abstract**—Hand sign recognition has become a significant domain of study and development in the field of computer vision and artificial intelligence. Particularly, American Sign Language (ASL) is a complex perceptible language that encompasses hand gestures, facial expressions, and body language to convey information. However, a communication gap persists between the deaf community and the hearing world, necessitating for more efficient and intuitive modes of communication. One of the promising solutions is the creation of a system capable of accurately recognizing and translating hand gestures into corresponding ASL letters in real-time. In this paper, first image collection is done in a live video stream after that these images are pre-processed, and dataset is created from these images. Later, the dataset is divided into training and testing part on which multiple machine learning algorithms are applied and a model/classifier is created using Random Forest as it yielded highest accuracy of 99.9%. Finally, the model is used to recognize hand signs in live video stream from these predicted hand signs and a string is generated which later can be converted to speech of user preference.

**Keywords**—Random Forest, ASL, Hand Gesture Recognition, Text to Multilingual Speech.

## I. INTRODUCTION

Machine Learning driven Sign Language Recognition with Optimized Analysis and Multilingual Speech Generation is an evolving field of study focused on creating real-time system capable of accurately interpreting and translating hand gestures into American Sign Language (ASL) letters or words. ASL is a complex language in which hand gestures, facial expressions, and body language are the modes to convey information. It serves as a primary mode of communication for the deaf and hard of hearing community. Which has led to development of the sign language interpreters and text-to-

speech systems, however a accessibility limitations persists between these communities and the hearing world.

Hand gesture recognition and text to speech for ASL can be a more preceptive and efficient form of communication method for the deaf and hard hearing community. This system aims to bridge this gap by using computer vision techniques and deep learning algorithms to detect and recognize hand gestures in real-time, subsequently translating them into ASL letters or words. The challenge in this system includes

accurately detecting both static and dynamic gestures, managing occlusions, and processing information swiftly.

The project aims developing a robust hand gesture recognition system for ASL, this project endeavors to offer a more preceptive and efficient means of communication for individuals within the deaf and hard of hearing community. The successful implementation of this project could significantly impact on the lives of millions who rely on ASL as their primary form of communication.

## II. LITERATURE SURVEY

This paper [1] offers a comprehensive examination of various hand gesture recognition methodologies and their practical uses. The authors discuss various feature of extraction methods, such as color, texture, and shape-based approaches, and assessing their performance on the data sets. Additionally, the author discussed about tracking of Hand, Segmentation and its features extraction. Furthermore, the research also explores the efficacy of other machine learning techniques, such as decision trees, neural networks, and support vector machines, in hand gesture detection.

In this paper [2] Introduces an instantaneous hand gesture recognition system employing a convolutional neural network (CNN). The authors introduce a novel CNN architecture,

# Machine Learning driven Sign Language Recognition with Optimized Analysis and Multilingual Speech Generation

which integrates features extracted from both color and depth images of the hand. The system demonstrates remarkable accuracy in recognizing 11 distinct hand gestures in real-time.

This paper [3] discusses the creation of a system using machine learning methodologies for recognition of hand gestures. The researchers evaluate the performance of various feature extraction methods, such as HOG, LBP, and SURF, alongside different classifiers, such as KNN, SVM, and decision tree. The findings indicate that the combination of HOG features and KNN classifier yields the highest accuracy.

This paper [4] suggested a hand gesture recognition system employing on principal component analysis (PCA) in combination of support vector machines (SVM). The authors use PCA to reduce the dimensionality of the hand gesture features, while SVM is employed for gesture classification. The system attains notable accuracy in recognizing 10 hand gestures.

In this paper [5] a comparative study is presented of different hand gesture recognition methods, including rule-based, template matching, and machine learning-based approaches. The authors assess the effectiveness of these approaches on a dataset comprising of 10 hand gestures and scrutinizing their accuracy and computational complexity. Additionally, Researchers Cui and Sun employ background technique to analyze hand distribution in video files. The hand shape of average hand size is used for segmentation to detect movements and ascertain hand curvature.

