International Journal of Innovative Research in Computer Science & Technology (IJIRCST)
ISSN: 2347-5552, Volume-12, Special Issue-1, March-2024
https://doi.org/10.55524/CSISTW.2024.12.1.65
Computing for Sustainable Innovation: Shaping Tomorrow's World (CSISTW-2024) on 16th-17th February, 2024
Organized by B.N. College of Engineering and Technology, Lucknow, India

Review on VisioSense: Navigating the Landscape of IndianSign Language Detection and Recognition

Saurabh Mishra[1], Neha Singh[2], Mayank[3], Suraj Singh[4], Shubham Mishra[5]

[1,2,3,4] B.Tech Student, Department of CSE, B.N. College of Engineering & Technology, Lucknow, UP, India
[5]Assistant Professor, Department of CSE, B.N. College of Engineering & Technology, Lucknow, UP, India

Correspondence should be addressed to Saurabh Mishra; msaurabh2000@gmail.com

Copyright © 2024 Made Saurabh Mishra et al. This is an open-access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT: This paper introduces a system for instant recognition of Indian Sign Language (ISL) and gesture identification using grid-based features. Addressing communication barriers between hearing-impaired individuals and society, the system achieves high accuracy without external devices like gloves or Microsoft Kinect sensors. Leveraging face detection, object fixation, and skin color technologies for hand detection and tracking, the Laptop's camera captures ISL gestures. Grid-based feature extraction represents hand movements as feature vectors, classified through the k-nearest neighbor algorithm. Hand gesture classification employs hidden Markov models, achieving 99.7% accuracy for static tasks and 97.23% for orientation recognition.

KEYWORDS: ISL Recognition, Gesture Recognition, Grid-Based Feature Extraction, k-Nearest Neighbor, Hidden Markov Model.

I. INTRODUCTION

Indian Sign Language (ISL) is a language used by deaf and hard of hearing people to communicate with other people. The research presented in this article relates to ISL as defined on the Talking Hands website [1]. ISL uses gestures to represent complex words and sentences. It has 33 movements, 10 of which are numbers and 23 are letters. Among the ISL letters, the letters "h" and "j" are represented by hand gestures, while the letter "v" is similar to the number
2. The system is trained to use gestures from ISL as shown in Figure 1. Many people have difficulty understanding ISL movements. This creates a communication gap between those who understand ISL and those who do not. People can never find an interpreter to interpret gestures when necessary.

To facilitate this communication, a solution that can perform hand movements and gestures has been developed in ISL. It has a windows camera to capture movements and gestures and a server to process frames from the Laptop camera. The purpose of this system is to use fast and accurate tools.

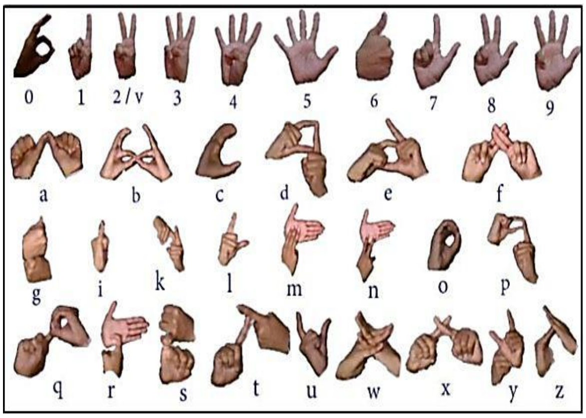

Figure 1: Hand poses in ISL

The framework portrayed in this paper effectively classifies all the 33 hand postures in ISL. For the beginning, motions containing as it were one hand were considered. The arrangement portrayed can be effortlessly expanded to two-handed signals. Within the following segment of this paper, the related work relating to sign dialect interpretation is talked about. Segment III clarifies the procedures utilized to prepare each outline and interpret a hand pose/gesture. Area IV talks about the test after actualizing the strategies talked about in area III. Area V portrays the windows application created for the framework that empowers real-time Sign Dialect Translation. Section VI talks about long term work that can be carried out in ISL interpretation.

