# Hand Sign To Text Conversion Using ML

Sonia Waghmare, Praveen Rathod, Ved Satdeve, Rohit Gaware, Ritesh Khapare

*Computer Department, Savitribai Phule Pune University*

*Abstract*— Hand sign language is an essential form of communication for individuals with hearing impairments. However, the language barrier between sign language users and non-sign language users often hinders effective communication. This paper presents a novel approach to bridge this gap by developing a system that translates hand signs into text using machine learning techniques, augmented with a built-in voice output. The system includes a voice output feature that is built-in to improve the user experience. Real-time communication is made possible once the hand sign has been identified and translated into text by the system, which also produces a corresponding audio output. Text-to-speech algorithms are used to synthesise the voice output, which gives non-sign language users a spoken version of the sign language message.

**Keywords**— ***Image recognition, Hand Gesture, LSTM Algorithm, MediaPipe, Python packet manager, Machine Learning***

## INTRODUCTION

Hearing is one of the most essential mortal senses, yet not everyone possesses this gift. According to the World Health Organization, roughly 360 millions of people worldwide are affected by disabling hail loss. The most natural and suggestive system for hail- disabled people is sign language. Non-deaf people noway essay to learn sign language in order to communicate with those who are deaf. The result is that deaf persons come insulated. still, the gap between the hearing community and the general population could be reduced if the computer could be designed to restate sign language to textbook format. A sign language is a way of communicating by using the hands and other corridor of the body. The main idea is to convert sign language to textbook. The frame help speech bloodied to communicate with the society using sign language. This leads to the elimination of the middleman who generally acts as a medium of restatement. This would contain a stoner friendly terrain for the stoner by furnishing textbook affair for a sign gesture input. The signs are captured at real time by using web cam. The uprooted features are compared by using bracket algorithm. In order to calculate the sign recognition, the features are compared with testing database. Eventually, honored gesture is converted into textbook and textbook get displayed on examiner. Our proposed system provides an occasion for a deaf and dumb people to communicate withnon-signing people without the need of an practitioner. It's veritably easy to get two fully different signs mixed up which leads to bad miscommunication, So we've used proper dataset which train our model similar that miscommunication can be avoided.

Our team used machine learning to create a hand sign to text translation model that also outputs voice. For the purpose of identifying and translating hand signs into text in this, we employed the LSTM algorithm, Mediapipe, and OpenCv.We implemented this model in the backend Django application and the front end Vue JS. Simple registration and login pages make up this web application.

## EXISTING SYSTEM

Existing systems for hand sign to text conversion rely on computer vision techniques and machine learning algorithms to recognize and interpret hand signs captured by a camera or sensor, and then convert them into textual representations. Examples include pose-based hand sign recognition, image-based hand sign recognition, sensor-based hand sign recognition, and hybrid approaches. Pose-based hand sign recognition uses deep learning models to classify hand poses, image-based hand sign recognition uses computer vision techniques to process 2D images of hand signs, sensor-based hand sign recognition uses wearable sensors to capture hand movements and gestures, and hybrid approaches combine multiple modalities to enhance accuracy and robustness.

## I. SYSTEM IMPLEMENTATION

This application comprises of four modules: text conversion, feature extraction from the dataset, and model development with LSTM and TensorFlow.The dataset for the hand sign was constructed via the dataset acquisition module. It consists of 30 videos, each running at 30 frames per second, totaling 900 frames. The feature extraction module, which uses Mediapipe and OpenCv to capture real-time video, comes after the dataset is created. For the sequential data, we have utilised the Long Short Term Memory (LSTM) algorithm, a form of RNN.

LSTM networks can capture long-term dependencies and manage variable-length input sequences, making them an excellent choice for converting hand signs to text. Sequential input data are processed using dataset preparation, image preprocessing, and LSTM network architecture to generate corresponding text outputs. In order to extract useful characteristics from the input images, convolutional layers are used. In terms of accurately recognising and translating hand signs into textual representations, LSTM-based architectures for hand sign to text conversion have shown encouraging results. However, the architecture, hyperparameter tuning, and quality and diversity of the training dataset all have a significant impact on the system's performance.

For deploying our application, we used the Django Python web framework and Vue.js. A well-liked JavaScript framework called vue.js is used to create single-page apps and user interfaces. Declarative rendering, component-based architecture, and reactive data binding are some of its primary characteristics.The Model-View-Controller (MVC) architectural pattern is followed by the high-level Python web framework Django, which also offers a potent ORM, URL routing, and template engine. It offers capabilities including model definition, querying, and migrations as well as support for many database backends.

## II. SYSTEM EVALUATION

### Evaluation Method:

People are increasingly using hand gesture to text (HAND SIGN TO TEXT CONVERSION) systems to interface with computers and other devices. These systems recognise hand gestures and translate them into text using machine learning. The effectiveness of HAND SIGN TO TEXT CONVERSION systems can be evaluated using a variety of different evaluation indicators. Accuracy and usability are the two categories into which these measurements normally fall.The accuracy of a system's hand gesture recognition is measured. This is often gauged by the accuracy rate, or the proportion of correctly identified gestures.How simple it is for people to interact with the system is measured by usability. The user satisfaction rate, or the proportion of users who are happy with the system, is often used to gauge this.

HAND SIGN TO TEXT CONVERSION systems can be made more accurate and user-friendly by a number of aspects, such as expanding the training dataset. The system's ability to recognise hand movements will improve with more training data implementing a more complex machine learning algorithm. More advanced algorithms are better able to recognise hand motion patterns utilising an improved user interface. Users may interact with the system more easily if the user interface is improved.People could interact with computers and other gadgets more easily with the help of HAND SIGN TO TEXT CONVERSION systems. We can increase the accessibility of these technologies to a wider group of users by increasing their accuracy and usability.

