IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

SIGN LANGUAGE RECOGNITION USING MACHINE LEARNING

A. LAHARI
Information Technology GVPCEW
Visakhapatnam, India

T. Mounika
Information Technology GVPCEW
Visakhapatnam, India

V. Harika
Information Technology GVPCEW
Visakhapatnam, India

P. Rajeswari
Information Technology GVPCEW
Visakhapatnam, India

ABSTRACT:

Sign language is a visual language that uses hand gestures to convey meaning, primarily used by the deaf and hearing-impaired community. Sign language recognition systems aim to bridge the communication gap between individuals who use sign language and the wider hearing population by translating sign language gestures into written or spoken language. Existing systems rely on a range of techniques including computer vision, machine learning models such as convolutional neural networks (CNNs), and traditional methods like support vector machines (SVMs). However, these existing models face limitations in recognizing a broad vocabulary and complex grammatical structures of sign languages, often leading to ambiguity and misinterpretation. The proposed system addresses these challenges by expanding the model to include more signs, resulting in a more effective framework for real-time applications. By integrating advanced computer vision techniques and machine learning algorithms, the system can recognize hand gestures based on features such as orientation, centroid, and finger positions. Leveraging a comprehensive dataset for diverse signing styles enhances generalization and fosters more effective communication for individuals with hearing impairments.

KEYWORDS: Sign Language, Gesture Recognition, SVM, Random Forest, Machine Learning.

I. INTRODUCTION

Sign language is a visual-gestural language used by deaf and hard-hearing individuals for communication purposes. It employs hand movements to convey meanings. Unlike spoken languages that utilize oratory faculties to produce sounds mapped against specific words and grammatical structures, sign language uses visual faculties and is governed by its own complex grammar and syntax. A sign language recognition system provides an easy, efficient, and accurate mechanism to transform sign language into text or speech, enabling communication between the hearing-impaired community and the general public. These systems leverage various classification methods to recognize alphabet flow and interpret sign language words and phrases. Sign languages are collections of hand gestures with specific meanings, employed by hearing-impaired individuals to communicate in everyday life. There are over 300 different sign languages worldwide, with each having its unique vocabulary and grammar. However, the percentage of the population fluent in any sign language is relatively low, making it difficult for the specially-abled to communicate freely with others. Sign language recognition systems offer a means to bridge this gap by translating gestures into a commonly spoken language such as English. In the realm of communication accessibility, sign language recognition technology plays an essential role in facilitating real-time translation of sign language gestures into spoken or written language. Advanced computer vision and machine learning algorithms enable SLR systems to recognize a wide range of signs accurately. The ongoing development in this technology aims to expand its ability to recognize more signs ultimately enhancing communication for individuals with hearing impairments.

II. RELATED WORK

In the realm of sign language recognition, efforts have been directed towards enhancing accessibility and inclusivity for individuals reliant on sign language for communication. Previous work has strived to overcome the constraints of existing models by exploring some methodologies and technologies. One notable approach involves expanding the vocabulary of models by incorporating additional signs, aiming to boost their effectiveness in real-time scenarios. Through the integration of advanced computer vision techniques and machine learning algorithms, we have sought to broaden the spectrum of recognizable signs, facilitating the formation of simple words and sentences to cater to various recognition tasks.

Machine learning algorithms, such as Random Forest and Support Vector Machines, have played a pivotal role in the development of sign language recognition systems. These algorithms predict labels based on features extracted from hand gestures, including orientation, centroid, finger status, and thumb position. By leveraging diverse datasets and enabling real-time processing, efforts have been made to enhance the accuracy and efficiency of recognition systems. Furthermore, the incorporation of speech synthesis functionality has improved usability by allowing for the vocalization of recognized signs, thereby aiding individuals with hearing impairments in communication.

Overall, advancements in sign language recognition technology signify progress towards improving accessibility and inclusivity. By addressing the limitations of existing models and leveraging innovative methodologies, these systems offer enhanced capabilities for recognizing and interpreting sign language gestures. Such developments hold promise for fostering better communication outcomes and fostering greater understanding and engagement among individuals who rely on sign language as their primary mode of expression.

III. EXISTING SYSTEM

Existing sign language recognition models leverage a combination of computer vision and deep learning methodologies, integrating convolutional neural networks (CNNs), recurrent neural networks (RNNs), and traditional machine learning techniques like Support Vector Machines (SVMs). These methodologies enhance accuracy and efficiency, thereby improving accessibility.