II. RELATED WORK

Significant progress in Sign Dialect recognition, especially motion recognition, has utilized methods like gloves and Microsoft Kinect sensors. Our system, distinguished by its sensor-free approach, utilizes the Windows phone camera, minimizing costs and enhancing convenience. While accurate, costly solutions like Microsoft Kinect and SignAloud gloves lack portability. In software-based solutions, our approach avoids the need for specific gear, relying on the Windows phone camera. Utilizing hand anthropometry principles, our system localizes and eliminates the palm, employing grid-based features for user-invariant hand posture recognition inspired by a k-NN model. For motion recognition, hand centroid tracking provides data, and our system incorporates HMM Signal recognition, using a 9-element vector for input to the Gee classifier, trained through Baum-Welch re-estimation equations, inspired by previous work.

III. IMPLEMENTATION

Using a Laptop Camera, signals and signs performed by the individual utilizing ISL are captured and their outlines are transmitted to the server for preparation. To form the outlines prepared for acknowledgment of gestures and hand postures, they have to be pre-processed. The pre-processing to begin with includes confront evacuation, adjustment and skin color division to evacuate foundation points of interest and afterward morphology operations to decrease clamor. The hand of the individual is extricated and followed in each outline. For acknowledgment of hand postures, highlights are extricated from the hand and nourished into a classifier. The perceived hand posture lesson is sent back to the Laptop gadget. For classification of hand signals, the middle hand postures are perceived and utilizing these perceived poses and their halfway movement, a design is characterized which is spoken to in tuples. This is often encoded for Gee and nourished to it. The signal whose Gee chain gives the most elevated score with forward-backward calculation is decided to be the recognized signal for this design. An outline of this handle is portrayed in

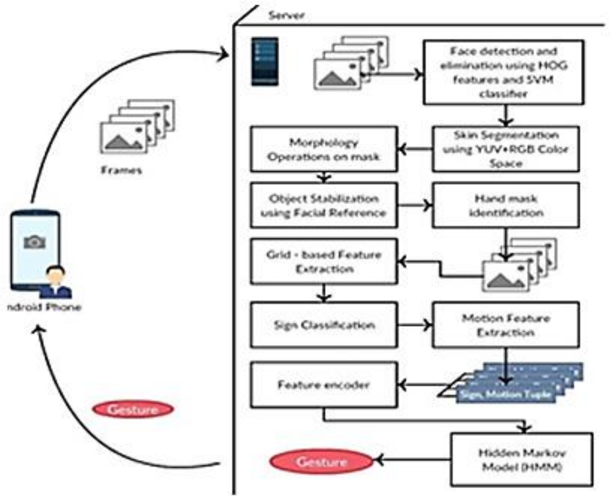

Figure 2: Flow Diagram for Gesture Recognition

A. Dataset Used

For the digits to 9 in ISL, a normal of 1450 pictures per digit were captured. For letters in ISL barring 'h', 'j' and 'v', around 300 pictures per letter were captured. For the 9 gesture-related middle of the road hand postures such as Thumbs_Up, Sun_Up, almost 500 pictures per posture were captured. The dataset contains an add up to 24,624 pictures. All the pictures consist of the sign demonstrator wearing a full sleeve shirt.Most of these pictures were captured from an standard webcam and a number of them were captured from a Laptop camera. The pictures are of changing resolutions. For preparing HMMs, 15 motion recordings were captured for each of the 12 one-handed pre-selected gestures characterized in [1] (After, All The Leading, Apple, Great Evening, Great Morning, Good$80 < U < 130 136 < V < 200 V > U R > 80 \,\&\, G > 30 \,\&\, B > 15$I $|R - G| > 15$(2)Night, I Am Too Bad, Pioneer, It would be ideal if you Allow Me Your Write, Strike, That's Good, Towards). These recordings have slight varieties in arrangements of hand postures and hand motion so as to form the HMMs strong. These recordings were captured from a Laptop camera conjointly including the sign demonstrator wearing a full sleeve shirt.

B. Pre-Processing

- **Face Detection and Elimination:** ISL hand poses are crucial, and facial features are unnecessary. Face detection using Histogram of Oriented Gradients (HOG) descriptors and a linear SVM classifier resolves issues during hand extraction. Utilizing an image pyramid and sliding window, this method outperforms Haar wavelet-based detectors, significantly reducing false positive rates. After face detection, the face contour is identified, and the face-neck region is blackened out, as shown in Figure 3.

