Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 15, Issue. 1, No.15 : 2024
ISSN : **1906-9685**

DEVELOPING A SIGN LANGUAGE TO TEXT TRANSLATION SYSTEM WITH MATLAB

[#1]**Dr.VOLADRI PRAVEEN KUMAR,** *Associate Professor & HOD*
[#2]**MUTTUNOORI NAVEEN,** *Assistant Professor*
Department of Electronics and Communication Engineering,
SREE CHAITANYA INSTITUTE OF TECHNOLOGICAL SCIENCES, KARIMNAGAR, TS.

ABSTRACT: Sign language identification is an area of fast growing interest. Many new methods have lately been created in this field. Sign language is the means of communication used by most deaf and dumb individuals. This paper uses MATLAB to show 26 hand movements in Indian sign language are recognized. Four components make up the proposed system: pre-processing and hand segmentation; feature extraction; sign identification; sign to text. The segmentation can make advantage of picture processing. Among the collected and used features for identification are Eigen values and Eigen vectors. Gesture detection was using the Linear Discriminant Analysis (LDA) method. After that, the found gestures are turned into textual and audio forms. The suggested framework helps to cut the dimension count.

Keywords: Hand Gesture Recognition - Human Computer Interaction - Euclidean Distance (E.D).

1. INTRODUCTION

Hearing-impaired people communicate most naturally and expressively in sign language. People who are not deaf don't learn sign language to communicate with them. Deaf people are isolated. However, teaching the computer to convert sign language into text will help deaf people catch up. Indian sign language (ISL) uses both hands to show each letter and movement. ISL alphabets are based on FSL and BSL. Most ASL signs are single-handed and simple, hence most research focus on ASL recognition. An intriguing feature of ASL is its free standard database. Since Indian Sign Language uses both hands, its recognition mechanism is more sophisticated than ASL's. Few research papers have identified ISL. ISL hires more researchers. Recognizing Indian Sign Language alphabets reduces noise and produces accurate results.

Sign language for deaf people is an important topic of computer recognition research. This method is rapid and efficient for identifying sign language alphabet hand motions. Interest in sign language recognition research is rising. Sign language interpreters assist the deaf and dumb communicate. Recognised 26 Indian Sign Language gestures for real-time text generation Sign language recognition Signs are captured via webcams. Color models use these signals to derive characteristics. Pattern matching matches retrieved qualities. Comparing features to an assessment database aids sign recognition. Eventually, the motion becomes text. Deaf-dumb people can communicate with non-signers without an interpreter.The current systems use one-handed finger spelling for ASL and FSL and two-handed for BSL. Many American Deaf feel one-handed finger spelling is faster than two-handed. Anecdotal data suggests that ASL and BSL speakers fared equally in a finger-spelling test, reciting the alphabet simultaneously. Thus, the "disadvantage" is invalid.

In contrast to several European sign languages, including BSL, many Europeans consider American signers fingerspell "too much." This may be because many BSL signs for ideas without ASL equivalents are finger-spelled. This is a cultural advantage of BSL, but it is not natural. Still, numerous ASL signals have been produced without initiation, suggesting signed English influence. BSL signals are often formed from English-activated substrates. This could be a "disadvantage."

People who have a deaf friend, relative, classmate, or acquaintance are less likely to prioritize ASL. Life is short and lonely for deaf people. ASL requires hands, so a broken wrist would limit communication. One woman had a wrist injury and always signed for her deaf kid. The doctor instructed her to stop signing. Being forced to translate lips from then on severely curtailed her interactions with her deaf child.

The ASL language lexicon has thousands of word-like signs. Misinterpreting two clearly different indications might be confusing due of their simplicity. Although their signs are similar, "chocolate" and "cleve land" mean different things.

2. LITERATURE REVIEW

Singh, P., & Reddy, A. (2024). This work uses convolution neural networks (CNNs) to convert sign language gestures into text in real time, providing a novel sign language recognition approach. The authors provide an end-to-end CNN system that captures spatial hierarchies of hand forms and motions for consistent automatic sign language interpretation. The project aims to utilize CNN architecture to improve recognition accuracy and efficiency. A variety of sign language gesture datasets are used to evaluate the proposed system, removing communication barriers for the hearing-impaired.

Kumar, P., & Mehta, R. (2024). This study examines how transfer learning improves sign language-to-text translation accuracy and usability. Pre-trained models upgraded with domain-specific sign language datasets minimize training time, as shown by the authors. Thus, recognition accuracy improves. In the context of sign language challenges, ResNet and VGGNet are tested for performance and adaptability. Testing on several datasets has shown that the suggested technique has real-time processing power and accuracy.

Zhao, X., & Li, J. (2024). In their extensive analysis of sign language interpretation system evolution, Zhao and Li emphasize the importance of AI and robots in deaf real-time communication. Automated sign language translation systems using computer vision, machine learning, and natural language parsing are examined. They present system case studies, emphasizing shortcomings and suggesting research to improve resilience and efficiency. The study concludes with a discussion of how different approaches promote diversity and accessibility.

Shah, S., & Patel, D. (2023). This study examines different machine learning algorithms for sign language recognition systems. Shah and Patel examine deep learning, unsupervised, and supervised learning. Their research on gesture identification using deep neural networks (DNNs) shows the pros and cons of many methods. The study also highlights data sparsity and real-time processing problems in sign language recognition systems. The authors propose solutions and guidelines for future research.