However, existing system face several limitations. Firstly, they suffer from a limited vocabulary, restricting their ability to interpret a wide range of signs and hindering effective communication. Additionally, ambiguity and misinterpretation persist due to the complexity of sign language gestures, leading to inaccuracies in translation. Moreover, the lack of comprehensive coverage across various sign languages hampers inclusivity for users from diverse linguistic backgrounds. Furthermore, existing systems struggle to accurately identify subtle or complex hand movements, limiting their utility in real-world scenarios.

IV. PROPOSED SYSTEM

The proposed system builds upon the foundation of the existing model by addressing its limitations through an expanded framework. By augmenting the model with additional signs, the proposed system aims to enhance its effectiveness for real-time applications. This expansion facilitates the recognition of a broader range of signs and enables the formation of simple words and sentences, catering to both continuous and isolated recognition tasks. To achieve this, advanced computer vision techniques and machine learning algorithms are integrated into the system, enabling real-time recognition of hand gestures based on various features such as orientation, centroid, finger status, and thumb position.

The proposed system performs prediction of labels using machine learning algorithms, specifically Random Forest and Support Vector Machine models based on these predicted labels, the system constructs words and sentences, thereby facilitating meaningful communication through sign language. Additionally, to enhance usability, the system incorporates speech synthesis functionality, allowing for the vocalization of the recognized signs. Overall, the proposed system represents a significant advancement in sign language recognition technology, offering enhanced accessibility and inclusivity for individuals who rely on sign language for communication.

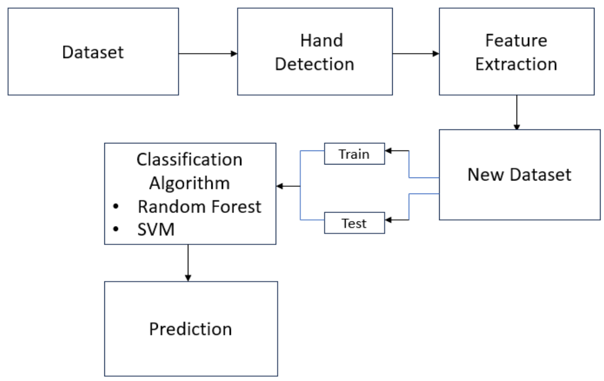


Fig-1. Proposed System

V. DATASET USED

We created a comprehensive dataset of hand gestures for sign language recognition, capturing real-time video frames using a webcam. The dataset includes 36 distinct classes, representing the letters A-Z and digits 0-9. For each class, we collected around 300 images to ensure a diverse and robust dataset. The images were organized into a structured directory system, with each class having its own dedicated folder. This organization allows for easy access and management of the dataset during training and evaluation of sign language recognition models. To capture the gestures, we prompted a user to perform each desired sign in front of the camera, gathering multiple samples of each gesture to enhance the variety and representativeness of the dataset. Overall, the dataset contains approximately 10,800 images of hand gestures, providing a rich and diverse collection that encompasses the alphabet and digits. This comprehensive coverage offers a strong foundation for training and evaluating sign language recognition models, enabling them to learn and generalize patterns across different contexts and hand positions.