Figure 3: Face Detection and Elimination Operation

- **Skin Color Segmentation:** To identify skin-like regions in the image, the YUV and RGB based skin color segmentation is used, which provides great results. This model has been chosen since it gives the best results among the options considered: HSV, YCbCr, RGB, YIQ, YUV and few pairs of these color spaces [14]. The frame is converted from RGB to YUV color space using the equation mentioned in [22]. This is specified in equation (1).

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} +0.299 & +0.587 & +0.114 \\ -0.147 & -0.289 & +0.436 \\ +0.615 & -0.515 & -0.100 \end{pmatrix} \cdot \begin{pmatrix} R \\ G \\ B \end{pmatrix} + \begin{pmatrix} 0 \\ 128 \\ 128 \end{pmatrix} \quad (1)$$

[14] notices the criteria for deciding on the off chance that pixel is to be classified as skin utilizing the taking after conditions:

$$\begin{cases} 80 < U < 130 \\ 136 < V < 200 \\ V > U \\ R > 80 \,\&\, G > 30 \,\&\, B > 15 \\ |R - G| > 15 \end{cases} \quad (2)$$

The division cover hence contains less noise and less wrong positives come about. An outline of the segmentation cover appears in Figure 4.

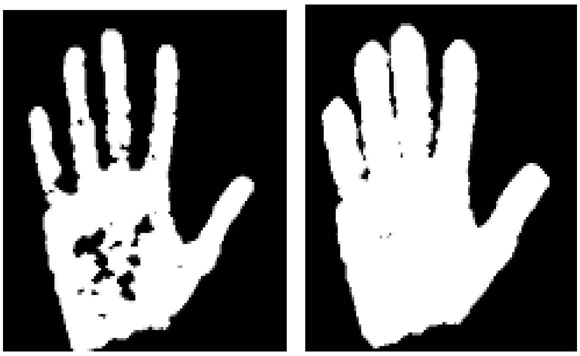

Figure 4: Skin segmentation mask and effect of morphology operations. (Left) Segmentation mask; (Right) Mask after application of Morphology operations

- **Morphology Operations:** Morphology operations were performed to evacuate any noise produced after skin color division. There are 2 types of mistakes in skin color segmentation:
  - Non-skin pixels classified as skin
  - Skin pixels classified as non-skin

Morphology includes 2 essential sub-operations:
Disintegration: Here, the dynamic zones within the cover (which are white) are diminished in size
Enlargement: Here, the dynamic zones within the cover (which are white) are expanded in size Morphology Open operation handles the 1st mistake. It Involves disintegration taken after by widening. The 2nd mistake is handled by the Morphology Near operation. This involves dilation taken after by disintegration. The result of applying morphology operations can be seen in Figure 4.

- **Use a Facial Mask to Obtain a Stable Product:** There must be a fixed camera position to follow the hands accurately. It's one thing for the camera to move while waving. If the cue announcer, for example an ISL user, does not move his hand but the cameraman holds his hand, the movement will be detected. This problem can be solved by using a fixed object. Assuming that the face is always included in the video motion, the character of the presenter's face is permanently tracked manually. The tracker started using coordinates obtained from face detection before removing the face. The tracker detects the position of the face bit and moves all frames in the direction of the detected face joint. The system uses the Kernelized Correlation Filter (KCF) tracker used in the OpenCV library to track faces in each frame. Tracking is done on the image before the face is darkened. Figure 3.4. shows the constant product.

Figure 5: (a) Actual movement of the camera. The position of the object relative to the frame changes. (b) Fixed effect – the position of the subject relative to the frame remains unchanged.

C. Hand Extraction and Tracking

Hand extraction and tracking are vital in sign language, representing gestures. After frame preprocessing, a black and white image isolates skin, excluding the face. Since frames feature either one visible hand or two touching, the primary structure is the human hand. Calculating contour areas identifies the hand, removing the largest contour, often the face. Face extraction is crucial; without it, the face dominates the image area, requiring manual division. as shown in Figure 6.