The following are some difficulties in assessing HAND SIGN TO TEXT CONVERSION systems the absence of a uniform evaluation procedure. There is no solitary accepted method for assessing HAND SIGN TO TEXT CONVERSION systems. As a result, evaluating the state of the art and comparing various systems is challengingvariations in hand movements. Depending on the person, the situation, and the surroundings, hand gestures might change. This makes it challenging to create systems that can precisely identify all hand motions a lack of sufficient training datasets. For training, HAND SIGN TO TEXT CONVERSION systems need a lot of data. However, there isn't yet a training dataset for HAND SIGN TO TEXT CONVERSION systems that is substantial enough. The effectiveness of these systems is so constrained.

Despite these difficulties, HAND SIGN TO TEXT CONVERSION systems have the potential to be an important tool for those who need to engage with computers more naturally or who have limitations. We can improve the accuracy, usability, and accessibility of HAND SIGN TO TEXT CONVERSION systems by addressing the difficulties involved in evaluating them.

HAND SIGN TO TEXT CONVERSION systems, which use hand gestures to interact with computers and other devices, are becoming more and more popular. These systems use machine learning to recognise hand gestures and convert them into text. Several different evaluation metrics can be used to assess the efficacy of HAND SIGN TO TEXT CONVERSION systems. These measurements usually fall within the categories of accuracy and usability.A system's ability to recognise hand gestures accurately is evaluated. The proportion of correctly identified gestures, or accuracy rate, is a common metric for determining this. Usability refers to how easy it is for users to interact with the system. This is frequently determined by looking at the user satisfaction rate, or the percentage of users who are content with the system.

## Results of the Evaluation:

The results of the evaluation of hand gesture to text using machine learning are promising. The accuracy of these systems has been shown to be high, with some systems achieving accuracies of over 85%. The usability of these systems has also been shown to be good, with users generally finding them easy to use.

Accuracy: The accuracy of sign detection systems has been shown to be high, with some systems achieving accuracies of over 85%.Our training model shows upto 85% of accuracy while training .

Usability: The usability of sign detection systems has also been shown to be good, with users generally finding them easy to use.

Challenges: There are still some challenges that need to be addressed before these systems can be widely adopted. The variety of hand gestures is a problem. Lack of a sufficient training dataset is another problem. We require a background that is flat and free of distortions or other elements that might prevent our system from correctly predicting the hand sign.

Potential: For anyone who desire to communicate with computers more naturally, including those with impairments, sign detection systems have the potential to be a useful tool. We can improve the effectiveness, usability, and accessibility of sign detection systems by addressing the difficulties in evaluating them.

## III. DISCUSSION

The LSTM technique is used in our implementation of hand sign to text conversion to convert hand signs to text. Our accuracy rate for this assignment was 85%.Our own dataset, which consists of 30 frames per second, has been trained. The goal of hand sign to text conversion, sometimes referred to as sign language recognition, is to close the communication gap between the hearing-impaired or deaf community and the wider public. We may promote inclusion and successful communication for people who predominantly use sign language by translating hand signs into text or spoken language.

There are numerous methods for converting hand signs into text, from wearable technology to computer vision-based techniques. To recognise and decipher the signs, computer vision systems capture and examine video input of hand motions. These techniques frequently employ deep learning algorithms and neural networks to precisely classify and decipher the hand gestures. By recording depth information and finer details of hand movements, specialised cameras such as depth cameras or RGB-D sensors can improve the accuracy of these devices. When converting hand signs into text, recurrent neural networks (RNNs) of the LSTM (Long Short-Term Memory) kind are widely used. LSTM networks perform exceptionally well at capturing long-term dependencies in the input data when used for sequence-to-sequence modelling.

In the context of hand sign to text translation, LSTM networks can be trained to recognise and interpret the sequential hand gestures captured from video or sensor data. The network uses a series of hand sign representations, such as hand landmarks, joint angles, or image frames, to anticipate the appropriate textual representation.