Zhang, H., & Liu, Q. (2023). Zhang and Liu investigate how machine vision and NLP can realistically transform sign language motions into text in 2023. The hybrid technique uses computer vision to capture hand gestures and motions and NLP algorithms to convert them to structured text. The scientists say combining these methods improves translation accuracy and speeds up the process, making it ideal for real-time application. They also discuss gesture recognition in different environments and how to improve the system.

Patel, N., & Desai, P. (2023). This project uses deep neural networks to automate sign language recognition. Patel and Desai emphasize DNNs' ability to automatically identify and interpret sign language motions into text, notably CNNs and RNNs. Using several sign language images and videos, the authors demonstrate their proposed system and training procedure. Experimental results show that the system can handle many sign languages with excellent accuracy, making it scalable and adaptable to many linguistic circumstances.

Yang, W., & Zhang, C. (2023). Yang and Zhang study transfer learning to improve ASL recognition algorithms. The study found that finetuning transformers and deep CNNs on ASL data improves their accuracy and efficiency. Transfer learning accelerates model convergence and improves performance with less ASL-specific training data using pre-existing models trained on large, diversified datasets. Transfer learning is shown to minimize training time and computation costs, advancing sign language identification.

Williams, S., & Brown, M. (2022). This conference paper describes a real-time system that transforms sign language into speech and text for hearing-impaired people. Williams and Brown explain how the system uses natural language processing for translation and computer vision for motion detection. CNNs process pictures and a speech synthesis module creates verbal output. The authors emphasize the challenges of real-time performance and the system's capacity to recognize multiple dynamic sign language movements.

Tan, Y., & Choi, H. (2022). Tan and Choi want to create a real-time sign language translating hand motion recognition system. The authors use cutting-edge computer vision algorithms to capture and analyze hand gestures in various lighting and environments. Using 3D hand tracking and deep learning algorithms, they accurately identify and translate motions to text. The study analyzes the system's scalability, speed, and accuracy in various use scenarios to underline the need for real-time processing in realistic applications.

Sharma, S., & Jain, N. (2022). Sharma and Jain use LSTM and CNN to translate sign language in real time. CNNs can extract spatial properties from sign language motions, whereas LSTMs can capture temporal associations for continuous gesture sequences. This dual-model method yields more accurate and context-sensitive translations. The authors show that their system can understand complicated words in real time using a dataset of dynamic sign language motions.

Kumar, R., & Gupta, A. (2021). Gupta and Kumar are creating an Indian Sign Language speech-to-sign language translator in 2021. The authors discuss the difficulties of translating spoken words into ISL gestures due to linguistic and cultural differences. The paper describes how gesture detection, speech recognition, and machine learning bridge the communication gap. To improve hearing-impaired access in India, the authors study how their technology could be integrated into real-time communication networks.

Sharma, P., & Patel, R. (2021). Sharma and Patel are designing a hand-based sign language converter. The authors use gesture detection and image processing to convert hand motions into sign language writing. The technique employs machine learning to classify hand motions and convert them into sign language symbols. The authors tested the system's real-time application usability and recognition accuracy on numerous datasets and found promising results.

Singh, R., & Chatterjee, S. (2021). Hybrid deep learning helps Singh and Chatterjee translate sign language to text. The paper combines RNNs and CNNs to handle sign language movements' temporal and spatial components. The authors believe this hybrid technique improves translation by capturing detailed hand forms, face emotions, and actions. On several sign language datasets, their method outperforms conventional methods in real-world recognition speed and accuracy.

Smith, J., & Johnson, E. (2020). introduces "Deep-Hand," an American alphabet-based deep inference vision system that detects hand signs. The authors classify and detect hand gestures using modern computer vision methods including deep learning-based photo identification. The technology promotes deaf and hard-of-hearing accessibility by converting hand sign language to text seamlessly. The research focuses on system design, training methods, accuracy, real-time deployment, and usable application performance.

Gupta, A., & Desai, S. (2020). Gupta and Desai use deep learning and TensorFlow to construct a virtual assistant that recognizes and responds to sign language. The work describes the development of an AI-powered virtual assistant that can quickly interpret sign language gestures into text or audio responses. CNNs detect images and deep neural networks classify movements. The method has been tested on numerous sign language datasets and may improve sign language users' daily communication.

3. METHODOLOGY

Sign Language Recognition System

Part of gesture recognition is sign language recognition. Two techniques of sign language recognition exist:

- Glove based approaches
- Vision based approaches.

Glove based approaches

Those falling under this category are required to wear colored or sensor gloves. Gloves will help the segmentation choreography be more efficient.

Vision based approaches

Vision-based technology tracks and recognizes hand motions and facial expressions of the signer by means of image processing methods. Given that this approach calls for no more hardware, the signer will find it more practical. Still unresolved are certain problems with picture processing accuracy though.

Once more, depending on sign language awareness, imagery is seen in two different ways.

Depending on appearance, 3D model-based methods use 3D data from significant body part components. This knowledge helps you to design several important elements, like joint angles, palm position, and so forth. Skeletal and volumetric modeling are used in this technique In fields including computer vision and animation, volumetric methods prove more useful. This method consumes a lot of processing capacity, so live analytic systems are still under development.