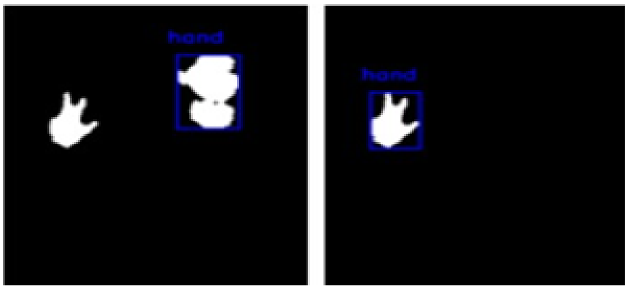

Figure 6: Importance of removing the face before hand extraction

To keep track of hands, the average hand of the group is calculated in each frame. If the hand moves, the coordinate of the center of mass of the hand will change. Then determine the slope of the line formed by the average cell size in the current frame and the average cell size in the previous frame. According to the slope value, the movement of the hand is determined as follows:
If x -1 < slope < 1 and if the difference between the x coordinates of the two is good, move the hand to the left.
If < br >x -1 < down < 1 and the difference between the x coordinates of the two centers of mass is negative, the hand moves to the right.
If x Slope > 1 and the difference between the y coordinates of the two centers of mass is positive, the hand moves forward.
If x Slope > 1 and the difference between the y coordinates of two magnitude points is negative, the hand moves downward.
Figure 7 illustrates the determination of the slope for hand movement tracking. Notably, the camera-recorded hand movement may differ from the actual presenter's actions. The system utilizes the OpenCV library to calculate contour areas and cell contour centroids. To mitigate noise during hand tracking, a 20-pixel radius imaginary circle is established around the prior hand center. If the new cluster location falls within this radius, it's deemed noise, and the movement is disregarded. When the average size surpasses the 20-pixel threshold, it signifies a hand gesture. The radius then reduces to 7 pixels until motion ceases, restoring to 20 pixels, effectively minimizing noise and ensuring precise hand movement tracking.

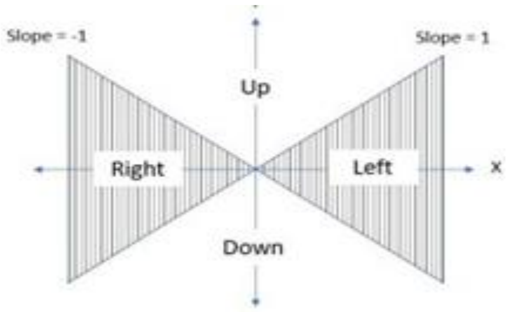

Figure 7: Determining Hand Motion using Slope

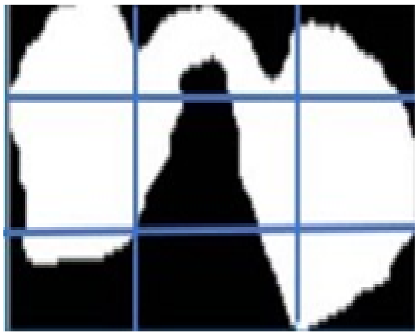

Figure 8: The Hand Pose ‘a’ in ISL Fragmented by a 3x3 Grid

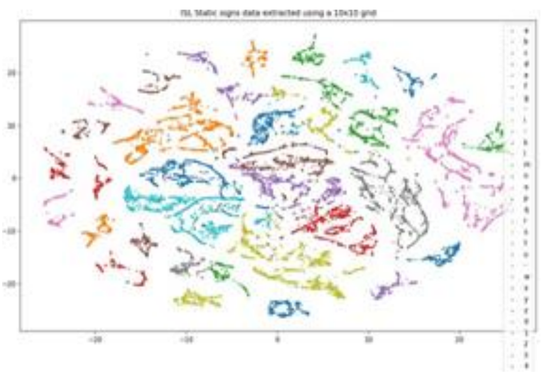

Figure 9: ISL Hand Poses' Data Visualized Using PCA and TSNE Dimensionality Reduction

D. Feature Extraction Using Grid-Based Segmentation Technique

Using M x N grid, the extracted hand test is divided intoM * N blocks. Using this grid, an eigenvector with M * N eigenvalues can be obtained; where each block returns an eigenvalue. In each block, feature values are calculated as a percentage of existing contours. This situation is stated in equation (3).

$$\text{Feature value} = \frac{\text{Area of hand contour}}{\text{Area of fragment}} \quad (3)$$

In the absence of a contour, the property value defaults to 0. Figure 8 illustrates a 3x3 grid created from the model. This method offers versatility as design features vary with hand orientation, mesh, and part size. The feature vector precisely represents hand shape and position. Employing M*N features ensure each move is distinctly represented by a different group. Figure 3.8 exemplifies this using a 10x10 grid with 100 features per sample. Dimensionality reduction from 100 to 2 is achieved through PCA and t-SNE, reducing width to 40 features with PCA, and further to 2 with t-SNE. Figure 3.8 visually demonstrates distinct groups representing different hand paths, and indicators emerge upon analyzing the extracted features.

