An Efficient Approach For Interpretation Of Indian Sign Language Using Machine Learning

[1]Dr.Sneha Jondhle, [2]Mr.Harsh Jondhle, [3]Ms. Laxmi Ramkrishna Sahu, [4]Ms. Swati Suhas Rodge, [5]Mr.Parth Kirti Panchal.

[1]Associate Prof, [2]Asst.Prof,[3,4,5]UG Student,[1,2]EXTC Engg, [3,4,5]Computer Engg.

[1]Shivajirao S. Jondhale College of Engineering, Dombivli, [2,3,4,5]Shivajirao S. Jondhle College of Engineering & Technology, Asangaon, Maharashtra, India.

[1]snehajondhle@gmail.com, [2]harshjondhle@gmail.com, [3]sahulaxmi2435@gmail.com, [4]swatirodge01@gmail.com, [5]parthpanchal1352003@gmail.com

Abstract - One way to communicate nonverbally is via sign language. Individuals with hearing impairments utilizing sign language as a means of communicating their emotions. However, because they are not familiar with the meaning of the sign language movements, individuals typically have trouble understanding the hand signals used by those with unique challenges. When a person with speech or hearing impairments want to speak with a regular person, a translator is typically required, and vice versa. This project proposes a technique for translating hand motions used in ISL for numbers (1–9), English alphabets (A–Z), and some English words to understandable text and vice versa, allowing people with disabilities to communicate with people more effectively[1].

Keywords- Machine Learning, Neural Network Classifiers, Gesture Recognition, Image Processing Techniques, CNN.

I. INTRODUCTION

Sign Languages are not universal. There are about 300 distinct use of sign languages throughout the globe. This is the case because people from various ethnic origins inherently developed sign languages. It's possible that India lacks a typical sign language. Various parts of India have distinct dialects and lexical variations of Indian Sign Language. However, there have been recent attempts to harmonize ISL. ISL hand gestures can be categorized into two main groups: static gestures and dynamic gestures[1].

Approximately 50 lakh individuals in India have speech or hearing impairments, according to the 2011 census. However, the number of highly qualified translator for sign language in india is fewer than 300. Because they have trouble interacting with other regular people, people with speech or hearing impairments often experience feelings of loneliness and isolation. Their social and professional lives are greatly impacted by this. In light of the difficulties that people with certain requirements encounter, an automated real-time system that can translate words from English into ISL and vice versa has been proposed in this paper. For those with special needs, this system facilitates effective communication with the outside world[1].

II. AIMS AND OBJECTIVE

a) Aim

The paper's objective is to recognize Indian Sign Language alphabets based on the appropriate gesture. While American Sign Language has extensively studied comprehension of sign language and gestures, its Indian counterpart has hardly touched on these subjects. Our approach to this problem is to recognize objects from images (which can be obtained from, say, a webcam) and then use machine learning and computer vision techniques to retrieve pertinent characteristics and perform subsequent classification. Rather than utilizing expensive technology like gloves or Kinect for gesture recognition[9].

b) Objective

An ISL reports main goals are to: preserve and promote ISL as an essential language and cultural legacy; improve accessibility and inclusion for Deaf people; integrate ISL into education at all levels; grow the ISL teachers and interpreters' capacity; push for government support and recognition; empower the Deaf community through leadership and cultural expression; increase public awareness and sensitization; expand ISL interpretation services across various domains; encourage partnerships and collaborations; track and evaluate project impact; use technology to enhance ISL learning; plan for long-term sustainability; and engage in international cooperation to advance ISL and Deaf culture on a global basis.

 DOI : 10.35291/2454-9150.2024.0268

III. LITERATURE SURVEY

Paper 1: A depth based ISL Recognition using Microsoft Kinect.

It is now crucial for a system to recognize sign language in order to close the communication gap between the hearing and speech impaired and the able-bodied. In this paper, an effective algorithm for converting meaningful English text and speech from an input hand gesture in Indian Sign Language (ISL) is introduced. Microsoft Kinect (preferred because its performance is independent of ambient light and object color) is used by the system to record hand gestures. The dataset used is made up of 4600 depth and RGB images (taken with a Kinect Xbox 360) with 140 distinct ISL gestures from 21 subjects. These gestures include fingerspelling (signs for letters and numbers) and double-handed signs[5].

Paper 2: Hand Gesture Recognition Software based on ISL

A great way to communicate with communities of people who have intellectual disabilities is through hand gestures. It helps establish connections between computers and people. This system's potential for expansion is evident in public settings where deaf individuals use sign language to speak with those who are able to hear.. This article presents a method for consistently identifying gestures using ISL, wherein each gesture is performed with both hands. The task of recognising gestures remains challenging. To resolve this issue downloading the key is required. These essential pointers are helpful for removing incompatible frameworks and decomposing Include sign language movements in the sequence of characters[2].