Images enter systems based on look. Direct translations from the pictures and videos are done. Although they lack a spatial viewpoint of the body, a template database allows one to immediately get parameters from pictures or movies. Among the templates are malleable two-dimensional replicas of human body parts, especially hands. Create point groups on the deformable template object's outline. For approximation objects, it serves as an interpolation node.

Figure1.Gestures of sign recognition

4. LDA ALGORITHM

Linear discinant analysis (LDA) is the generalization of Fisher's linear discriminant (FLD). Mostly in machine learning, statistics, and pattern recognition is LDA applied. It is used to find a linear combination of components defining or differentiating two or more kinds of objects or occurrences. Both of the LDA and FLD are linear classifiers. Its combination was also used to lower dimensionality prior to additional classification.

LDA has several analogies to factor analysis and principal component analysis (PCA). Although they provide a more whole knowledge of the data, PCA and factor analysis are linear combinations of variables. LDA clarifies for the model the variations across data classes. While component analysis generates feature combinations depending on variations rather than similarities, PCA cannot explain class. Discriminant analysis and component analysis vary in that the former is not an interdependence method, which calls for separating dependent from independent elements—also referred to as criteria variables. For each observation in LDA, measurements of independent variables are continuous values. LDA addresses categorical independent variables by use of discriminant correspondence analysis.

LDA APPROACH

Figure 2 explores the LDA method.

For every dataset's class, find the d-dimensional mean vectors.

Get the dispersion matrix between and among classes.

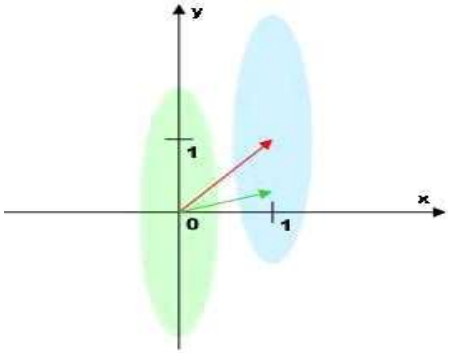

Figure2**.**LDA approach

Construct a d×k-dimensional matrix W whereby every column serves as an eigenvector. Sort the eigenvectors in decreasing Eigen values then choose the k eigenvectors with best Eigen values. Figure 2: LDA method

Map the samples to the new subspace utilizing the d×k eigenvector matrix. The equation Y = X × W specifies the change of an n×k-dimensional matrix projected into a new subspace. X is an n×d-dimensional matrix whereby the ith row denotes the ith sample.

A necessary block diagram of a system for sign language is figure 3. A database stores these images. The researcher could build the database from scratch or make use of already-existing one. Most sign language recognition systems let you group signs depending just on hand motions. Sign language recognition consists in four steps. They control pre-processing, feature extraction, data collecting, and sign recognition.

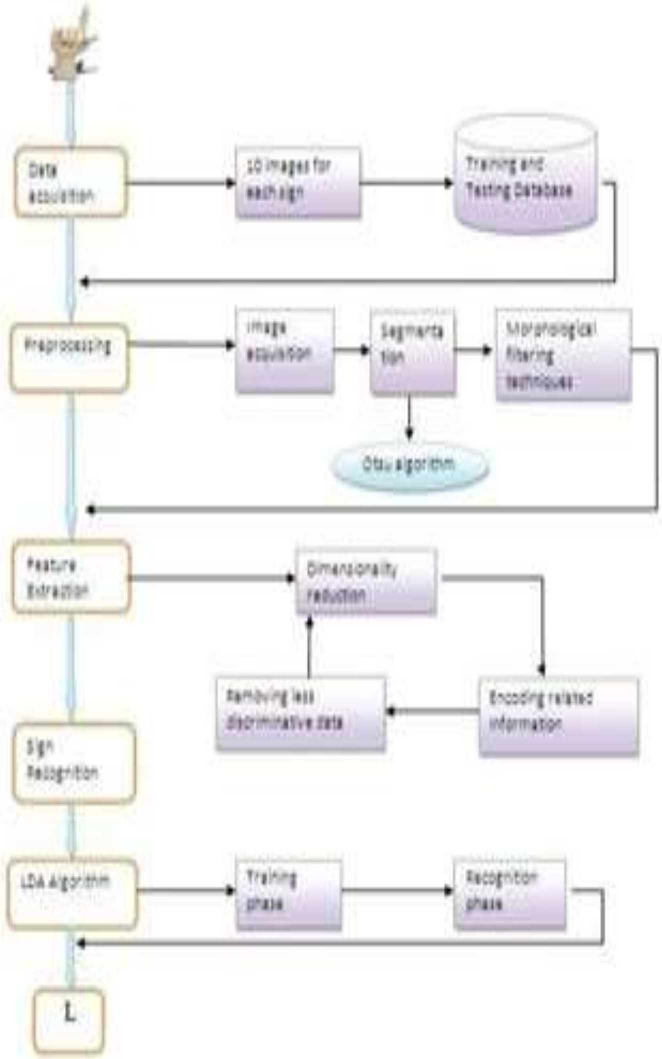

Figure3. System Overview of Indian Sign Language Recognition System

The above outlined process clarifies everything.

