[1]Kumaravelu Sangeetha

[1]Navaneetha Krishna Bose

[2]Duraimutharasan

Enhancing Recognition of Indian Sign Language via Fusion of SURF-based SVM and CNN Models for Seamless Integration into a Regional Language Translation System

Journal of Electrical Systems

Abstract: - Efficient communication is fundamental for human interaction however, people with speech and hearing impairments frequently find it difficult to communicate with others without the help of a translation. For the deaf and mute people, sign language is an essential form of nonverbal communication, highlighting the necessity of having a system that can understand and recognize sign language. In order to enable seamless communication for those with hearing and speech impairments, this research presents a novel method for the automatic identification of finger typing in Indian sign language. The suggested framework processes input signs, employs skin colour-based segmentation, and utilizes various image processing techniques to detect and transform sign shapes and there is a voice module which will convert the sign language into speech in respective regional language. After that, the discovered region is converted into a binary image. The binary picture that is produced is then subjected to the Euclidean distance transformation. On the image that has been modified by distance, row and column projection is used. Central moments are utilized in conjunction with HU's moments for feature extraction. Neural networks and SVM are used for classification. As for India is considered there are so many languages which are spoken all over the country. When the same sign language is used the interpreter finds it difficult to understand so we provide a system when ever the sign language is used it can take up any type of languages which are being used. Revolutionizing Communication for Speech-Impaired Individuals: A Comprehensive Framework Recognizing Indian Sign Language with BOVW, SURF, SVM, and CNN. Integration with Real-time Video Stream Analysis and Multimodal Output is being generated as voice.

Keywords: Indian Sign Language Recognition, SVM, CNN

1.Introduction

A vital means of communication for those who are hearing impaired is sign language providing them with a means to express their thoughts and feelings. In contrast to the effortless communication experienced by the reading, writing, body language, gestures, and speech to the general public, individuals facing speaking difficulties encounter unique challenges. For these individuals, reliance on sign language becomes imperative, presenting a heightened barrier to effective communication with the larger majority. It is a complex system, consisting of varied hand shapes and actions that create different gestures, orientations, and facial expressions. These elements come together to convey nuanced meanings and facilitate effective communication within the deaf-dumb community. For deaf-dumb individuals, communicating in public places can be particularly challenging when interacting with those who are not trained in sign language. Common settings like banks, hospitals, and post offices may pose communication barriers, leading to difficulties in expressing needs, asking questions, or seeking assistance. This lack of effective communication can result in frustration and isolation for deaf-dumb individuals, impacting their ability to navigate everyday situations independently. In some cases, deaf individuals may resort to seeking the help of sign language interpreters to facilitate communication with hearing individuals. Interpreters play a crucial role in converting spoken language from sign language and the reverse way around. But this fix is not always practical or readily available, and the need for interpreters can be both costly and limited in availability.There's considerable interest in to address these issues through developing technological solutions, sign language recognition systems. These systems aim to bridge the communication gap by enabling deaf-dumb individuals to communicate more easily in various public settings. By leveraging technology, these individuals can express themselves, ask questions, and access services without the need for a human interpreter, providing greater independence and inclusivity in their daily lives. The development of such systems involves advanced techniques like image processing, machine learning, and artificial intelligence to effectively decipher and convert sign language motions into written language or speech. The deaf community in

[1 *]Department of Computer Science & Engineering, AMET University, Kanathur, Chennai 603112, Tamilnadu, India

[2]Dean-Faculty of Advanced Computing Sciences, AMET University, Kanathur, Chennai 603112, Tamilnadu, India

*Corresponding Author: Kumaravelu Sangeetha

*Email: sanrith1217@gmail.com

India relies heavily on Indian Sign Language (ISL) for communication. Acknowledged as a well-established and standard system, ISL facilitates discussion among deaf individuals within the country. While English is often spoken alongside ISL, the latter has its own set of symbols and gestures for expressing thoughts and ideas. In ISL various symbols are assigned to different alphabets of the English language, forming a structured and comprehensive method of communication. This includes not only word-level gestures but also finger spelling, providing a rich and versatile vocabulary for users. The integration allows for a nuanced and expressive form of communication.