Paper 3: Real-time Indian Sign Language (ISL) Recognition.

In this work, a real-time system with grid-based characteristics is presented to identify hand poses and gestures from ISL. The goal of this system is to improve communication between the general public and people who are deaf or or speech. The current methods are either not real-time or offer comparatively poor accuracy. Excellent results are obtained from this system for both parameters. 33 hand positions in addition to a few ISL gestures can be recognized by it. A smartphone's video captures sign language, which is then sent to a remote server for handling into frames. It's user-friendly because it doesn't require the application gloves or the Microsoft Kinect sensor, among other external devices[4].

Paper 4: Real time-Employee Emotion Detection system (RTEED) using Machine Learning.

Employee health and well-being are now the most important considerations in the workplace. Because it will have an impact on both an employee's productivity and the contribution of the team. In the end, automatic facial expression analysis using machine learning has developed over the past few decades into an intriguing and busy study subject. In order to automatically identify employee emotions in real time using machine learning, the Real Time Employee Emotion Detection System (RTEED) has been proposed in this work. The RTEED system assists the company in monitoring employee wellbeing, and any identified emotions are sent to the appropriate employee through messages. Employees are able to make better judgements, focus better at work, adopt bealthier lifestyles, and work in more productive ways as a result. Multi-PIE at CMU Machine learning models are trained using face data. Each employee will have a webcam that can record their real-time facial expression. The RTEED system is intended to recognize sis different emotions through the captured image, including happy, sadness, surprise, fear, disgust, and rage[10].

IV. EXISTING SYSTEM

There are two primary categories of ISL hand gestures: (i) static gestures and (ii) dynamic gestures. The ISL hand gestures for the numbers 1–9, the English alphabet (A–Z), and certain English words are static. Two primary functions are carried out by the suggested system: Conversion from gesture to text (i) and from speech to gesture (ii). Classifiers of neural networks are used for the translation of motions into words. The Google Speech Recognition Speech to gesture conversion is done via API. In this effort, spoken English words are accurately translated into ISL gestures and conventional ISL gestures are translated into English. To do this, different classifiers for neural networks are developed and tested for their ability to recognize gestures[1].

V. COMPARATIVE STUDY

Sr. No.	Author	Project Title	Publication	Technology	Purpose
1.	T Raghuveera, R Deepthi, R Mangalashri and R Akshaya	A depth based ISL Recognition using Microsoft Kinect.	ScienceDirect,2018	Support Vector Machine	The system combines the three SVM-trained feature classifiers to increase the average recognition accuracy to 71.85%.
2.	Mr. Sanket Kadam, Mr. Aakash Ghodke, Prof. Sumitra Sadhukhan	Hand Gesture Recognition Software based on ISL	IEEE,2019	Gesture Recognition Technique	The model where Deaf individuals are sending messages to regular people through communication..
3.	Kartik Shenoy, Tejas Dastane, Varun Rao, Devendra Vyavaharkar.	Real-time Indian Sign Language (ISL) Recognition.	IEEE,2021	k-Nearest Neighbors algorithm	It can identify 33 ISL hand postures. as well as a few gestures.

DOI : 10.35291/2454-9150.2024.0268

VI. PROBLEM STATEMENT

A lot of technical issues occurred during the course of the procedure for developing new products while researching the project-related details. Capturing hand and finger gestures was a significant challenge faced while conducting the research. to avoid using gloves or any other sensors and only use a video camera. Using suitable training methods and focusing on real-time video gesture identification [3].

VII. PROPOSED SYSTEM

One can use the SURF as a feature detector or as a feature descriptor. Applications like object identification and categorization of images are among the many uses for it. It is a reliable and quick algorithm for comparing and representing images. It functions within an image as a blob detector. Using the determinants of Hessian matrices, the SURF features are computed by locating the image's interest points that contain the significant features. The scale invariant descriptors are created for every interest point that was discovered in the preceding procedure.