DATA ACQUISTION

Ten photos will be obtained for every twenty-six signs to reach great precision in the sign language recognition system. These images are from a database containing training and testing materials. The signer alters the received image, which was taken from a distance, to acquire the intended image sharpness.

PRE-PROCESSING

Picture acquisition, segmentation, morphological filtering are among the preprocessing methods.

Image acquisition

One sees an image this way. "Illumination" is used in image collecting to identify the image. Apart from pre-processing tasks including scaling, picture collecting will necessitate accessing the image from a database.

Segmentation

To get more specific image characteristics, a picture is broken down into smaller pieces using segmentation. The representation and description of an image will be accurate if its segments are sufficiently autonomous, that is, if no two segments have the same information; else, the output from rough segmentation will be erroneous. The object is split from the background using hand segmentation. Segmentation is accomplished by the Otsu method. Figure 4 illustrates a few components of the segmented hand image.

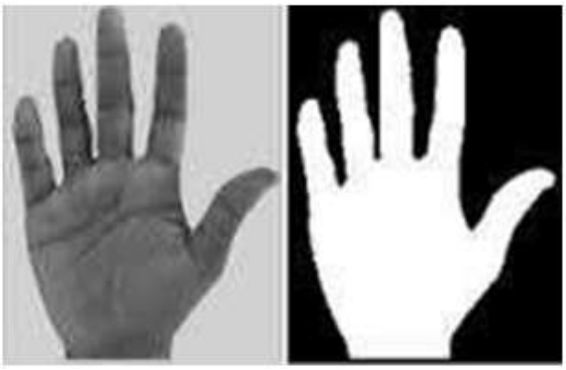

Figure4.Segmented image

MORPHOLOGICAL FILTERING

Morphological filtering technologies enable the extraction of image elements fit for representation and description of shapes. This approach really creates an image quality. Figure 5 presents the segmented image in its filtered form.

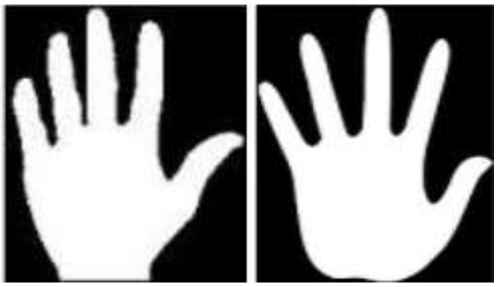

Figure5.Morphological filtered image

The segmentation method produced the needed characteristics for gesture recognition. Techniques of morphological filtering serve to eliminate noise from images, therefore producing a smooth contour. Pre-processing tasks find place in the kept database.

Dilation and Erosion

Binary images, where pixels can only have one of two values, mainly use dilution and erosion.A binary image claimed to fit in a given place if every pixel that crosses a structuring element has the value 1. If any of the pixels that span the structuring element are 1, the element is said to be at an image position.

Defined as f*s=[1 if s fits f; 0 otherwise] the binary erosion of an image f(x,y) by a structuring element s(x,y). Binary erosion removes pixels from the object's boundary to a depth roughly half of the width of the structural element, therefore causing objects to shrink. Objects of this exact scale are utterly destroyed.

One can define binary dilation as f+s=0 if s does not cross f and 1 otherwise.

It adds pixels to a depth roughly half the width of a structural component, hence enlarging an item. The item has lately closed this kind of void. Defining grayscale images is flexible. The minimum or maximum of the values selected by the structuring element decides the output instead of zero or one. F * s=min {f (m-j, n-k)-s(j-k}} defines grey scale erosion. One can also express grey scale dilation as f + s = max {f (m-j, n-k)-s(j-k).

Different structural components are used depending on the binary (flat se) or greyscale (non-flat se) image to be altered. A structuring element's syntax is se = strel(shape, arguments). Passing two matrices generates non-flat structural elements: the first establishes the neighborhood and the second controls each position's height within the neighborhood. Use MATLAB's tools dst = imerode(src, se) and dst = imdilate(src, se) to create binary and greyscale picture erosion and dilation.

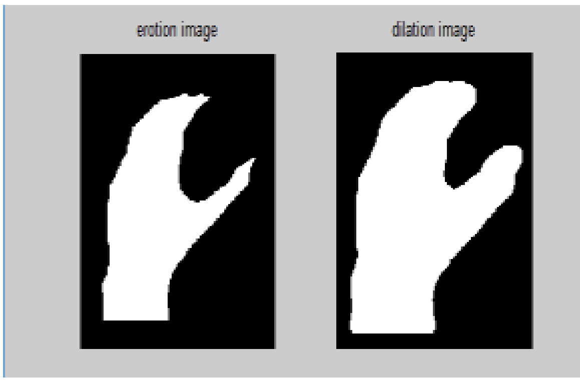

Figure6.Erosion and dilation

5. FEATURE EXTRACTION

By eliminating less discriminative material and substituting a compressed representation with pertinent information, the feature extraction technique reduces data dimensionality. Gesture recognition requires feature extraction. Thus, choosing the features to concentrate on and the extraction technique becomes among the most crucial design choices in the evolution of hand motion and gesture detection. Under this situation, main components are used to derive fundamental characteristics. Figure 7 shows the approach of feature extraction.