This paper focuses on the automated detection of motionless motions in Indian Sign Lanaguage. The indicators that are taken into account for identification encompass the alphabets. The objective aims to create a system that is capable of precisely identify these static signs, contributing to improved interaction accessibility for the community of the deaf in India. Methodology likely involves advanced machine learning algorithms, feature extraction methods, and image processing strategies to train a model capable of recognizing and distinguishing the unique characteristics of each static gesture in the ISL alphabet. By automating the recognition process, this research aims to enhance the efficiency and effectiveness of communication for individuals using Indian Sign Language, fostering greater inclusivity and understanding in both deaf and hearing communities. The exclusive use of sign language by those with speech impairments intensifies the difficulty in expressing themselves and understanding others. This underscores the necessity for systems that can recognize sign language seamlessly interpret then transform translating spoken or written language into sign language and vice versa.

Recognizing this gap, researchers globally are dedicating their efforts to the creation of technologies for automatically recognizing sign language. These innovative technologies aim to bridge the communication divide, offering a more accessible and efficient means for individuals with speech impairments to interact with the broader community. As a result, the ongoing exploration and advancement of sign language recognition systems represent a pivotal step towards fostering inclusive communication and breaking down barriers for those affected by speech impairments.

The delayed standardization of Indian Sign Language (ISL) can be attributed to evident factors. The initiation of Indian Sign Language studies in India dates back to 1978; however, due to the absence of a standardized form, its utilization remained limited to short-term courses. Compounding the challenge, the gestures employed in deaf schools exhibited considerable variations, attending these schools, around 5% of the deaf population as a whole.

The lack of a standardized ISL posed difficulties in achieving widespread understanding and communication consistency. It wasn't until 2003 that significant progress was made with the standardization of Indian Sign Language, garnering increased attention from researchers. This pivotal development marked a turning point, laying the groundwork for a more unified and accessible form of ISL, thereby addressing the longstanding issues associated with varied gestures and fostering greater cohesion within the deaf community [1] Despite India's status as a diverse nation, encompassing nearly 17.7% of the global community, Notably little research has been done in this field. This stands in stark contrast to the efforts seen in other countries [2–3]. The primary focus lies on the real-time classification and recognition of Indian sign language provided by a deaf-mute user. Therefore, the algorithm's simplicity and speed are crucial. The system approach entails segmenting the hand based on statistics related to skin color, turning the segmented image into a binary format, and then applying feature extraction techniques to the binary image. These techniques include distance transformation, Discrete Fourier Transform, and probability distribution properties such as central moments. [4]. Figure 1 illustrates the gestures corresponding to each alphabet