VIII. ALGORITHM

Step 1: Start

Step 2: User Input as Text

```
# Load a sample dataset
i = datasets.load()
X = data,y = i.target
# Split the dataset into a training and testing set
X_train, X_test ,y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
clf = SVC(kernel='linear')
# Train the classifier on the training data
clf.fit(X_train, y_train)
# Make predictions on the test data
y_pred = clf.predict(X_test)
```

Step 3: Comparing input with database

```
# Define the CNN model
model = keras.Sequential()
# Convolutional layers
model.add(activation='relu',layers.Conv2D(32, (3, 3),
        input_shape=(64, 64, 3)))
model.add(layers.MaxPooling2D((2, 2)))
model.add(layers.Conv2D(64, (3, 3), activation='relu'))
model.add(layers.MaxPooling2D((2, 2)))
model.add(layers.Conv2D(64, (3, 3), activation='relu'))
# Fully connected layers
model.add(layers.Flatten())
model.add(layers.Dense(64, activation='relu'))
model.add(layers.Dense(10)   # 10 output classes for classification
```

Step 4: Produce output as image

```
# Load an image
image = cv2.imread('your_image.jpg', cv2.IMREAD_GRAYSCALE)
# Initialize SURF detector
surf = cv2.xfeatures2d.SURF_create()
keypoints, descriptors = surf.detectAndCompute(image, None)
image_with_keypoints = cv2.drawKeypoints(image, keypoints, None, (0, 255, 0), 4)
# Display the image with keypoints
cv2.imshow('SURF Keypoints', image_with_keypoints)
cv2.waitKey(0)
cv2.destroyAllWindows()
```

Step 5: Conversion complete

```
# Compile the model
model.compile(optimizer='adam', loss=tf.keras.losses.SparseCategoricalCrossentropy(from_logits=True), metrics=['accuracy'])
# Train the model
model.fit(train_dataset, epochs=10)  # Replace train_dataset with your training data
# Evaluate the model
test_loss, test_acc = model.evaluate(test_dataset)
```

Step 6: Take input as Speech

Step 7: Conversion of Speech to Text

```
#Speech_to_text:
r = sr.Recognizer()
image_names = os.path.join(settings.MEDIA_ROOT, 'sign words')
dir_list = os.listdir(image_names)
dir_list = list(map(str.lower, dir_list))
lst = [os.path.splitext(x)[0] for x in dir_list]
print(lst)
def SpeakText(command):
root = pyttsx3.init()
root.say(command)
```

 DOI : 10.35291/2454-9150.2024.0268

```
root.runAndWait()
SpeakText("Now Please Speak Something")
try:
with sr.Microphone() as source:
r.adjust_for_ambient_noise(source, duration=0.2)
audio2 = r.listen(source)
MyInput = r.recognize_google(audio2)
MyInput = MyInput.lower()
print("Did you say ", MyInput)
```

Step 8: Proceed to Step 3 , Step 4 , Step5

```
SpeakText(MyInput)
if MyInput in lst:
detected_sign = MyInput
else:
detected_sign = 'thank you'
# p =os.path.join(image_names, detected_sign+".jpg")
return render(request, 'users/speaknow.html', {'path':
detected_sign+".jpg"})
```

Step 9: Stop

IX. MATHEMATICAL MODEL

1. SVM

Support Vector Machines classify data by evaluating a hyperplane that raises the dividing line between classes in the training data. Hyperplane can be formulated as

$$f(x) = a^T x + c$$

Where, a = dimensional coefficient, c = offset

2. Logistic Regression

Logistic regression (LR) is typically used to assess the likelihood that an instance will fit a given class. Logistic regression describes the connection between one or more categorical outcomes and categorical predictors using a logit transformation of the dependent variable, where the logit model predicts the logit of the dependent variable from the independent variable [3].

$$\hat{p} = h_\theta(x) = \sigma(x^T\theta)$$

Where,

$$\sigma(t) = \frac{1}{1+e^{-t}}$$

3. K Nearest Neighbor

K-Nearest Neighbor (KNN) was a method which looks at K dataset instances close to the observation. The method itself then evaluates the inspection variable y to be anticipated using the output [8]. For calculating the distance of two observations, Euclidean distance is used, and the equation is as follows:

$$d(x_i,y_i) = \sqrt{(x_{i,1}-y_{i,1})^2 + \ldots + (x_{i,m} - y_{i,m})^2}$$

4. Decision Tree

The decision-tree method recognizes various outcomes, especially event consequences, using a graph-based framework. model trees target variables may have a discrete range of values, but in trees, the leaves stand in for class designations, and the branches for feature links that imply class labels. The equation of entropy is provided underneath,

$$E = -\sum_{b=1}^{n} p_{ab} \log_2 p_{ab}$$

X. SYSTEM ARCHITECTURE

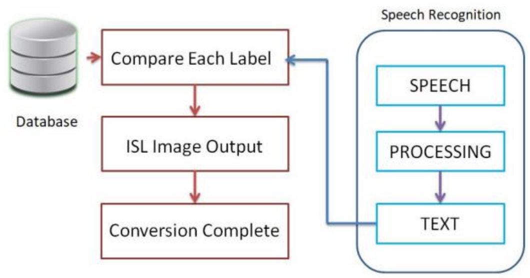

Fig.1: System Architecture

XI. ADVANTAGES

Accessibility: ISL provides a channel of communication for Deaf individuals in India who may not have access to spoken languages. It enables Deaf individuals to interact with each other.