In this paper [6] a novel approach is adopted to enhance outcomes, researchers integrate support methods and skin color prior to mindfulness monitoring, advocating for education and guidance. Researcher Fracnke achieved better results than the support algorithm. The method exhibits impressive simulation and detection rate up to 97%, the follow-up rates up to 99%, and with the recognition rates up to 99%. This has become a trending research domain for deep learning-based gesture recognition. Despite the capabilities of existing algorithms and peripherals for gesture recognition, cost and consumable limitations hinder meeting the requirements.

This paper [7] describes a static American Sign Language (ASL) hand gesture recognition system is created utilizing deep convolutional neural networks. The architecture of this system is crafted to be lightweight, facilitating easy deployment and mobility. The authors used different image processing techniques for background subtraction and frame segmentation to enhance accuracy in real-time scenarios. The authors achieved a testing accuracy rate of 96% while emphasizing mobility, cost-effectiveness, and ease of deployment environments with limited computational resources.

In this paper [8], a system for recognizing hand gestures is created utilizing convolutional neural networks (CNN). The system generated by Convolution neural networks (CNN) helps in keeping track of complex movements of various body parts, such as hands, face and arms, and extracting the prime features of the movements. The researchers in this paper trained and tested dataset regularly with the aid of images to check the effectiveness of CNN. A general extraction of around 86% was achieved for images.

In this paper [9], a segmentation technique is introduced to detect hand gestures from input images, thereby enhancing recognition accuracy. The authors evaluated and compared segmentation techniques including SkinMask, YCbCr, and HSV (hue, saturation, and value) for hand gestures. The YCbCr method involved deriving CR component followed by hole filling, erosion, and binarization operations. SkinMask employed color segmentation to detect hand-colored pixels, while HSV used threshold masking to identify dominant features. For hand gesture classification, Softmax classification is used with features extracted via convolutional neural network. This paper resulted in improved recognition accuracy compared to existing cutting-edge systems.

In this paper [10], a system is devised to comprehend and interpret signs. By utilizing vision-based hand gesture recognition and detection an isolated recognition model is created. The machine learning based SLR model was assessed under controlled environment with the help of 4 candidates. The model utilized convex hull for feature extraction and KNN for classification, which resulted in yielded a accuracy of 65%.

In this paper [11], a simple feature extraction method is proposed to extract features from video sign language by using Discrete Cosine Transform (DCT) and the area of objects in a binary image. A straightforward neural network model is created to recognize gestures based on features extracted from the video stream. Additionally, a system with audio is integrated to play specific words associated to the gestures. An accuracy rate of 91% was achieved using the proposed model.

In this paper [12], an effective real-time method for recognizing gestures made by hands is proposed. The process begins with triggering of hand detection followed by tracking with a specific gesture. Subsequently, the hand segmentation is achieved utilizing color and motion cues. Finally, a scale-space feature detection is incorporated into gesture recognition to overcome the limitations such as aspect ratio which is prevalent in many hand-gesture techniques created based on learnings. The researchers achieved a satisfactory performance for image browsing using proposed method.

In this paper [13], an analysis is conducted on various widely used classification techniques, random forest, Naïve Bayes, XGBoost, K-Nearest Neighbor (KNN), logistic regression, Convolution Neural Network (CNN), Stochastic Gradient Descent Classifier (SGDC), and Support vector classifier (SVC). The authors found that among traditional machine-learning classifiers like SGDC, Naïve Bayes, SVC, KNN, Logistic Regression and XGBoost are outperformed by the random forest and the sign language MNIST dataset exhibiting superior predictive accuracy. However, they obtained the best results using CNN algorithm.