E. Classification

- ***Using K-Nn to Identify Isl Hand Gestures:*** Looking at the diagram in Figure 9, it can be seen that the data is organized in groups. There are many groups for the same hand pose. For classification, an algorithm that can distinguish different types of data is needed. K-nearest neighbor (k-NN) was found to be suitable for this data classification. Manual extraction from each frame of the current feed must be extracted using the grid-based segmentation discussed previously. Then the pattern M*N map is represented by the distance measurement. Distance measurement can be done using brute force where the Euclidean distance between the sample and each sample fits the included class and the smallest k distance is selected. Other top options include KD tree and Package tree. The best way to calculate distance depends on the size of the file. Brute force works well on small files[24]. For low-dimensional data, KD Tree works well, while for high-dimensional data, Tree works best [24]. The classifier selects the most frequently occurring classes among k nearest neighbors.
- **Gesture classifications using HMM:** Some sort of standard is needed to cope with this change. Hidden Markov Model (HMM) is a statistical model that can be similar to variations [25]. There are two types of HMM: Continuous HMM and Discrete HMM. In continuous HMM, the number of observation signals at each point of the observation range is infinite, but in discrete HMM it is finite. HMM tuaj yeem hla los yog sab laug mus rau sab xis. In a left-to-right HMM, changes can only occur in 1 direction.g. If the HMM moves to the next state, it cannot return to the previous state, as shown in Figure 10. However, in crossing HMM, one can move from one state to another. The initial state probability (π) and transition probability of the HMM from left to right are shown in figure 10.

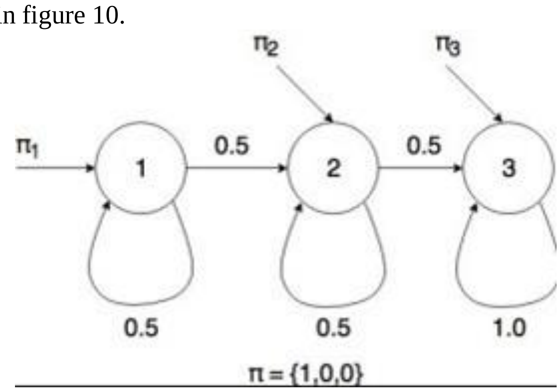

Figure 10: HMM chain for gestures with 3 hidden states (e.g. "Good day")

The human brain perceives gestures as a combination of some intermediate hand postures and hand movements performed in a certain order. ISL gestures that use this concept include tweens and hand gestures. Therefore, a discrete left-to-right HMM is used in this system. This uses segmented hand center of gravity movement and proportions the movements to classify the given analysis as belonging to one of 12 predefined movements.

The idea of this HMM is to analyze a sequence taken from a video. The number of tracking signals is the sum of the number of tracking directions and the average hand poses studied on the system. In this process, 4 gestures are tracked (as described in Section III C) and the system is trained using 9 average gestures such as “Thumb_Up”, “Sun_Up”, “Fist” as shown in Figure 11.

Figure 11:Thumbs_Up and Sun_Up writing hand poses

Motion tween recognition also uses grid-based feature extraction. Intermediate gesture recognition is similar to gesture recognition used to recognize letters and numbers, but only occurs in the absence of motion. Therefore, the total number of visible symbols is 13, meaning that periodic analysis can have points with values 0-12. At each frame, a tuple is created with the following format: M>; where M represents the movement of the cell relative to the previous frame and S represents an indicator when no movement was detected. If it is detected, the movement will be matched with the corresponding detection symbol ("up": 0, "right": 1,"left": 2, "down": 3). If no movement is found, hand poses are drawn accordingly. This map is predetermined. Therefore, there will be a tracking symbol for each frame, and the video sequence will form a tracking sequence that encodes gestures and facial expressions. For example, time series – [None>, None>, None>, Up>, Up>, Up> , , Up>, None>, None>, None>] Represents the "Good Afternoon" indicator. After encoding this as a sequence of observations, we get [4, 4, 4, 0, 0, 0, 5, 5, 5]. Several graphs of this indicator are shown in Figure 12.