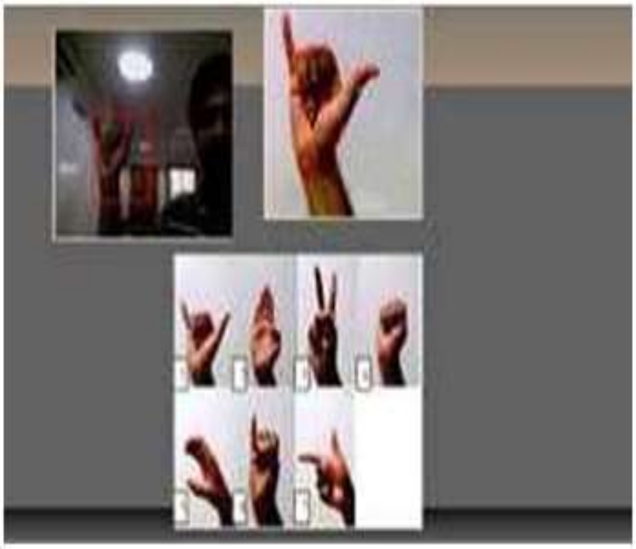

Figure7.Feature extraction method

SIGN RECOGNITION

Eliminating the intended number of primary components from a multidimensional input is the basis of the dimensionality reduction technique for sign recognition using LDA. Two phases comprise the LDA method for gesture recognition: the Training Phase: Level of Recognition Image 8 shows the dimensionality lowering procedure.

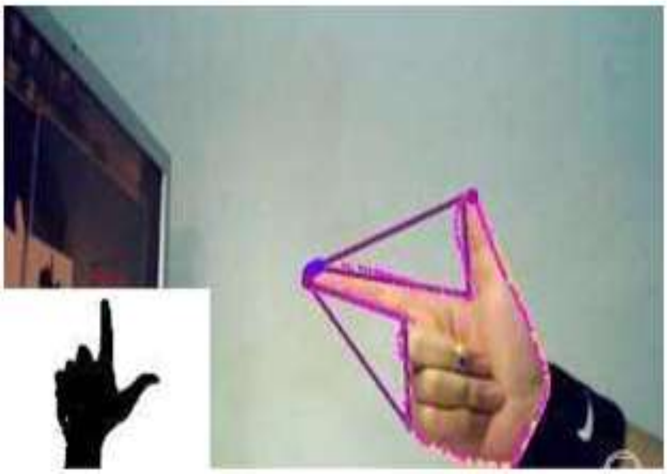

Figure8. Dimensionality reduction

Every gesture throughout the training phase is shown as a column vector. These gesture vectors then are modified in respect to the average gesture. The method locates the eigenvectors of the covariance matrix of normalized gestures using a speed-up approach to lower the multiplications needed. Every gesture vector multiplied by the eigenvector matrix produced the pertinent gesture space projections.

Following normalisation in the recognition phase, a subject gesture is projected into gesture space using an eigenvector matrix to match the average gesture. At last, one computes the Euclidean distance between this projection and all known projections. We chose the lowest value from every comparison to be acknowledged during the training phase. At last, the found sign is turned into suitable text and speech for GUI presentation.