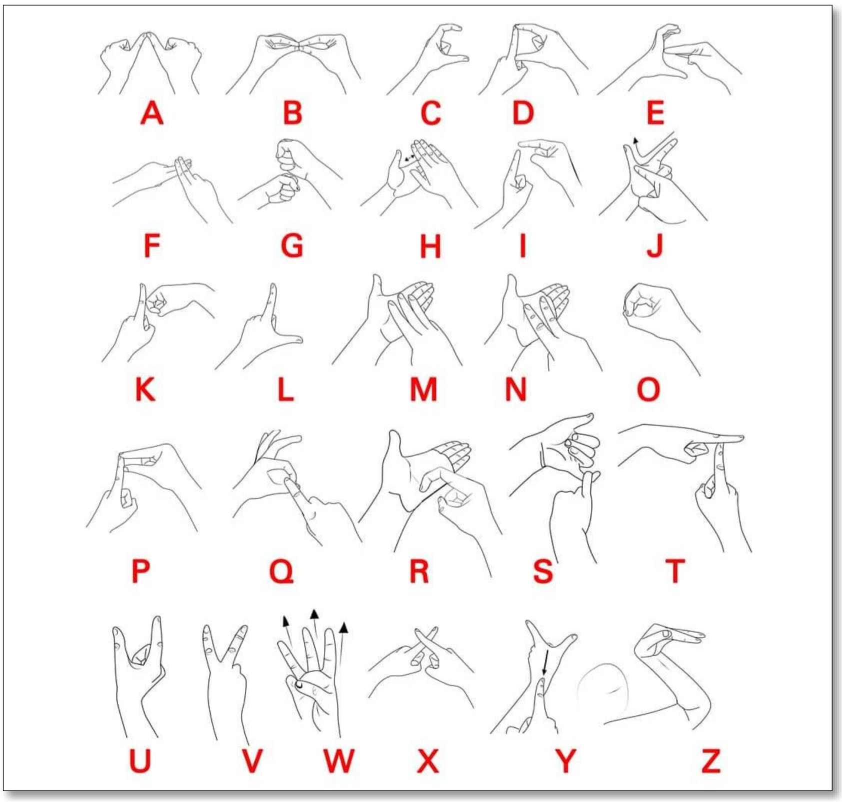

Figure 1 showcases of visual representation

It is characterized by its complexity, comprising a diverse set of features that pose challenges to standardization efforts. The language incorporates both static and dynamic signs, introducing a dynamic aspect to communication. Furthermore, it includes signs, adding an additional layer of intricacy to the lexicon. One of the notable complexities arises from regional variations within India, where different geographical areas may have distinct signs for the same alphabet. This diversity is influenced by cultural nuances, historical factors, and local preferences. As a result, attempts to introduce a standardized scheme for ISL encounter difficulties due to the need for flexibility to accommodate these regional differences. The absence of a comprehensive and universally accepted collection of signs hinders the development of consistent models and systems for sign language recognition. Without a standard dataset, the training and validation of algorithms become more challenging, affecting the dependability and accuracy of systems for recognizing sign language.

In essence, the multifaceted nature of ISL, involving variations in signs, linguistic nuances across regions, and the lack of an established source, manifests The complexity of Indian Sign Language. Addressing these challenges requires a nuanced approach that considers the cultural and regional diversity inherent in ISL, as well as collaborative efforts to establish a common ground for standardization and technological advancements in the field.

Two primary approaches for sign language recognition are the sensor oriented approach and the vision oriented approach. In the sensor-oriented methodology, customized hardware such as glove is applied to identify finger motions. These gestures are then translated into comparable electrical impulses facilitating sign interpretation. Notably, the approach holds an advantage due to its lack of specialized hardware requirements. This characteristic makes it a more spontaneous and user-friendly option, particularly favoured by signers [5]. The absence of additional equipment simplifies the implementation of the vision-based gesture recognition system, contributing to its popularity among users and scholars working in the sign language sector technology.

The authors of this work provide a way for creating a reliable, real-time system for identifying numerals and alphabets. Unlike approaches that involve superior technologies, it focus on recognizing directly from visuals, that are being captured. It discusses achieved accuracy of the proposed system's results. The emphasis is on acquiring timely, accurate, and effective recognition of signs. The ultimate goal is To reduce the gap in communication between individuals are able-bodied. By leveraging accessible technology like webcams and focusing on image-based recognition, the methodology seeks to provide a useful and optimal vision which enhancing communication accessibility within the Indian Sign Language community.

2. Related works

Various authors have adopted diverse methodologies in their exploration of sign language, influenced by the unique characteristics and gestures inherent to each form.