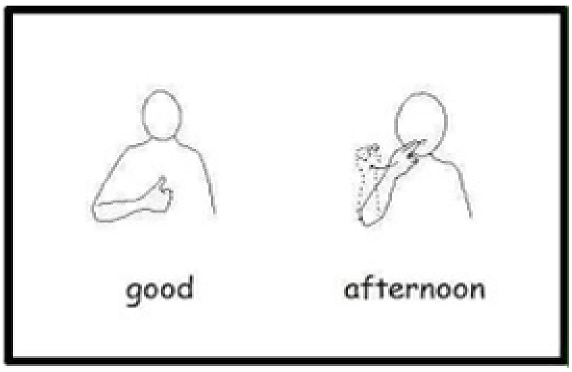

Figure 12: Frames of Gesture “Good Afternoon”

Gesture recognition in this system employs a chain of 12 Hidden Markov Models (HMMs), one for each gesture. The number of hidden states in each chain is determined by categorizing gestures into connected hand movements. For instance, the "Good After" gesture in Figure 12 is segmented into three states: "Thumb_Up," "Up," and "Sun_Up." All HMM chains mirror a left-hand-initial state probability and a right-hand HMM with similar initial-state transition probability, as illustrated in Figure 1.10. The state transition probability is of order n x n, and the emission probability matrix is of order n x 13 for n hidden states. The emission probability matrix begins with empirically determined probabilities, enhancing the chances of the HMM model converging to the maximum likelihood after training. Following initialization, Baum-Welch algorithm trains the HMM's estimated probability and state transition probability, while forward-backward algorithm determines the highest score for new observations in the trained HMM chains.

- **During Segmentation:** The motion recognition module should only provide video tracks corresponding to movements. Continuous movements are not possible without time division. We use the rule that if the hand leaves the frame, it will mark the end of the current direction and the direction recognition process will be completed on the current system taken from the orientation. When you get it again When entered into the frame, the beginning of the new movement will be marked. Use this rule to spend time apart.

IV. TEST RESULTS

The results discussed in this section are obtained from a computer with 8 GB RAM, Intel i7 processor with 4 virtual CPUs and 2 GB NVIDIA GPU. The operating system used is Ubuntu 16.04.

To better classify hand poses, the appropriate network size must be determined for hand segmentation and feature extraction. We use 6 different grid sizes to extract features from the same training data: 5x5,10x10, 10x15, 15x15, 15x20 and 20x20. These features are then placed into the k-NN classifier. Features are then extracted from the test equipment using a large grid. Features extracted from the test material are classified by a pre-trained k-NN classifier. An ideal network size will create different groups of motions so that the k-NN classifier can recognize them with high accuracy. To determine the best grid size, the accuracy of learning the k-NN classifier is calculated and compared with the test data. Choose the correct one according to equation (4):

$$Accuracy = \frac{Number\ of\ samples\ classified\ correctly}{Total\ number\ of\ samples} \qquad (4)$$

As can be seen from the figure 10x10 grid size has the highest accuracy with 99.714%. In our application, a 10x10 grid is chosen to extract the cell image. The average time required to extract features from a 300x300 image using a 10x10 grid is approximately 1 millisecond. Test data is 30% of the original motion data.

Figure 13 shows the confusion matrix of the k-NN classifier when trying to use the annotation. It can be calculated from the confusion matrix that the model can distinguish all movements. The time used for one frame is listed in Table I. As a result, our application reaches a frame rate of 5.3 fps. After processing the frames and extracting the time series from the step sequence, the

HMM model (including the 12 HMM chain) takes an average of 3.7 milliseconds to classify the movements.

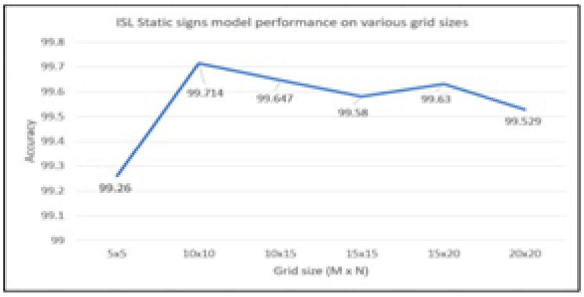

Figure 13: Comparison of k-NN classification accuracy of features extracted from cell data using different columns