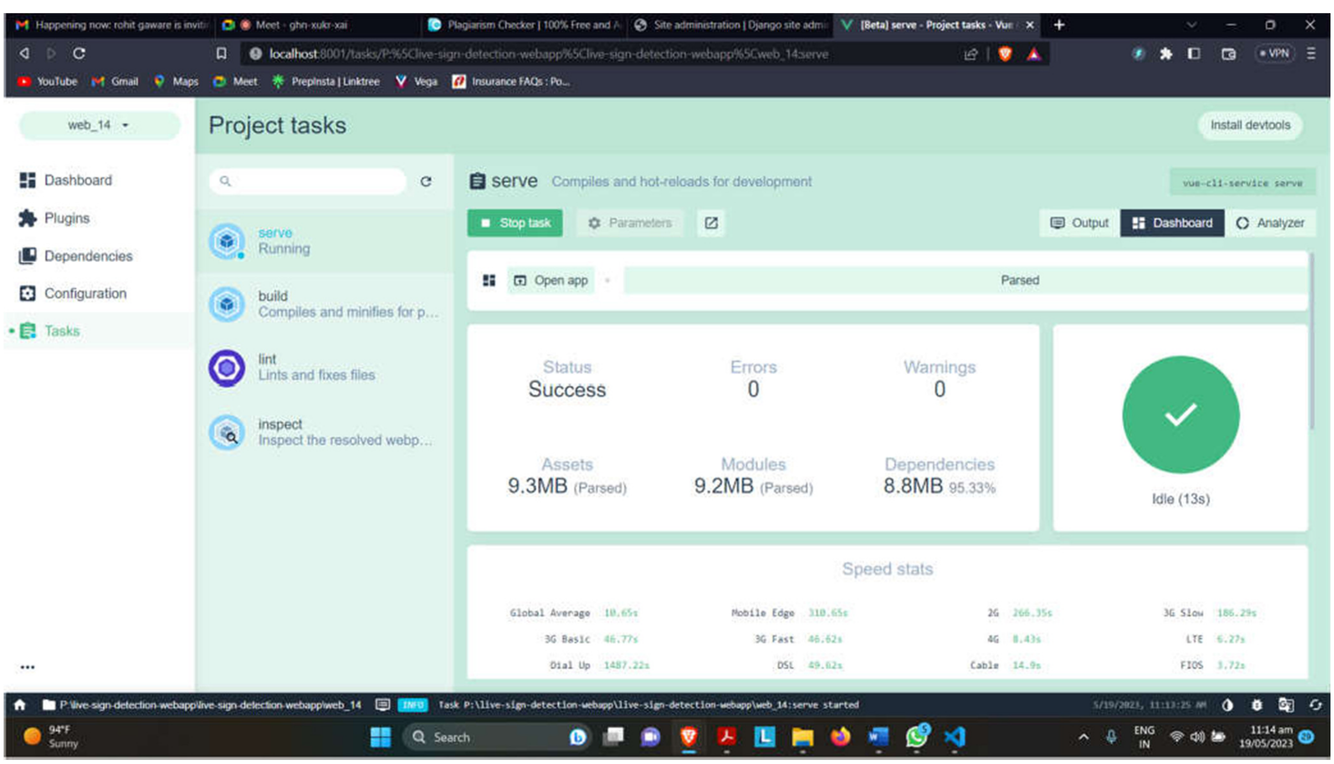
Because hand sign movements can vary in length, it is essential for hand sign recognition that LSTM networks can accommodate variable-length input sequences. To accurately recognise sign language, LSTM networks can capture the temporal dynamics and dependencies between various hand motions.

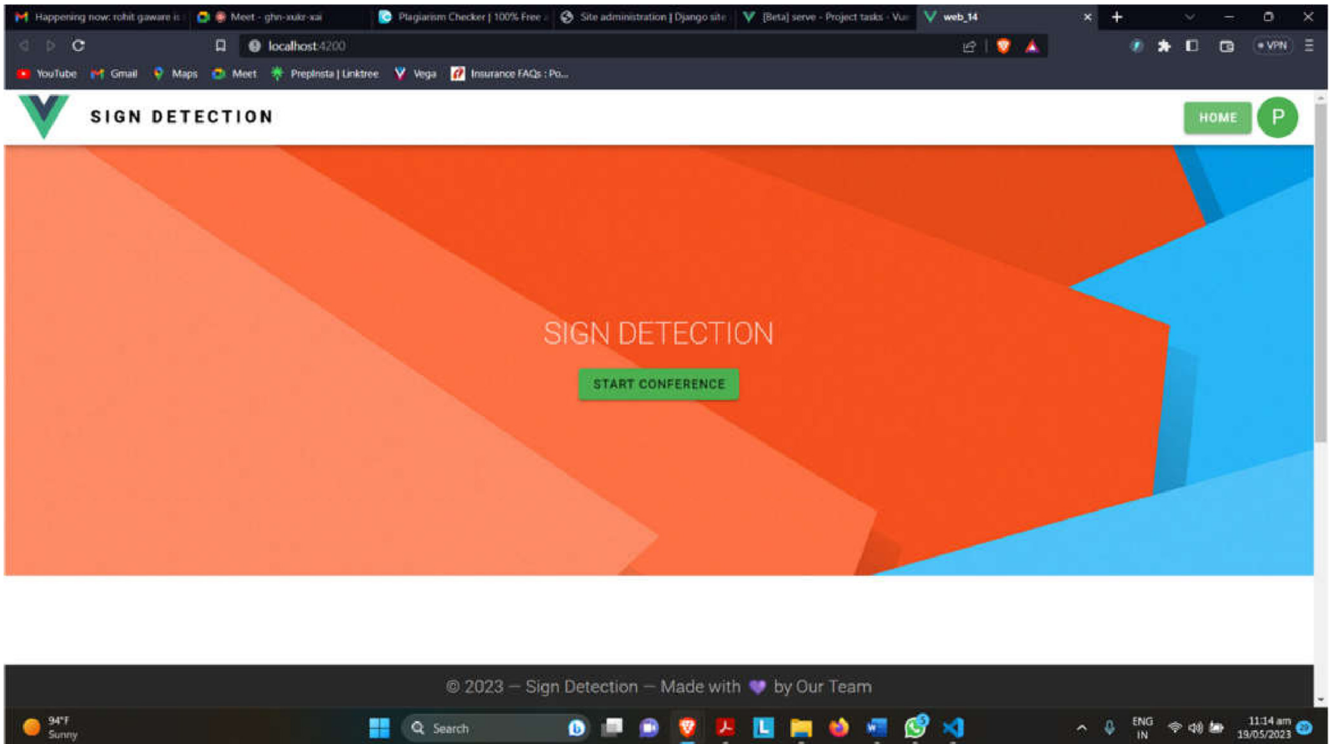
It's crucial to remember that the availability and calibre of the training data strongly influence how well an LSTM-based hand sign to text conversion system performs. Training strong models requires assembling a wide and representative dataset that includes a range of sign gestures and variations.