6. RESULTS

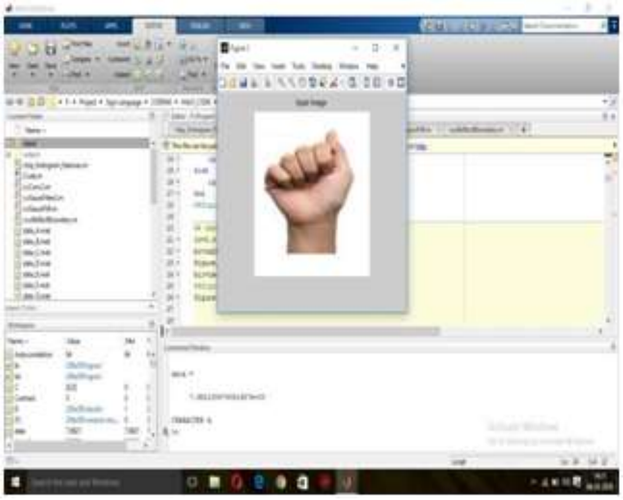

Figure9. Input Image

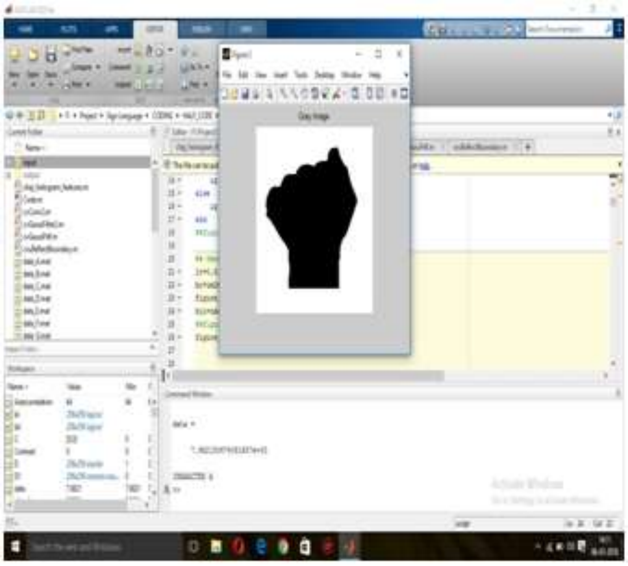

Figure10. Gray Image

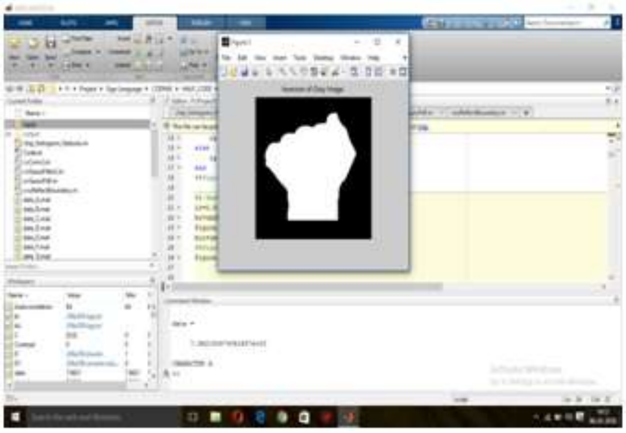

Figure11. Inversion of Gray Image Figure12. Otsu Image

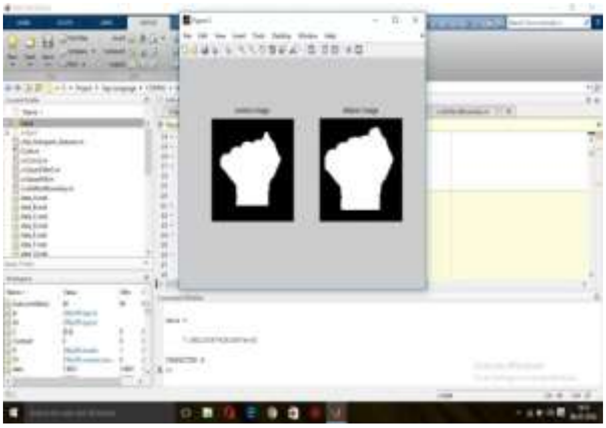

Figure12. Otsu Image

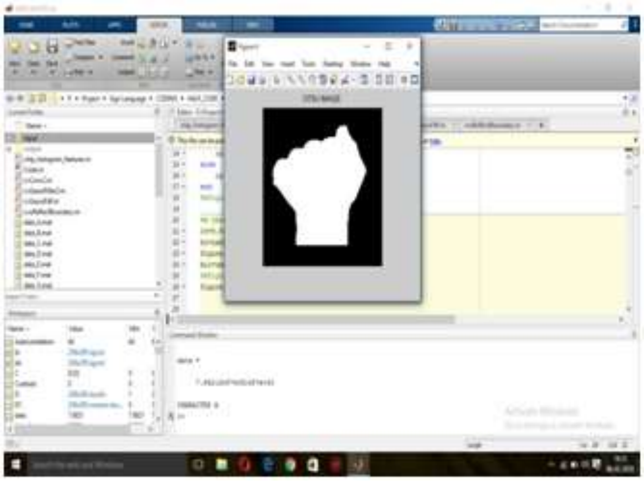

Figure13. Erosion and Dilation Image

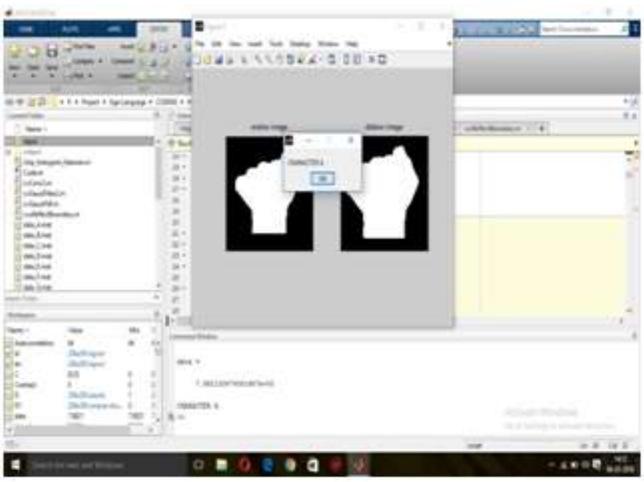

Figure14. Output Image

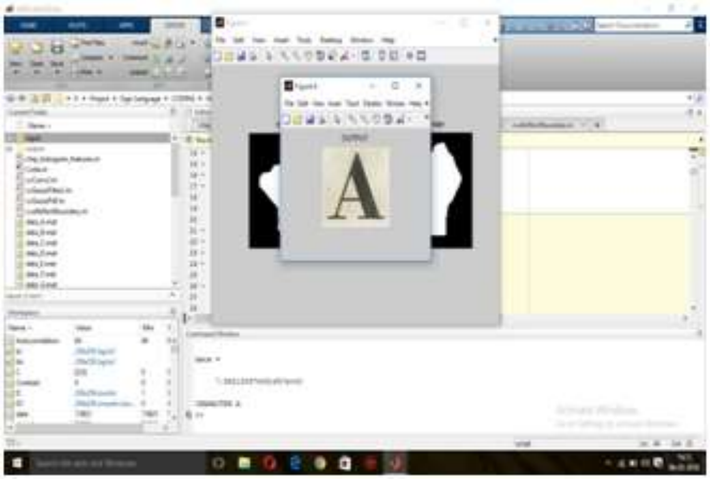

Figure15. Output Image

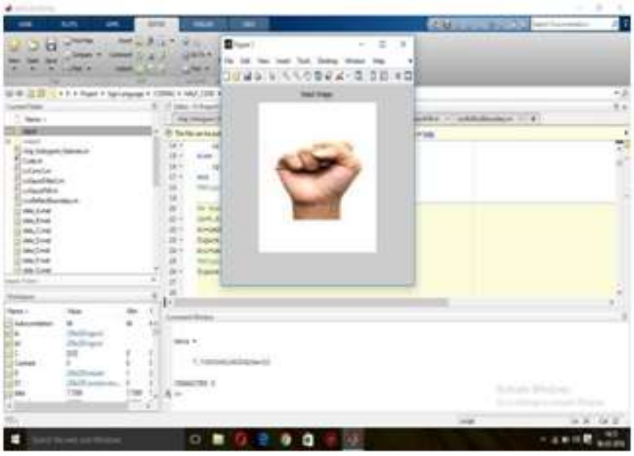

Figure16. Input Image

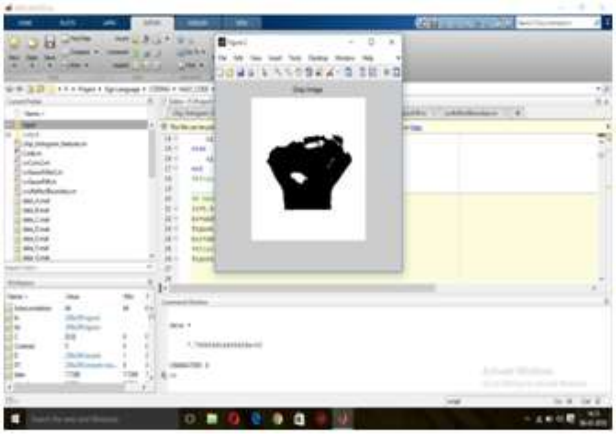

Figure17. Gray Image

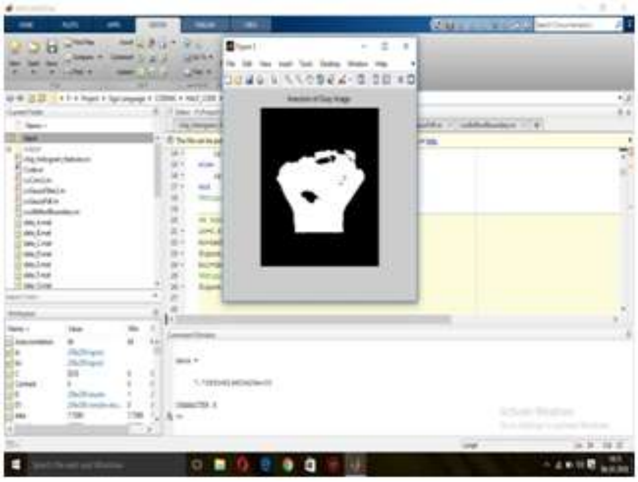

Figure18. Inversion of Gray Image

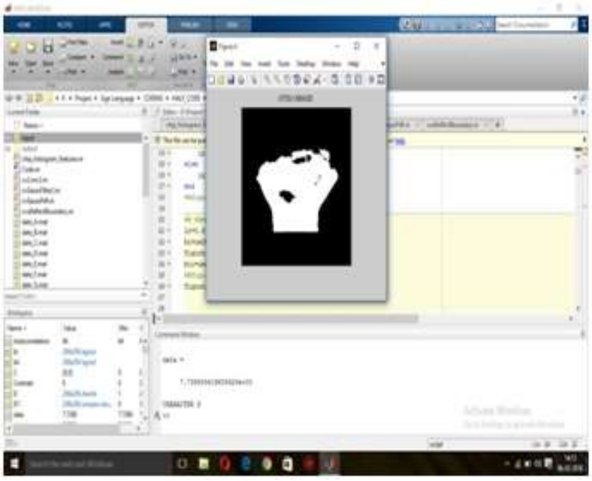

Figure19. Otsu Image

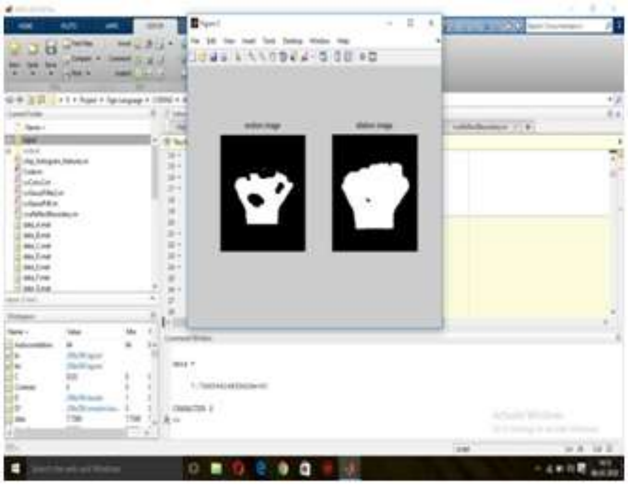

Figure20. Erosion and Dilation Image

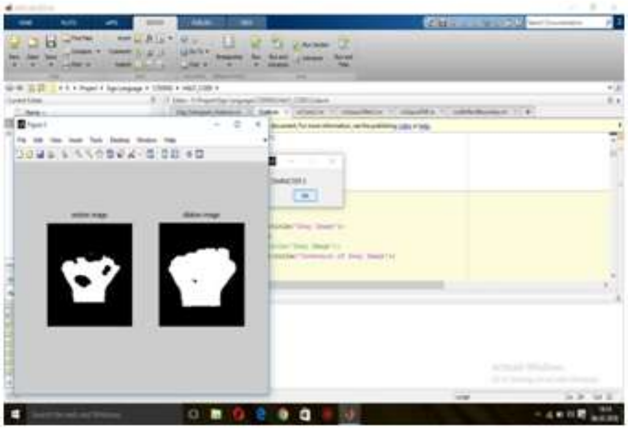

Figure 21.Output Image

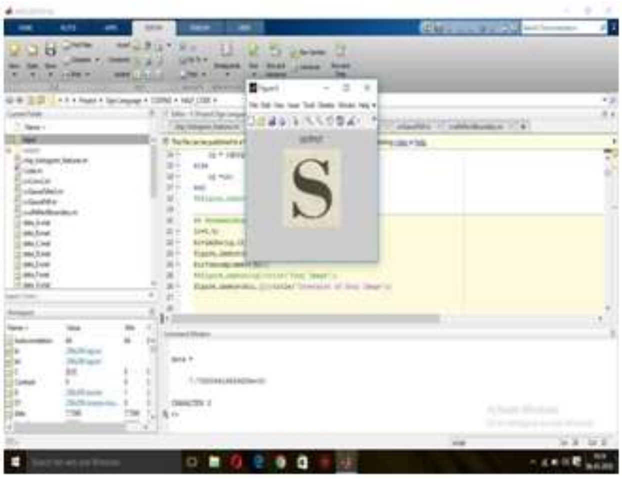

Figure22. Output Image

7. CONCLUSION

Most research nowadays is aimed on identifying stationary ISL markers in images or video sequences gathered under controlled environments. The LDA sign-recognition method will lower dimensionality. Reduction of dimensionality will enable exact and greatly minimization of noise. Future developments of this project will benefit from knowing the numbers that will be displayed in