J. Singha and colleagues, in their work outlined in reference [6], introduced a instantaneous recognition method for sign gestures. A methodology employed a classification technique based on Eigen value-weighted Euclidean distance. This approach involves assigning weights to Eigen values to measure the distance between features, enabling the accurate classification of various sign gestures in real-time scenarios. The use of Eigen value-weighted Euclidean distance reflects a sophisticated and efficient approach to sign language recognition, contributing to advancements in this field. Author has, a detailed in their work referenced as [7], they introduced a sign language in real time recognition system utilizing Principal Component Analysis (PCA). For data acquisition, a total of 260 images were employed, encompassing 10 images for each of the 26 signs considered. The segmentation of the images was accomplished using Otsu's method, a widely used algorithm for automatic thresholding in image processing. The application of PCA in real-time sign language recognition suggests an approach that leverages dimensionality reduction to extract essential features for effective classification. The use of a substantial dataset with 26 distinct signs indicates a comprehensive evaluation of the proposed methodology. Furthermore, the adoption of Otsu's method for image segmentation underscores a well-established technique for image thresholding, contributing to the overall accuracy and efficiency of the sign recognition system presented in their work. The majority of the studies referenced, as mentioned in [8], primarily focused on employing techniques such as pattern recognition and feature extraction. However, it has been observed that in many instances, relying on a system with a single feature is insufficient to achieve robust results. In response to this limitation, researchers have introduced hybrid approaches, combining multiple methodologies or features. These hybrid methods aim to address the complexities and variations present in sign language recognition more effectively, offering a more comprehensive solution to the challenges posed by this intricate field.

A. Nandy and colleagues, as outlined in their work referenced as [9], employed Blended techniques Using Euclidean distance and K-Nearest Neighbor (KNN) for classifying motions based on Features of an oriented histogram. However, it apparent as it exhibited poor performance, especially in cases involving similar gestures. In their work, K. Manjushree and team [10] focused on one-handed sign categorization using feature matching and an oriented gradient histogram. On a different note, S. Kanade and collaborators [11] developed a system using a custom dataset, incorporating Principal Component Analysis (PCA) features and Support Vector Machines (SVM). Their system demonstrated commendable accuracy levels. A. Sahoo, as detailed in their work [12], installed an Indian Sign Language recognition system (ISL) encompassing both hand character signs. Geetha. M and colleagues, as documented in their research [13], employed B-Spline approximation for matching the shapes of (ISL) alphabets as well as numerical. Q. Chen and team [14] introduced a technique for identifying hand gestures utilizing all the algorithms. Additionally, they elucidated the use of random context-free grammars to achieve comprehensive gesture recognition. This approach involves leveraging specific features and machine learning algorithms for accurate gesture identification, while the incorporation of random context-free grammars enhances the overall system.

3. Proposed Work

Regarding the field of sign language interpretation, development of a highly accurate system necessitates the availability of efficient and robust data. In this particular study, the authors addressed the challenges associated with sign detection and classification by utilizing a custom-built dataset. The data flow through various stages in the development as depicted in Figure. 2, includes all the important features

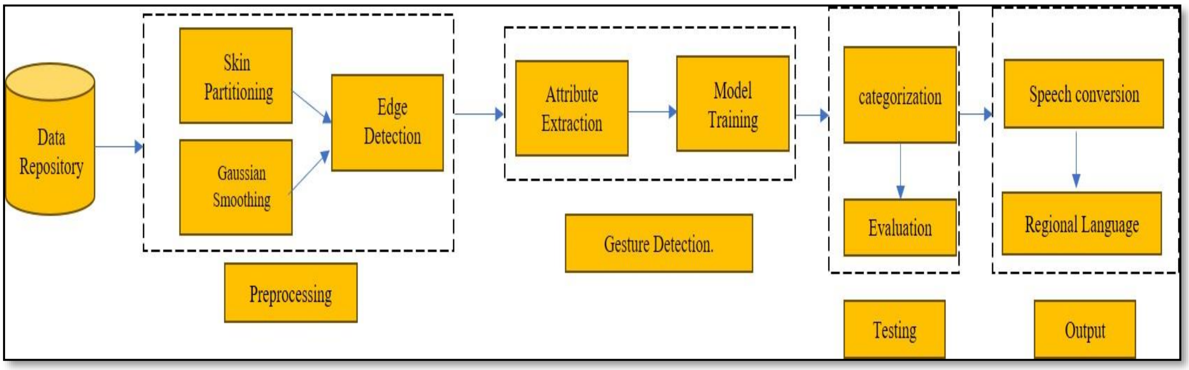

Figure 2: flowchart illustrating the study's suggested methodology.