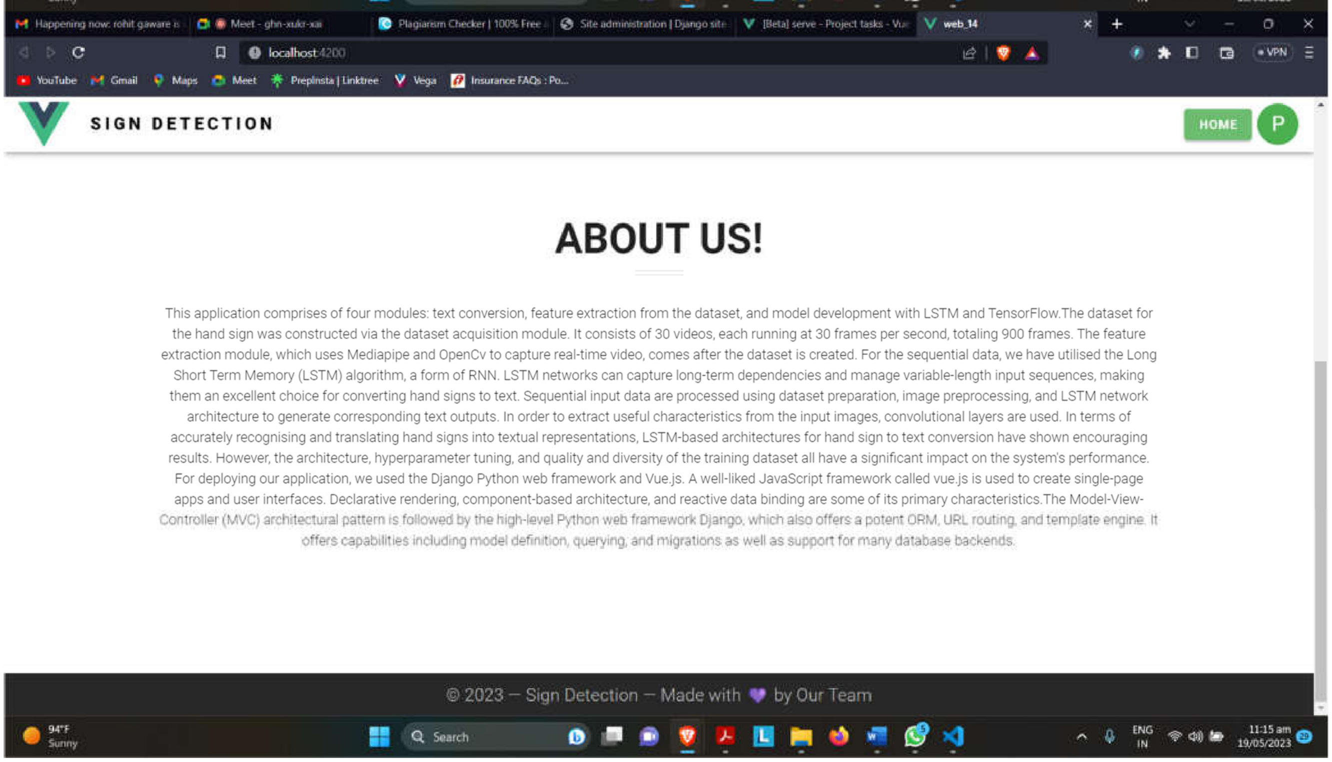
The accuracy and resilience of hand sign recognition systems can also be improved by merging LSTM networks with additional methods, such as computer vision algorithms for hand tracking or feature extraction.

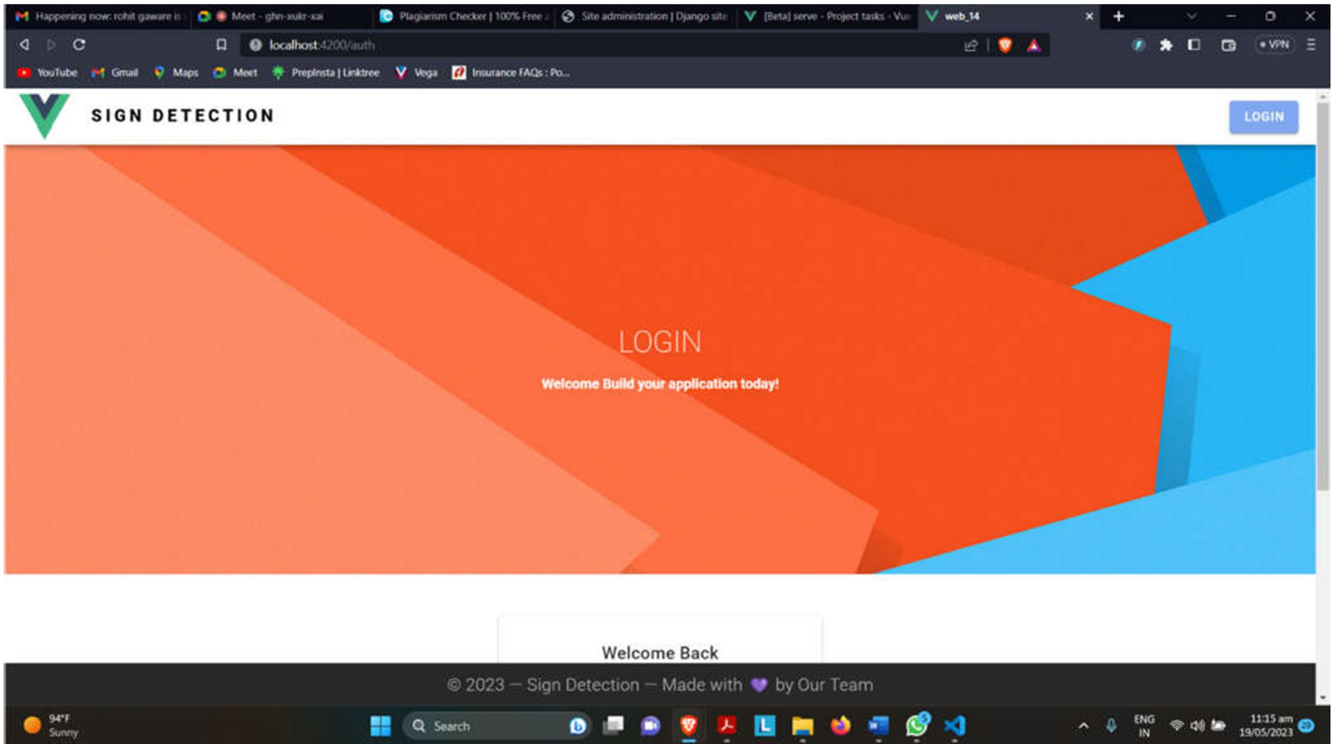
Using Django and sqlite as the database, we integrated the backend. We built a login and registration process into the Django Rest Framework for our online application. We have the vue.js framework for front-end use, which is useful for single-page applications and has the greatest react.js and angular.js feature combo. Real-time video is transformed to a concurrent jpeg format during the conversion process, and this jpeg file is then translated to byte code to make it easier to process.Using pyttsx3, we have combined the speech and text.