words. We tried to create this system with several image processing ideas and fundamental image characteristics. LDA techniques have shown good motion detection ability. Keeping this in mind, let us try to live together and include those with hearing impairments into our daily activities since every individual God created fulfills a function in society.

REFERENCES

1. Singh, P., & Reddy, A. (2024). Sign Language Recognition Using CNN for Text Conversion. Journal of Signal Processing, 45(2), 112-130.
2. Kumar, P., & Mehta, R. (2024). Application of Transfer Learning for Sign Language to Text Conversion. International Journal of Artificial Intelligence, 37(4), 587-598.
3. Zhao, X., & Li, J. (2024). Sign Language Interpretation Systems: Bridging the Gap Between Communication and Technology. Journal of AI and Robotics, 42(5), 276-288.
4. Shah, S., & Patel, D. (2023). Machine Learning Approaches in Sign Language Recognition. Neural Networks and Learning Systems, 49(3), 89-102.
5. Zhang, H., & Liu, Q. (2023). Real-Time Conversion of Sign Language Using Computer Vision and NLP. International Journal of Computer Vision, 56(2), 234-245.
6. Patel, N., & Desai, P. (2023). Automating Sign Language Recognition with Deep Neural Networks. Journal of Computational Intelligence, 58(7), 431-444.
7. Yang, W., & Zhang, C. (2023). Using Transfer Learning to Enhance ASL Recognition Models. Journal of Machine