This paper [14] introduces a new Leap Motion controller gen.2 based hand gesture recognition system. The system incorporates the Spatial Fuzzy Matching (SFM) algorithm, which was first demonstrated by matching and combining spatial data to produce a fused gesture dataset. A suggested initial frame correction technique based on SFM seeks to quickly initialize test gesture trajectories with respect to the gesture dataset in order to facilitate dynamic hand recognition. The fused dataset is small enough to run on standard laptops, which speeds up the recognition rate calculation. Using the

# Machine Learning driven Sign Language Recognition with Optimized Analysis and Multilingual Speech Generation

gathered dataset, the system recognized hand motions at 94%–100% for static gestures and more than 90% for dynamic gestures.

### III. MOTIVATION

The motivation behind developing American Sign Language (ASL) hand gesture recognition system originates from the necessity to facilitate communication for individuals from communities of hard of hearing or deaf. ASL, a holistic and innate form of communication employing hand gestures, facial expressions, and body language to convey messages. Through the interpretation of ASL hand gestures, computers and machines can serve as intermediaries, helping in communication between individuals with hearing disabilities and rest of the world. This advancement has the potential to enhance accessibility and inclusivity, enabling individuals with hearing impairments to more actively participate in both social and professional contexts. By harnessing hand gesture recognition technology for ASL, we can effectively bridge communication gaps and empower those whose ability to communicate might otherwise be restricted. Hand gesture recognition for ASL can also be used for people to learn sign language.

### IV. PROPOSED WORK

The aim of this proposed work is to construct a machine learning model with the ability to recognize American Sign Language (ASL) hand gestures. ASL serves as the principal language for the deaf and hard-of-hearing, with hand gestures playing a crucial role in conveying meaning. To achieve this aim, we will compile a dataset comprising of images of hand gestures commonly used in ASL. The images will be captured using a high-resolution camera ensuring controlled lighting conditions and consistent background settings. We will annotate the images with the corresponding ASL symbols, creating a labelled dataset of 26 letters, each with 100 images. Subsequently, after completion of the collecting the images then we will preprocess the dataset using the MediaPipe library to detect hand landmarks. Developed by Google, MediaPipe is open source, cross-platform and customizable framework for machine learning solutions which is easy to understand, offering some advanced machine learning features like face detection, hand recognition, prediction, object detection and more. For hand recognition, for each detected hand MediaPipe detects 21 key points, normalizing and saving the hand landmarks along with the corresponding labels in a pickle file. The dataset will then be divided into two subsets, for testing and training. We will use 80% from the created dataset for training and 20% for testing. Various algorithms like Naive Bayes algorithm, Random Forest algorithm, Multinomial Naive Bayes algorithm, Logistic Regression algorithm, Support Vector Machines (SVM) algorithm, Linear Support Vector Machines (SVM) algorithm, KNN (K Nearest Neighbour) algorithm, Decision Tree algorithm will be used to train and test the model, aiming to identify the most effective one.

Following, the model is trained the trained model will be saved using the pickle.dump method and store it in a file with the '.p' extension. Performance evaluation on the test set will be conducted, measuring accuracy. We will also perform cross-validation to assess the model's generalization ability.

Finally, the trained model will be deployed as a real-time application that can recognize hand gestures in live video streams. We will develop a user interface that displays the recognized ASL symbols in real-time and within this interface a string will be generated based on the predicted hand signs in the live video stream. This string can subsequently convert to speech in selected language to facilitating communication for communities of individual who are hard-of-hearing and deaf with others. The interface will incorporate several features to enhance usability, such as pausing string generation, resetting string, and removing last character from the string. These functionalities aim to assist deaf and hard-of-hearing users in composing messages efficiently, allowing them to convey complete message in single go rather than predicting individual words one at a time. Moreover, the converted speech can be rendered in a variety of languages based on user preference.

To sum up, the aim of this proposed work is to create a machine learning model for American Sign Language (ASL) hand gesture recognition system, employing Random Forest classification. The envisioned model has the potential to enhance the inclusivity and accessibility of communication for communities of individuals who are hard-of-hearing and deaf.