Table 1: Calculation time for Each Stage

Sr. No.	Phase	Time taken (in ms)
1	Data Transfer over WLAN	46.2
2	Skin Colour Segmentation and Morphological Operations	12.3
3	Face Detection and Elimination	99.9
4	Object Stabilization	14.8
5	Feature Extraction	11.2
6	Hand Pose Classification	1.7
Average Time per Frame		186.1

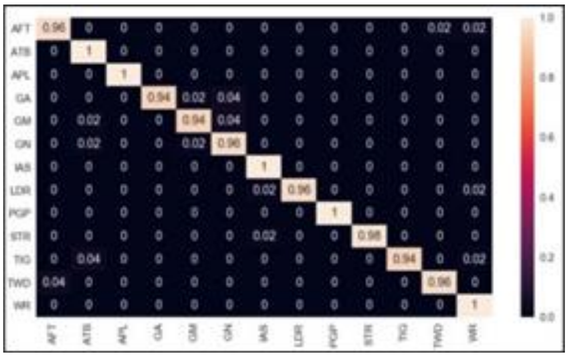

Figure 14: ISL Gesture Confusion Matrix Hmap

The 12 gestures and their abbreviations identified during the test are as follows:

- Then (AFT)
- Best wishes (ATB)
- Apple (APL)
- Good afternoon (GA) < br>
- Good morning (GM)
- Good night (GN)
- Sorry (IAS)
- Leader (LDR)
- Please Give Me Your Pen (PGP)
- Strike (STR))
- Awesome (TIG)
- Go (TWD)

Also Therefore, tests are performed against Wrong Moves (WR), i.e.movements that are not within the listed range. above. Similar to hand recognition, use 10x10grid and k-NN classifier to identify average hand poses. For each group pointing under light conditions, 50 snapshot tests were performed in which the person pointing wore a long shirt. Figure 14 shows the confusion matrix obtained by obtaining the results of these tests. Each cell of the confusion matrix contains the percentage of trials that yielded similar results. It clearly shows that the success rate of the classification is over 94%, with an average of 97.2%. Therefore, from the results obtained, it can be concluded that hand movement and hand classification are accurate enough to create the appropriate time for the hidden Markov model.

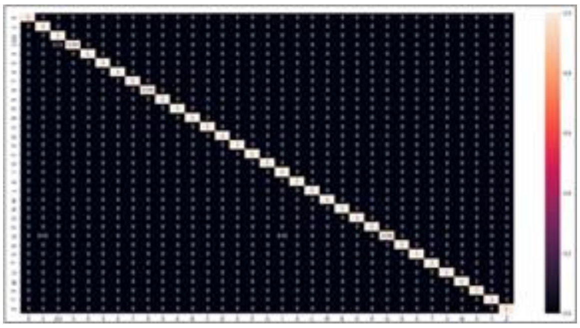

Figure 15: Confusion matrix heat map for k-NN classifier testing of ISL hand gestures

V. ISL APPLICATION

In this application, the desktop's camera is employed to capture user movements or gestures at a rate of 5 frames per second. All captured frames are continuously transmitted to a remote server for processing. The server handles the processing of frames, detecting poses or movements. Subsequently, the processed results are sent back to the application and presented at the top. The current implementation utilizes sockets to emulate client-server connections. Figure 16 displays a screenshot of the application.

Figure 16: (Left) hand gesture "a" recognized by the application; (right)The "Good Afternoon" gesture is recognized by the app

VI. FUTURE WORK

Currently, our research focuses on a single ISL hand gesture, but this method can expand to dual-handed movements with hand extraction algorithms. Incorporating natural language processing enables the system to recognize ISL sentences through multiple gestures in a single image. Manual extraction for skin color segmentation is no longer necessary, allowing for more flexibility in clothing choices. However, limitations arise for those who prefer shorter sleeves. Extending the system with object detection tools can address this, but it requires a diverse range of hand identification patterns for accurate detection in different positions, orientations, and backgrounds. Adequate lighting conditions are crucial, and employing robust color segmentation techniques adaptable to varying lighting conditions enhances feature extraction accuracy.