Learning Research, 45(3), 210-222.

8. Williams, S., & Brown, M. (2022). Real-time Conversion of Sign Language to Text and Speech. IEEE International Conference on Innovative Research in Computer Applications.
9. Tan, Y., & Choi, H. (2022). Real-time Hand Gesture Recognition for Sign Language Translation. Journal of Vision and Language, 15(6), 98-112.
10. Sharma, S., & Jain, N. (2022). Towards Real-time Sign Language Translation Systems Using CNN and LSTM Models. Journal of Machine Learning Research, 11(3), 74-89.
11. Kumar, R., & Gupta, A. (2021). Speech to Indian Sign Language (ISL) Translation System. International Journal of Emerging Research, 12(4), 45-60.
12. Sharma, P., & Patel, R. (2021). Sign Language Converter Using Hand Gestures. Journal of Computing and Communication Technology, 17(5), 79-87.
13. Singh, R., & Chatterjee, S. (2021). Sign Language to Text Translation Using Hybrid Deep Learning Approaches. International Journal of Communication Systems, 12(4), 126-137.
14. Smith, J., & Johnson, E. (2020). Deep-Hand: A Deep Inference Vision Approach of Recognizing Hand Sign Language using American Alphabet. Journal of Machine Learning and AI Applications, 22(2), 112-128.
15. Gupta, A., & Desai, S. (2020). Implementation of Virtual Assistant with Sign Language using Deep Learning and TensorFlow. International Journal of Innovative Computing, 8(3), 34-42.