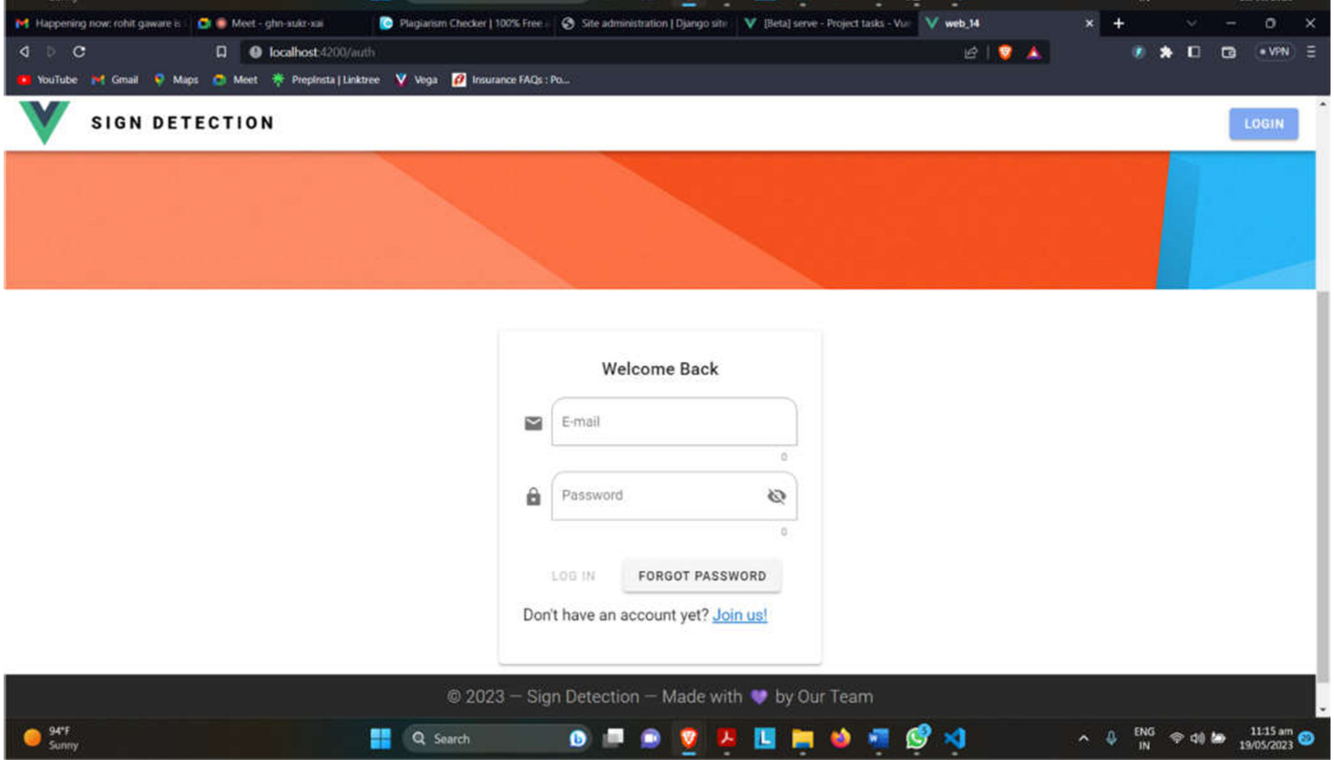
Output/Result:

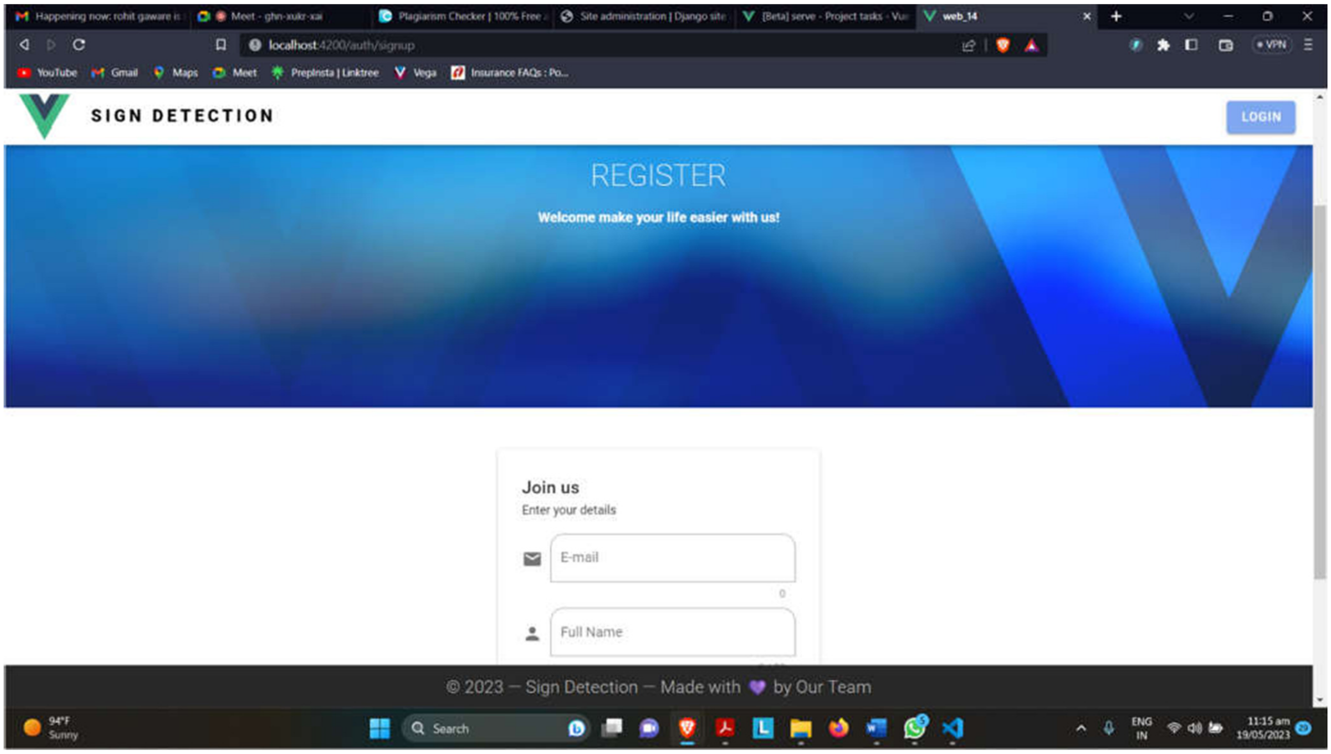


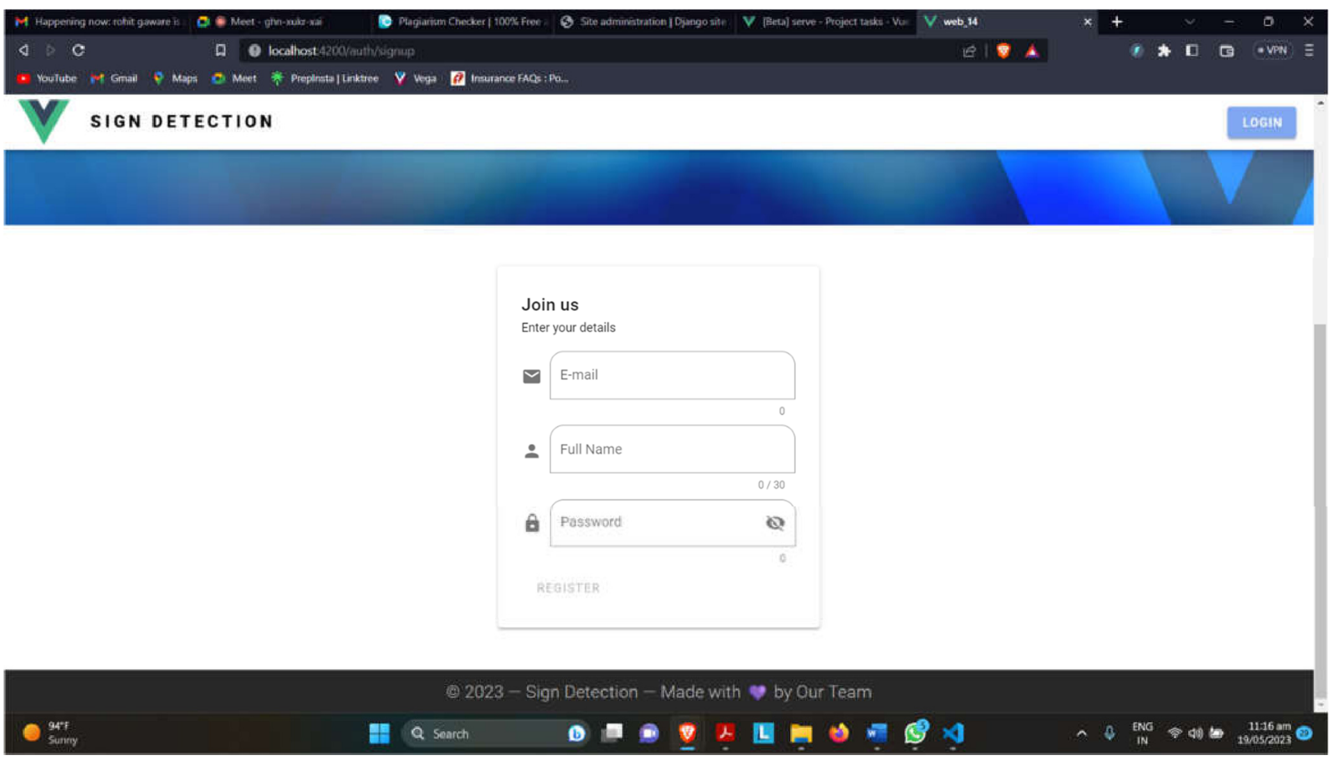

SIGN DETECTION
HOME
SIGN DETECTION
START CONFERENCE
© 2023 — Sign Detection — Made with by Our Team