Dataset: The foundation of the system is a custom-built dataset, carefully curated to encompass a diverse range of signs and gestures relevant to the targeted sign language.

Image Acquisition: This stage involves capturing images, possibly through a variety of devices like cameras or webcams, to represent the gestures in the dataset accurately.

Data Pre-processing: Pre-processing techniques are used to improve the data's quality and usefulness for analysis. This may involve tasks such as noise reduction, image enhancement, and normalization.

Feature Extraction: The system extracts relevant features from the pre-processed data to capture essential information that aids in distinguishing between different signs. Feature extraction is a critical step in the recognition process.

Speech Conversion: The data's are trained based on the image acquisition and it shows in regional languages

Sign Classification: Finally, the extracted features are utilized in the classification stage, where the system employs a model to categorize and recognize the sign language gestures. This step contributes to the overall precision as well as effectiveness system of recognition. By carefully managing the data flow through these stages, the authors aim to optimize the recognition process, ensuring that the system is well-equipped to handle real-time sign language interpretation for users. The use of a custom-built dataset enhances the system's adaptability and performance, contributing to its utility in practical applications.

3.1 User-centric dataset compilation

The process of dataset collection plays a pivotal role in research across various domains, serving as a foundational element for the advancement of machine learning or deep learning models. Despite its significance, dataset collection is rife with challenges. One major hurdle encountered during this phase was the absence of standardized datasets specifically tailored for Indian Sign Language. Consequently, as an integral aspect, It endeavoured to create a dataset by hand to address this issue comprehensively. This manual dataset creation initiative was undertaken to surmount the lack of readily available resources, aiming to contribute to the broader objective of developing robust system.

To initiate the dataset creation process, we utilized a webcam to capture videos featuring a diverse range of signs. The quality of the captured images and the mitigation of background noise were crucial considerations, highlighting the importance of the camera's positioning. Achieving optimal positioning is essential for maintaining picture quality and reducing unwanted environmental disturbances.

Two separate methods were used to take pictures in order to add variances to the collection.

The default method, incorporating skin segmentation to enhance image quality. This method is suitable for scenarios with a plain color background, providing a standardized approach to image capture and contributing to the diversity of the dataset.

In this approach, we implemented the averages. This required identifying each new object that appeared after a few early frames as the foreground and considered those frames to be the background. This method significantly simplified the extraction process. To ensure the model's adaptability to diverse scenarios, we created the dataset by incorporating both the skin segmentation and running averages approaches.

From the live footage, the signs were translated into frames, and these frames were taken out using a limit for pixel values. To streamline the pre-processing phase and optimize computational resources, The resolution has 255*255 Each sign folder in the dataset contained approximately 1200 images of the respective sign, resulting in a total of 43,200 images

For visualization, The pictures in the dataset can be observed in Figure. 2 (refer to Figure. 3). This meticulous dataset creation process ensures a robust foundation for the model's training and evaluation, enabling it to effectively recognize a diverse array of sign language gestures which is then related to voice system which can be .wav format where the image is recognised immediately the voice system is enabled and it gives the voice for that particular character, words, number or sentences.

3.2 Data Preprocessing

During this phase, the image undergoes preparation to facilitate the subsequent processes of feature detection and extraction. One key aspect of this preparation is ensuring uniformity of scale across all images by maintaining consistent dimensions.

To elaborate, the following steps are typically involved:

Resizing Images: The dimensions of each image are adjusted to fixed size. This resizing helps in creating a consistent format for all images, making it easier for algorithms to analyze and extract features. It ensures that the information within each image is represented uniformly.

Normalization: Picture pixels may be normalized to a standard range. This includes putting the pixel values on a standard scale, often in the range of 0 to 1 or -1 to 1. It helps in achieving consistent intensity levels, which can be crucial for accurate feature extraction.

Aspect Ratio Adjustment: In some cases, it might be necessary to adjust the aspect ratio of the images to prevent distortion. This ensures that the content of the images is not skewed or stretched, maintaining the original proportions of objects within the images.