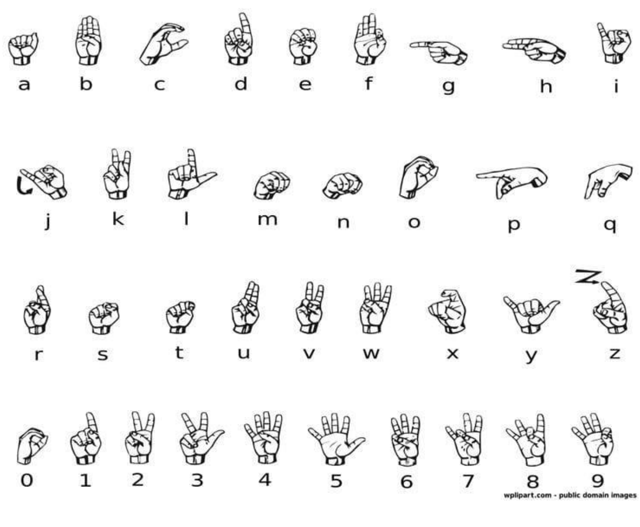


Fig-2. Sign Language Gestures from A-Z and 0-9

VI. METHODOLOGY

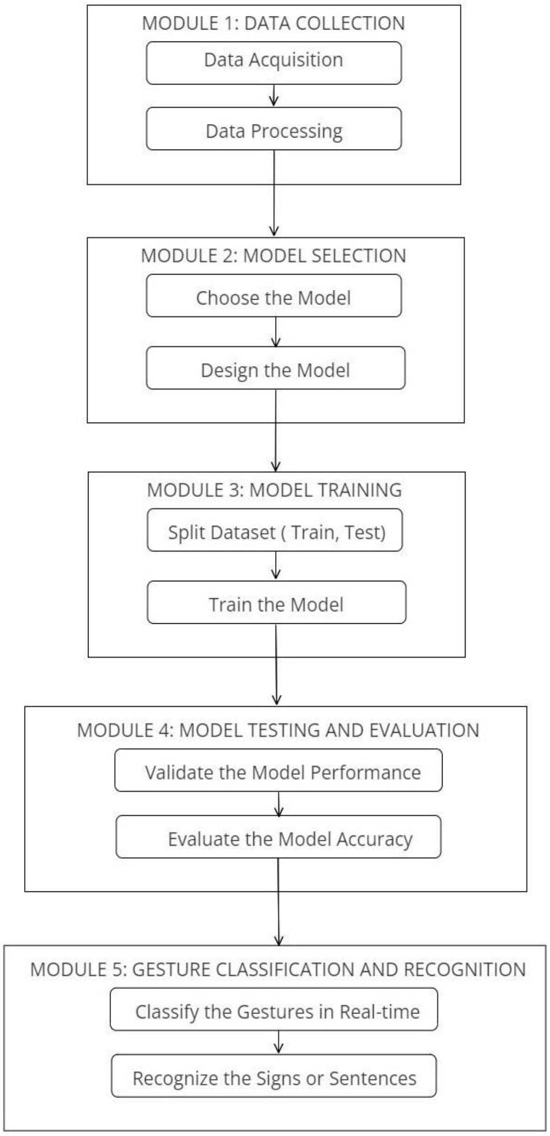


Fig-3. Dataflow of proposed system

A. Data Collection

Data collection module involves collecting image data using a webcam for a specified number of classes and dataset size. It imports necessary libraries, sets up a directory to save the data, initializes parameters such as the number of classes and dataset size, and opens a webcam stream. For each class, it creates a directory if it doesn't exist already and prompts the user to press 'Q' when ready to start capturing images. Then, it captures dataset size number of frames from the webcam, saving each frame as a JPEG image in the corresponding class directory. Finally, it releases the webcam stream and closes all windows. Data processing reads and processes images from the dataset directory. It detects hand landmarks using the MediaPipe Hands model and then extracts normalized coordinates of hand landmarks relative to the image size. We then save the extracted data and labels into a pickle file named 'data.pickle'.

B. Model Selection

In the process of selecting the appropriate model for our sign language recognition system, we conducted a thorough analysis of our dataset, consisting of diverse hand gestures representing letters and digits. After careful consideration, we identified Random Forest and Support Vector Machines (SVMs) as potential candidates due to their respective strengths in handling complex datasets and non-linear relationships within the data. Through rigorous testing and validation procedures, including parameter tuning and performance evaluation, we determined that both Random Forest and SVMs exhibited commendable accuracy and efficacy in predicting sign gestures. Consequently, we opted to integrate both models into our system, leveraging their complementary attributes to enhance the robustness and reliability of our sign language recognition framework. This strategic approach ensures the system's adaptability and proficiency in accommodating various sign language expressions, thereby advancing accessibility and inclusivity for individuals reliant on sign language for communication.

C. Model Training

The collected data is divided into two sets: a training set used to train the model and a test set used to evaluate its performance on unseen data using the train_test_split function from scikit-learn. We split the data into training and test dataset in 80:20 ratio. The training set is used to train the model, while the testing set is used to evaluate its performance. This split ensures that the model is tested on data it has not seen before, providing an unbiased evaluation. The model is trained using the training set (features and labels), adjusting its internal parameters to minimize the classification error. Training involves fitting the model to the training data and optimizing its performance using a chosen loss function.