### V. RESULTS AND DISCUSSION

The experimental outcomes are carried out using the python programming language. The process for ASL Hand gesture Recognition comprises four primary stages: data collection and preprocessing, followed by training and testing of the data, then final recognition of hand gestures for ASL in live video streams.

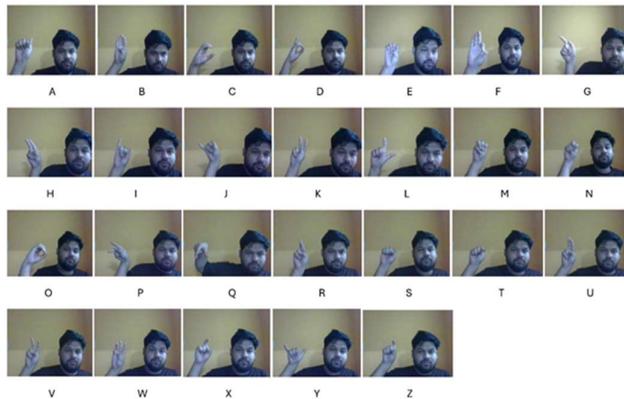


Fig. 1. American Sign Language from A-to-Z data.

The depicted figure shows the initial step, which involves gathering data for hand gesture recognition via the system's camera, using openCv library. For each character in ASL, spanning from A to Z, a total of 100 images have been collected for both training and testing.

# Machine Learning driven Sign Language Recognition with Optimized Analysis and Multilingual Speech Generation

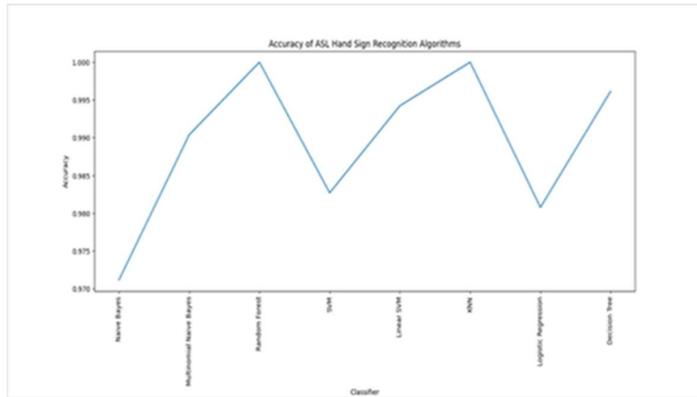


Fig. 2. The comparative performance of Random Forest compared with remaining algorithms. The graph shows the accuracy of the Naïve Bayes, Multinomial Naive Bayes, Random Forest, SVM, Linear SVM, KNN, Logistic regression, Decision trees and Random Forest achieves the highest accuracy at 99.9%, followed by SVM with 97.5%, KNN with 98.2% and Naive Bayes with 93.86%. Random Forest, characterized by its collaborative learning method integrate multiple decision trees to enhance the accuracy and robustness of the model. Additionally, Random Forrest demonstrates faster testing speeds compared to other algorithms.

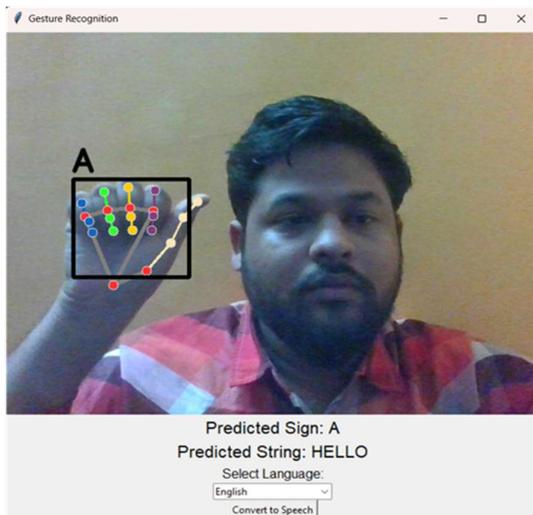


Fig. 3. Final output of ASL hand gesture recognition and string generated in live video.