Grayscale Conversion: Depending on the requirements, images may be converted to grayscale to simplify the analysis. Grayscale images contain only intensity information, which can be beneficial for certain types of feature detection algorithms.

Data Augmentation (Optional): In some cases, additional variations of the images may be generated through techniques like rotation, flipping, or cropping. This process, known as data augmentation, can enhance the model's ability to generalize patterns and improve overall performance.

By performing these preprocessing steps, the images are made more amenable to subsequent feature detection and extraction algorithms. Standardizing the scale and format ensures that the model can effectively learn and recognize patterns

In the default setting, the video frames captured are initially converted to HSV, specifically the images acquired against a plain background. This choice is made because it tends to differ, facilitating easier extraction. Subsequently utilizing calculations to filter pixels corresponding to the skin color. It is a binarization process, simplifying it to a binary format, and a blurring technique is employed to eliminate unwanted noise. The next step involves extracting the maximum contour from the result, under the notion that the contour has the greatest area. To refine the output and address potential errors, a median filter is applied. Additionally, morphological operations are employed for hand detection system. These operations aid in smoothing, refining, and optimizing the shape and structure of the detected hand contours, contributing to a more reliable and precise identification of the hands in the given image.

3.3 Feature Extraction Techniques

Feature extraction is a crucial step in image processing and computer vision where relevant information or distinctive characteristics are extracted from raw data, such as images, to represent and describe the objects or patterns of interest. In the context of image analysis, the process involves transforming the raw pixel values into a set of meaningful and compact features that can be used for subsequent tasks like classification, recognition, or analysis.

Color Histograms: This method quantifies the distribution of color intensities in an image. It's particularly useful when color information is critical, such as in identifying objects based on their color characteristics.

Local Binary Patterns : By comparing each pixel with its neighbouring pixels, Local Binary Patterns effectively characterizes the texture of an image and is useful for texture analysis and recognition..

Scale-Invariant Feature Transform: it highlights important areas in a picture. that are invariant towards scaling, rotation and translation. These key points serve as robust features for object recognition.

Principal Component Analysis: It transforms the data into a new coordinate system, reducing its dimensionality.. It's commonly used to compress information while retaining the most significant features.

Convolutional Neural Networks (CNN): CNNs automatically extract hierarchical features from images in deep learning. They consist of convolutional layers that extract features at different levels of abstraction.

Edge Detection: Techniques like the Canny edge detector identify boundaries and edges in an image, which can be important features for object recognition.

Texture Analysis: Methods like Gabor filters can be applied to extract texture information from images, helping in tasks that require understanding surface characteristics.

Shape Descriptors: Features like the area, perimeter, and circularity of shapes in an image can be used for shape-based recognition.The specifics of the data and the analysis's objectives determine which feature extraction technique is best. After feature extraction, the obtained features are often used as input for machine learning algorithms or other analytical tools to perform tasks like object recognition, classification, or segmentation. This stage encompasses the Building a Case of Visually Words (BOVW) through series of steps, including extracting features from images, clustering these features, constructing a codebook for the particular model, and generating histograms [15]), is widely employed in image classification. Analogous to BOW, where the The repetition of textual terms is counted to generate keywords and a frequency histogram, Instead of words, these characteristics serve as the visual "words" in the model. The resulting frequency histogram encapsulates the distribution of these visual words within the image.

The constructed vocabulary and associated histograms facilitate the prediction of the category of a comparable image. By comparing the frequency histograms, the model can discern similarities between images and make predictions based on the shared visual characteristics.

3.4 Result Indicator

The algorithm automatically predicts class labels, which are then converted and returned as numerical variable—into text and speech from the speech module to different regional languages of India. Fig 3 shows the proposed system. Better communication and user convenience are the goals of this. After the classifier has determined the label, it is sent as a key to a dictionary, which returns the value, which is the associated sign. The user is subsequently presented with this. Pyttsx3, a Python modules are used for conversion. Threading is done because it slows down the live image by causing the process very slow. As result, simultaneous achievement of sign prediction and text-to-speech translation is possible. This guarantees uninterrupted playback of the sound at all times.