D. Model Testing and Evaluation

In our model testing and evaluation phase, we employed a combination of computer vision techniques and machine learning algorithms to interpret hand gestures captured by a webcam in real-time. Utilizing the trained model, loaded from a saved file, we processed each frame of the video stream to detect hand landmarks using the MediaPipe Hands model. By extracting coordinates of these landmarks relative to the image size, we prepared the data for prediction. The model then classified each gesture, assigning it a corresponding letter or digit based on a predefined mapping. Throughout this process, we ensured visual feedback by overlaying the recognized letter onto the video feed and providing audio feedback through text-to-speech functionality. Additionally, we implemented functionality to handle user input, facilitating the formation and vocalization of words and sentences. This iterative process of prediction, feedback, and user interaction contributed to the refinement and validation of our sign language recognition system, ultimately enhancing its usability and effectiveness in real-world applications.

E. Gesture Classification and Recognition

When evaluating the effectiveness of a model intended for sign language recognition, model testing is a crucial stage. This stage involves measuring the accuracy, precision, recall and F1-score of the trained randomforest and svm models using the created dataset. The model testing and evaluation is done separately.
Formulas of evaluation metrics:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Specificity} = \frac{TN}{TN + FP}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

Measure	SVM (%)	Random Forest (%)
Precision	99.66710131183815	100.0
Recall	99.66666666666667	100.0
F1 Score	99.66665185119338	100.0
Test Accuracy	99.66666666666667	100.0

Fig-4. Metrics Table

VII. IMPLEMENTATION DETAILS

A. PYTHON

We have developed this system using python. To work with ML models, python is efficient. Python offers wide range of libraries like OpenCV, Keras, TensorFlow. The platform we used is Pycharm.

B. PYCHARM

Pycharm is a powerful developer tool that you can use to complete the entire development cycle in one place. It is a comprehensive integrated development environment (IDE) that you can use to write, edit, debug, and build code, and then deploy your app.

XIII. RESULTS

We gained an accuracy of 100% for randomforest model by training the model with 100 estimators i.e., 100 random forest trees, using images from the collected dataset. We gained an accuracy of 99.6% for the SVM model. Both models achieved high accuracy, showing effective generalization and robust performance in recognizing different gestures.

The below Fig.5 represents the confusion matrix for randomforest classifier where the predictions are done on test set. The no. of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) predictions that the model made is displayed in the confusion matrix for classes.

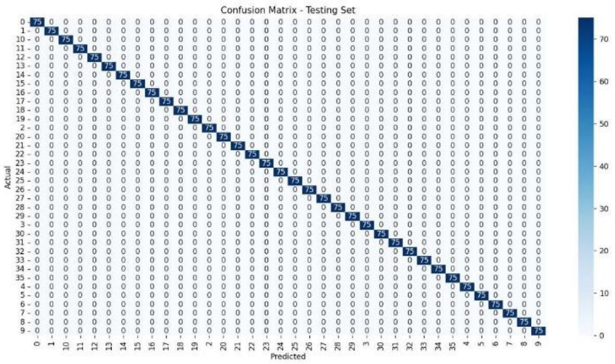


Fig-5. Confusion matrix for RandomForest

The below Fig.6 represents the confusion matrix for SVM classifier where the predictions are done on test set. The no. of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) predictions that the model made is displayed in the confusion matrix for classes.

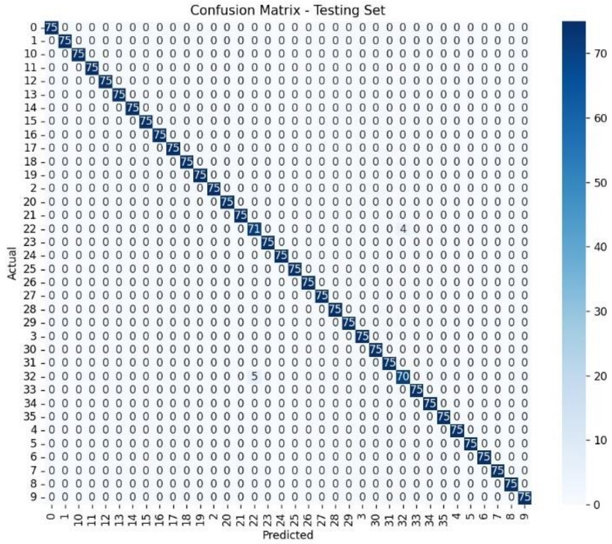
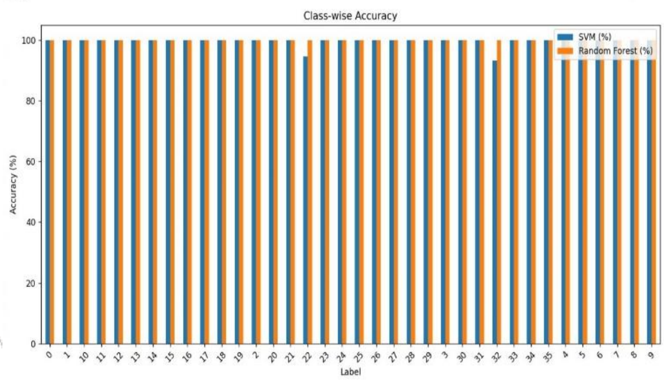


Fig-6. Confusion matrix for SVM

The below Fig.7 represents class-wise accuracy for randomforest and SVM.