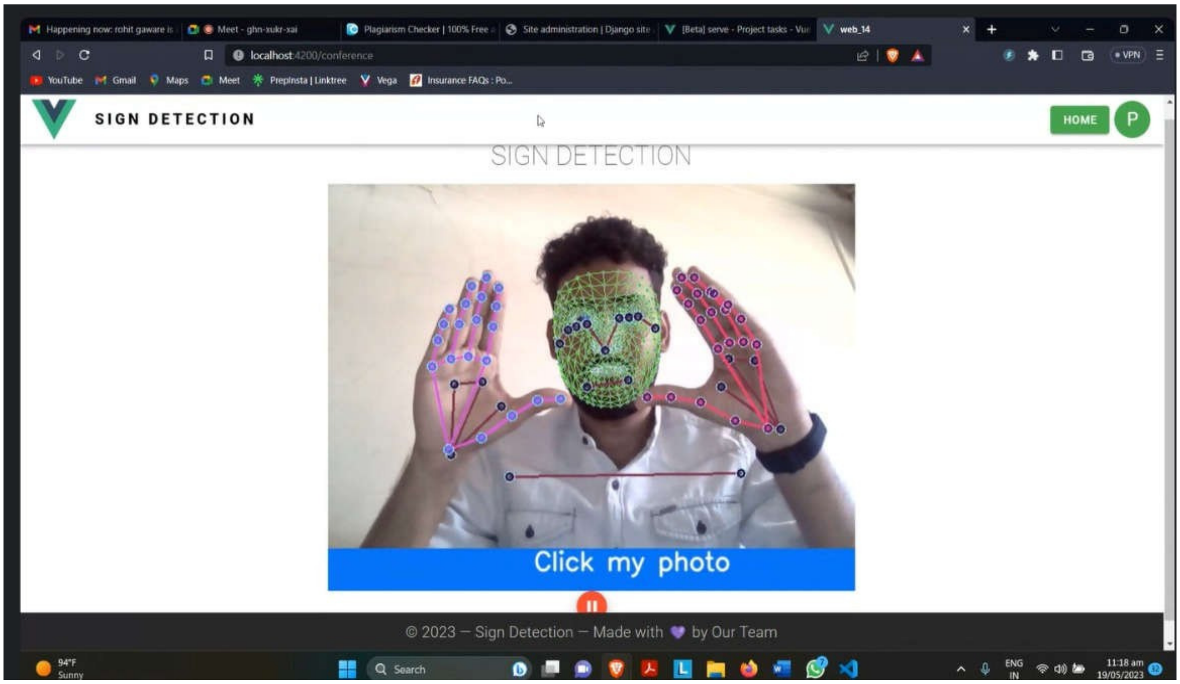

SIGN DETECTION
HOME
ABOUT US!
This application comprises of four modules: text conversion, feature extraction from the dataset, and model development with LSTM and TensorFlow.The dataset for the hand sign was constructed via the dataset acquisition module. It consists of 30 videos, each running at 30 frames per second, totaling 900 frames. The feature extraction module, which uses Mediapipe and OpenCv to capture real-time video, comes after the dataset is created. For the sequential data, we have utilised the Long Short Term Memory (LSTM) algorithm, a form of RNN. LSTM networks can capture long-term dependencies and manage variable-length input sequences, making them an excellent choice for converting hand signs to text. Sequential input data are processed using dataset preparation, image preprocessing, and LSTM network architecture to generate corresponding text outputs. In order to extract useful characteristics from the input images, convolutional layers are used. In terms of accurately recognising and translating hand signs into textual representations, LSTM-based architectures for hand sign to text conversion have shown encouraging results. However, the architecture, hyperparameter tuning, and quality and diversity of the training dataset all have a significant impact on the system's performance. For deploying our application, we used the Django Python web framework and Vue.js. A well-liked JavaScript framework called vue.js is used to create single-page apps and user interfaces. Declarative rendering, component-based architecture, and reactive data binding are some of its primary characteristics.The Model-View-Controller (MVC) architectural pattern is followed by the high-level Python web framework Django, which also offers a potent ORM, URL routing, and template engine. It offers capabilities including model definition, querying, and migrations as well as support for many database backends.
© 2023 — Sign Detection — Made with by Our Team

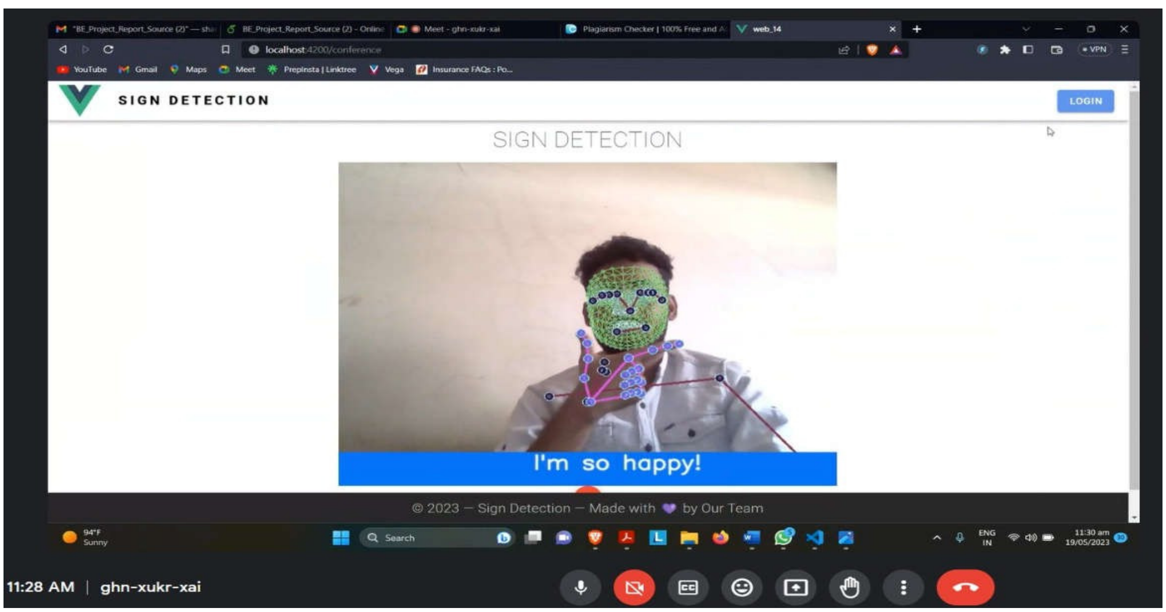

SIGN DETECTION
LOGIN
LOGIN
Welcome Build your application today!
Welcome Back
© 2023 — Sign Detection — Made with by Our Team