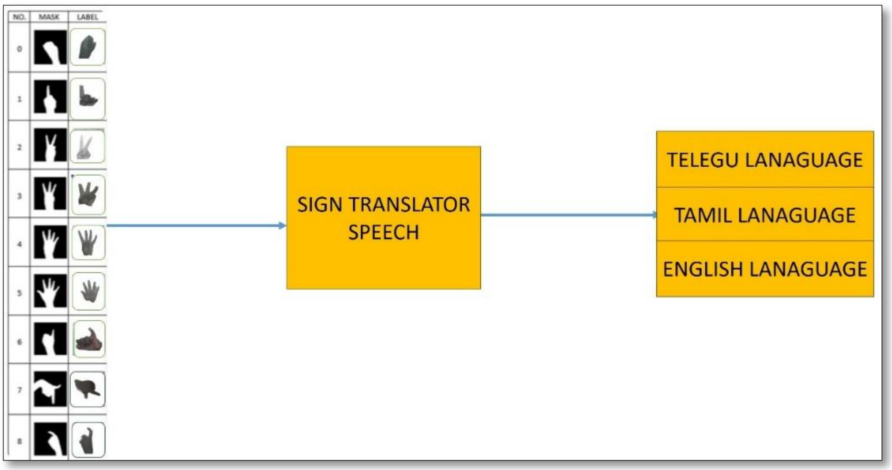

Fig 3. The proposed system

4. Experiment and results

The dataset consists of two sets. The training set makes up 90% of the entire data collection, while testing is done with the remaining 10%. SVM and CNN, the two classifiers, have both produced excellent accuracy on the images. Table 1 describes SVM, CNN and Speech Conversion Accuracy. CNN has outperformed other networks with fewer features. The system is capable of identifying 39 signals, which consist of 26 alphabets, 10 digits, and 3 words. However, a few adjustments might produce even better results, the current results are encouraging. 99% accuracy has being proven.

Table 1 SVM, CNN and Speech Conversion Accuracy

Label	SVM (%)	CNN (%)	Exact Speech Conversion %	Label	SVM (%)	CNN (%)	Exact Speech Conversion %	Label	SVM (%)	CNN (%)	Exact Speech Conversion %
0	100	100	100	**C**	100	100	100	**P**	100	100	100
1	100	100	100	**D**	100	100	100	**Q**	100	100	100
2	100	100	100	**E**	98	99	100	**R**	99	99	100
3	99	99	100	**F**	99	100	99	**S**	100	100	100
4	100	100	100	**G**	99	100	100	**T**	100	100	97
5	100	100	99	**H**	100	100	100	**U**	100	100	100
6	100	100	100	**I**	99	100	100	**W**	100	100	100
7	100	100	100	**J**	100	100	99	**X**	100	100	100
8	99	100	98	**K**	100	100	100	**Y**	100	100	96
9	99	100	100	**L**	99	100	100	**Z**	100	100	100
10	100	100	100	**M**	100	100	100	**Welcome**	99	99	99
A	100	100	100	**N**	99	100	99	**Hello**	98	100	98
B	100	100	100	**O**	100	100	100	**Thank You**	99	100	99

4.1 CNN AND SVM performance

Accuracy, which is just the ratio of correctly predicted observations to all observations, is the most basic performance metric. In terms of real positives (X), real negatives (Y), incorrect positives (Z), and incorrect negatives (M), the accuracy formula can be written as follows. [17]

$$\text{Accuracy Rate} = \frac{X+Y}{X+Y+Z+M}$$

Precision is defined as the ratio of correctly predicted positive observations to the total number of positive data. The recall is defined as the ratio of correctly predicted positive labels to the total number of positive labels. The F2 score is the weighted average of precision and recall.

$$\text{Predicate Accuracy} = \frac{X}{X+Z}$$

$$\text{Recall Accuracy} = \frac{X}{X+M}$$

$$\text{F2 Score} = 2*\frac{(\text{Recall}*\text{Precision})}{\text{Recall} + \text{Precision}}$$

The formula for the F2 score is shown above.. The table 2 shows the accuracy calculated based on the above calculation.