IX. CONCLUSION

This project is all about helping people communicate better through sign language. We started by gathering a bunch of hand gesture images, covering every letter from A to Z and numbers 0 to 9. We made sure to collect 300 images for each gesture, giving us a diverse dataset to work with. Then, we organized all this data into a pickle file to use for training our model. We trained our system to recognize these gestures in real-time. So now, when someone makes a sign in front of their webcam, our system instantly turns it into text and even speaks it out loud. It's all about making communication more accessible and inclusive for everyone. This project is our way of using technology to break down barriers and bring people closer together. Our goal is to foster greater understanding and connection within our society, ultimately contributing to a more inclusive and empathetic world.

X. REFERENCES

[1]. A. Kumar, K. Thankachan and M. M. Dominic, "Sign language recognition," 2016 3rd International Conference on Recent Advances in Information Technology (RAIT),

Dhanbad, India, 2016, pp. 422-428, doi: 10.1109/RAIT.2016.7507939.

[2]. A. S. Nikam and A. G. Ambekar, "Sign language recognition using image based hand gesture recognition techniques," 2016 Online International Conference on Green Engineering and Technologies (IC-GET), Coimbatore, India, 2016, pp. 1-5, doi: 10.1109/GET.2016.7916786.

[3]. M. Jha, H. P. Rajawat, H. Suhagiya and L. Kurup, "Indian Sign Language recognition using LSTM and mediapipe," 2023 International Conference on Advanced Computing Technologies and Applications (ICACTA), Mumbai, India, 2023, pp. 1-6, doi: 10.1109/ICACTA58201.2023.10393604.

[4]. K. K. Dutta and S. A. S. Bellary, "Machine Learning Techniques for Indian Sign Language Recognition," 2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC), Mysore, India, 2017, pp. 333-336, doi: 10.1109/CTCEEC.2017.8454988.

[5]. Q. Chen, N. D. Georganas and E. M. Petriu, "Real-time Vision-based Hand Gesture Recognition Using Haar-like Features," 2007 IEEE Instrumentation & Measurement Technology Conference IMTC 2007, Warsaw, Poland, 2007, pp. 1-6, doi: 10.1109/IMTC.2007.379068

[6]. M. M. Islam, S. Siddiqua and J. Afnan, "Real time Hand Gesture Recognition using different algorithms based on American Sign Language," 2017 IEEE International Conference on Imaging, Vision & Pattern Recognition (icIVPR), Dhaka, Bangladesh, 2017, pp. 1-6, doi: 10.1109/ICIVPR.2017.7890854.

[7]. X. Liu, J. Sacks, M. Zhang, A. G. Richardson, T. H. Lucas and J. Van der Spiegel, "The Virtual Trackpad: An Electromyography-Based, Wireless, Real-Time, Low-Power, Embedded Hand-Gesture-Recognition System Using an Event-Driven Artificial Neural Network," in IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 64, no. 11, pp. 1257-1261, Nov. 2017, doi: 10.1109/TCSII.2016.2635674.

[8]. J. Farooq and M. B. Ali, "Real time hand gesture recognition for computer interaction," 2014 International Conference on Robotics and Emerging Allied Technologies in Engineering (iCREATE), Islamabad, Pakistan, 2014, pp. 73-77, doi: 10.1109/iCREATE.2014.6828342.

[9]. S. Reshna and M. Jayaraju, "Spotting and recognition of hand gesture for Indian sign language recognition system with skin segmentation and SVM," 2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), Chennai, India, 2017, pp. 386-390, doi:10.1109/WiSPNET.2017.8299784

[10]. Feng-Sheng Chen, Chih-Ming Fu, Chung-Lin Huang, Hand gesture recognition using a real-time tracking method and hidden Markov models, Image and Vision Computing, Volume 21, Issue 8, 2003, Pages 745-758, ISSN 0262-8856, https://doi.org/10.1016/S0262-8856(03)00070-2.