This image shows that the system is recognizing all the signs made by the user with high accuracy and creating a string from the predicted hand signs.

## VI. CONCLUSION

This study concentrated on enhancing the accuracy of the ASL Hand gesture recognition. We employed MediaPipe to recognize the hand gestures and the hand key points. After training with various models, we find random forest is best and most effective model. Training with random forest

yielded an accuracy rate of 99.9% on our dataset which is very efficient for complex structured hand gestures of the sign language. The system's ultimate outcome is predicting all the characters from A to Z in live video stream with efficient accuracy in ASL which used to generate the string and the system is efficiently converting the generated speech to multiple language.

## REFERENCES

- [1] A Survey on Hand Gesture Recognition Techniques and Applications" by K. Selvakumar and R. Kavitha (2017).
- [2] Real-time Hand Gesture Recognition using Convolutional Neural Networks" by C. K. Kim, Y. H. Kim, and S. H. Lee (2018).
- [3] Hand Gesture Recognition using Machine Learning Techniques" by S. R. Maity, A. K. Das, and D. Bhattacharjee (2019).
- [4] Hand Gesture Recognition using Principal Component Analysis and Support Vector Machines" by P. S. Hiremath, R. Kulkarni, and S. K. Koppa (2017).
- [5] A Comparative Study of Hand Gesture Recognition Techniques" by N. Srivastava, R. Tripathi, and M. Singh (2017).
- [6] Shams and T. H. Kim, "Hand Gesture Recognition using Deep Learning Techniques: A Comprehensive Review," in 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2020.
- [7] R. Patel, J. Dhakad, K. Desai, T. Gupta and S. Correia, "Hand gesture recognition system using convolutional neural networks", 2018 4th international conference on computing communication and automation (ICCCA), pp. 1-6, December 2018.
- [8] A, Lakshmanan & Parasuraman, Dhanasekaran & Mahalakshmi, P. & Ashok, Sharmila, "Hand Gesture Recognition using Convolutional Neural Network". 1-5, i-PACT (2021).
- [9] Rahim, M., Miah, A., Sayeed A. and Shin J. (2020), "Hand Gesture Recognition Based on Optimal Segmentation in Human-Computer Interaction", Proceedings of the 3rd IEEE International Conference

## Machine Learning driven Sign Language Recognition with Optimized Analysis and Multilingual Speech Generation

- on Knowledge Innovation and Invention (ICKII), Kaohsiung, Taiwan, P. 163–166.
- [10] Kallingale, Amrutha & Prabu, P. (2021). “ML Based Sign Language Recognition System”. In 2021 International Conference on Innovative Trends in Information Technology. pp. 1-6. doi: 10.1109/ICITIT51526.2021.9399594.
- [11] M. P. Paulraj, S. Yaacob, H. Desa, C. R. Hema, and W. M. R. Wan Ab Majid, “Extraction of head and hand gesture features for recognition of sign language”, 2008 Int. Conf. Electron. Des. ICED 2008, pp. 1-6. doi: 10.1109/ICED.2008.4786633.
- [12] Y. Fang, K. Wang, J. Cheng, and H. Lu, “A real-time hand gesture recognition method.”, In 2007 at IEEE International Conference on Multimedia and Expo, pp. 995-998. IEE.
- [13] Bhushan, S.; Alshehri, M.; Keshta, I.; Chakraverti, A.K.; Rajpurohit, J.; Abugabah, A. An Experimental Analysis of Various Machine Learning Algorithms for Hand Gesture Recognition. *Electronics* 2022, 11, 968. <https://doi.org/10.3390/electronics11060968>
- [14] H. Li, L. Wu, H. Wang, C. Han, W. Quan and J. Zhao, "Hand Gesture Recognition Enhancement Based on Spatial Fuzzy Matching in Leap Motion," in *IEEE Transactions on Industrial Informatics*, vol. 16, no. 3, pp. 1885-1894, March 2020, doi: 10.1109/TII.2019.2931140.