Cultural Identity A crucial component of Indian Deaf culture is ISL. It allows Deaf individuals to interact alongside their contemporaries and express themselves culturally. ISL is a means of expression for culture as well as a channel of dialogue.

Education: ISL is used in Deaf education in India. It enables Deaf pupils to have access to information, participate in classroom discussions, and communicate with teachers and peers. This improves the quality of education for Deaf students.

Legal Recognition: In certain regions of India, ISL is now considered a recognized language. Legal recognition can aid in defending Deaf people's rights, such as their ability to receive services and education.

 DOI : 10.35291/2454-9150.2024.0268

XII. DESIGN DETAILS

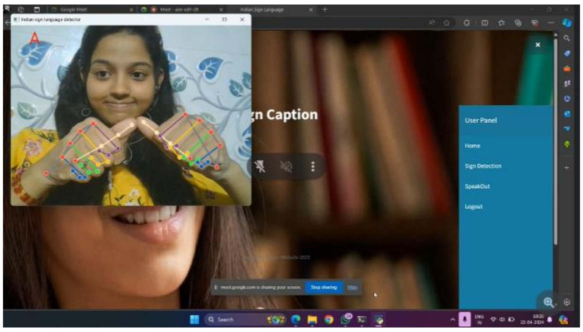

Fig 1:Sign Detection

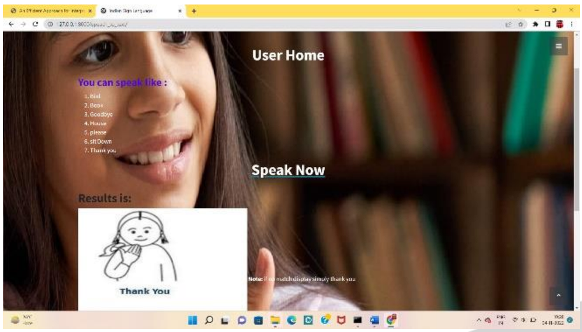

Fig 2:Speech To Text

XIII. CONCLUSION

Thus we have tried to implement the paper "An Efficient Approach for interpretation of Indian Sign Language using Machine Learning", Akash M; Sona M; Divyapriya S; Krishnaveni V, Dhivyasri S; Krishnaa Hari K B, IEEE, 2021. The conclusion as follow: From the results obtained, it is inferred that the SVM classifier along with the K-means clustering and BoV classifiers is best suited for gesture recognition. A user friendly application that can interpret Indian Sign Language has been developed using the most efficient SVM classifier (for gesture to text conversion) and Google Speech Recognition API (for speech to gesture conversion).

REFERENCE

[1] Dhivyasri S; Krishnaa Hari K B; Akash M; Sona M; Divyapriya S; Krishnaveni V, "An Efficient Approach for interpretation of Indian Sign Language using Machine Learning", in IEEE Xplore, 15 June 2021.

[2] Mr. Sanket Kadam, Mr. Aakash Ghodke, Prof. Sumitra Sadhukhan, "Hand Gesture Recognition Software based on ISL", IEEE Xplore, 20 June 2019.

[3] Muthu Mariappan H, Dr Gomathi V, "Real time Recognition of ISL", IEEE Xplore,10 October 2019.

[4] Kartik Shenoy, Tejas Dastane, Varun Rao, Devendra Vyavaharkar, " Real-time Indian Sign Language (ISL) Recognition", IEEE Xplore, 18 October 2018.

[5] T Raghuveera, R Deepthi, R Mangalashri And R Akshaya, "A depth based ISL Recognition using Microsoft Kinect", ScienceDirect, 2018.

[6] G. Ananth Rao a, P.V.V. Kishore, "Sign Language Recognition Based On Hand And Body Skeletal Data", IEEE 2018.

[7] G. Ananth Rao a, P.V.V. Kishore, "Selfie video based continuous Indian sign language recognition system", ScienceDirect, 2017.

[8] Cheok Ming Jin, Zaid Omar, Mohamed Hisham Jaward, "A Mobile Application of American Sign Language Translation via Image Processing Algorithms", in IEEE Region 10 Symposium, on IEEE explore, 2016.

[9] Sanil Jain and K.V. Sameer Raja, "Indian Sign Language Gesture recognition", IIT Kanpur March 16, 2015.

[10] Prof. Vishal R. Shinde, "Real time-Employee Emotion Detection system(RTEED)using Machine Learning", in IJREAM, ISSN : 2454-9150, Volume 09, Issue 01 Special Issue, April 2023.

 DOI : 10.35291/2454-9150.2024.0268