VII. CONCLUSION

The system detailed in this paper effectively tracks the informant's signature, employing fixed methods to eliminate face, skin color, and unwanted information, followed by manual deletion. It achieves a remarkable 99.7% accuracy in classifying all 33 ISL characters and 97.23% accuracy for 12 points. Utilizing HMM chains for each indicator and a k-NN model for classification, the system recognizes direction in approximately 0.2 seconds and gesture in 0.0037 seconds. Results indicate precise and instant recognition of ISL gestures, surpassing other methods in accuracy and speed discussed in literature. This approach, inspired by related studies, demonstrates effective communication and can extend to other sign languages, contingent on data meeting system requirements.

CONFLICT OF INTERREST

The authors declare that they have no conflict of interest.

ACKNOWLEDGMENT

The authors extend sincere gratitude to all contributors to this research paper. Saurabh Mishra,Mayank,Neha Singh, and Suraj Singh each played a vital role, contributing dedication and expertise that enriched the work. Special thanks to mentors Mr. Shubham Mishra and Ms. Neha Goyal for unwavering support, guidance, and valuable feedback, influencing the research trajectory. Their expertise enhanced data analysis and interpretation, bolstering the findings' robustness. Colleagues and friends offered constructive insights, contributing significantly to idea refinement. Lastly, heartfelt appreciation to families for patience, understanding, and continuous support, acting as a driving force in academic pursuits. To everyone mentioned, deepest thanks for essential contributions to this research paper's success; your efforts have not gone unnoticed.

CONFLICT OF INTERREST

The authors declare that they have no conflict of interest.

REFERENCES

[1] TalkingHands.co.in, "Talking Hands", 2014. [online]. Available:http://www.talkinghands.co.in/. [View: 21 July 2017].

[2] A. Agarwal and M. K. Thakur, "Sign language recognition using Microsoft Kinect," 6th International Conference on Contemporary Computing (IC3), September 2013.

[3] MailOnline, "English SignAloud glove hand gestures," 2016. [online]. can be found at:http://www.dailymail.co.uk/sciencetech/article-355 7362/SignAloudgloves-translate-sign-language-move ments-spoken-English.html. .[Viewed: 10 February 2018] .

[4] Alexia. Tsotsis, "MotionSavvy is a tablet application that understands sign language," 2014. [online]. Available:https://techcrunch.com/2014/06/06/motionsavvy-is-a-tablet-app-thatunderstands-sign-language/. [View February 10, 2018].

[5] P. Paudyal, A. Banerjee thiab S.K.S. Gupta, "SCEPTRE: A universal, non-intrusive and programmable gesture recognition technique," Proceedings of the 21st International Conference on Intelligent User Interfaces, p. 282-293, 2016.

[6] F. S. Chen, C. M. Fu, and C. L. Huang, "Gesture recognition using real-time tracking method and hidden Markov model," Image and Vision Computing, vol. 21, p. 745-758, 2003.

[7] P. Padmavathi, M. S. Saipreethy and V. Valliammai, "Character recognition in Indian Sign LanguageLanguage using neural networks", IJCACommentary on Recognition Models and Current Issues in Vision Taj RTPRIA, 2013.

[8] A.B. Jmaa, W. Mehdi, Y.B. Jemaa, thiab AB Hamadou, "Hand localization and finger feature extraction: Application to sign language digit recognition," International Conference on Intelligent Data Engineering and Automatic Learning, p. 151-159,2009.

ABOUT THE AUTHORS

	Saurabh Mishra Bachelor of Technology in C.S.E B.N. College Of Engineering And Technology, Lucknow Project Leader and Implementation Manager leads the Indian Sign Language Detection project, overseeing strategic guidance and project management. His team, with distinct roles, collaborates to innovate in research, data, and user experience.
	Neha Singh Bachelor of Technology in C.S.E B.N. College Of Engineering And Technology, Lucknow Data Acquisition and Preprocessing Expert Identify and collect diverse datasets of Indian Sign Language gestures for training and testing purposes.
	Mayank Bachelor of Technology in C.S.E B.N. College Of Engineering And Technology, Lucknow Research and Development Specialist Conduct in-depth research on Indian Sign Language (ISL) recognition using facial recognition techniques.
	Suraj Bachelor of Technology in C.S.E B.N. College Of Engineering And Technology, Lucknow User Experience (UX) and Accessibility Specialist Focus on creating an intuitive and user-friendly interface for the Indian Sign Language detection system.

Shubham Mishra
Assistant Professor, Department of CSE, B.N. College of Engineering & Technology, Lucknow, UP(India)
Project Guide