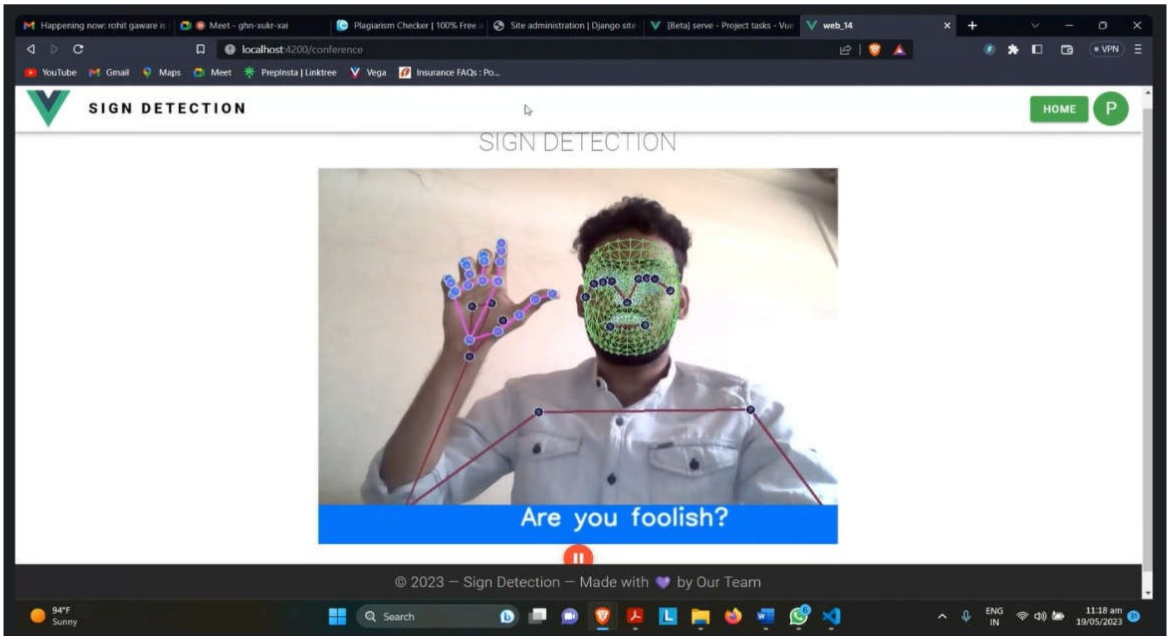

SIGN DETECTION
LOGIN
Welcome Back
E-mail
Password
LOG IN
FORGOT PASSWORD
Don't have an account yet? Join us!
© 2023 — Sign Detection — Made with by Our Team


SIGN DETECTION
LOGIN
REGISTER
Welcome make your life easier with us!
Join us
Enter your details
E-mail
Full Name
© 2023 — Sign Detection — Made with by Our Team


SIGN DETECTION
LOGIN
Join us
Enter your details
E-mail
Full Name
0 / 30
Password
REGISTER
© 2023 — Sign Detection — Made with by Our Team


SIGN DETECTION
HOME
SIGN DETECTION
Click my photo
© 2023 — Sign Detection — Made with by Our Team


SIGN DETECTION
LOGIN
SIGN DETECTION
I'm so happy!
© 2023 — Sign Detection — Made with by Our Team
11:28 AM | ghn-xukr-xai




## CONCLUSION

The interface solution's practical applicability for dumb and deaf individuals is constrained by its simplicity and usability in real-world situations. In this method, hand gesture to text conversion has been employed to aid the reduction of hardware components as an easy and practical technique to create human-computer interaction. The folks are able to communicate with one another with ease. People may use the system easily and without any complexity thanks to its user-friendly design.The programme saves money and does away with the need for pricey technologies

## REFERENCES

1] Ohnmar Win , ”Hand Gesture to Text and Speech Conversion”,2018

[2] Prof. Radha S. Shirbhate, Mr. Vedant D. Shinde , Ms. Sanam A. Metkari, Ms. Pooja U. Borkar, Ms. Mayuri A. Khandge,”Sign language Recognition Using Machine Learning Algorithm,2020

[3] Sanil Jain and K.V.Sameer Raja, ”Indian Sign Language Gesture recognition” , 2015

[4] Ankit Ojha, Ayush Pandey, Shubham Maurya, Abhishek Thakur, Dr. Dayananda P,”Sign Language to Text and Speech Translation in Real Time Using Convolutional Neural Network ”,2020

[5] Uday Khati, Prajitesh Singh, Achyut Shankar ,”Text Generation through Hand Gesture Recognition”,2020

[6] Impaired R.Priyakanth, N.M.Sai Krishna, Radha Abburi , ”Hand Gesture Recognition and Voice Conversion for Speech”,2020

[7] K. Manikandan, Ayush Patidar, Pallav Walia, Aneek Barman Roy,” Hand Gesture Detection and Conversion to Speech and Text,” 2019

.