Table2: Accuracy

CNN	SVM
99.89%	99.9%

5. Conclusion

The research offers a novel methodology that uses SVM and CNN to classify and recognize Indian Sign Language signals and translate three particular words into voice. The main goal aims to have a more instantaneous grasp of usefulness, which permits system mobility. The process involves generating a distinct input set, background dependence issue, and turning the system into one that is rotationally invariant. Static alphabets and numbers have been successfully taught to the system with a 99.5% accuracy rate which in term is being given to the speech conversion module which also shows 99.6% To achieve a more potent real-time application framework, additional indicators in the future, additional languages from various nations may be added to the dataset. For both identification system, various technique can be expanded to produce simple words and expressions. Accelerating the solution to applications that operate in real time is response time.

References

[1] Viswanathan Daleesha M, Idicula Sumam Mary. Recent developments in Indian sign language recognition: an analysis. Int J Comput Sci Inf Technol 2015;6(1): 289–93.

[2] Bantupalli Kshitij, Xie Ying. American sign language recognition using machine learning and computer vision. Master of Science in Computer Science Theses 2019; 21.

[3] Shivashankara S, Srinath S. A comparative study of various techniques and outcomes of recognizing American sign language: a review. In: International Journal of Scientific Research Engineering & Technology (IJSRET); 2017. ISSN 2278 – 0882. 6(9).

[4] Adithya V.P.R Vinod and Usha Gopalakrishnan "Artifical Neural Network based method for Indian Sign Language Recognition " IEEE Conference on Information & Communication Technologies(ICT), IEEE, 2013.

[5] Athira K, Sruthi CJ, Lijiya A. A signer independent sign language recognition with Co-articulation elimination from live videos: an Indian scenario. J King Saud Univ Comput Inf Sci 2022;34(3):771–8.

[6] Singha J, Das K. Recognition of Indian sign language in live video. Int J Comput Appl 2013;70(19):17–22.

[7] Shreyashi Narayan Sawant, "Real time Sign Language Recognition using PCA", International Conference on Advanced Communication Control and Computing Technologies(ICACCCT), IEEE 2014

[8] Rokade Yogeshwar I, , et alJadav Prashant M. Indian sign language recognition system. In: International Journal of Engineering and Technology July; 2017.

[9] Nandy Anup, Prasad Jay Shankar, Mondal Soumik, Chakraborty Pavan, Nandi Gora Chand. Recognition of isolated Indian sign language gestures in real time. In: International Conference on Business Administration and Information Processing; 2010.

[10] Manjushree K, Divyashree. Gesture recognition for Indian sign language using HOG and SVM. International Research Journal of Engineering and Technology 2019;6 (7).

[11] Kanade Sudhir S, Deshpande Padmanabh D. Recognition of Indian sign language using SVM classifier. International Journal of Scientific Research and Development 2018;2(3).

[12] Sahoo Ashok Kumar, Kumar Ravulakollu Kiran. Vision based Indian sign language character recognition. J Theor Appl Inf Technol 2014;67(3).

[13] Geetha M, Manjusha UC. A vision based Recognition of Indian sign language Alphabets and Numerals Using B-Spline Approximation International Journal on Computer Science and Engineering (IJCSE). 2012.

[14] Chen Q, Georganas ND, Petriu EM. Hand gesture recognition using Haar-like features and a stochastic context-free grammar. IEEE Trans Instrum Meas 2008;57 (8):1562–71. https://doi.org/10.1109/TIM.2008.922070.

[15] Sivic J, Zisserman A. Video Google: a text retrieval approach to object matching in videos. In: null. IEEE; 2003. p. 1470.

[16] Bay Herbert, et al. SURF: speeded up robust features. In: European Conference on Computer Vision (ECCV); 2006.

[17] Shagun Katoch, Varsha Singh , Uma Shanker Tiwary, Indian Sign Language recognition system using SURF with SVM and CNN. 2590-0056/© 2022 Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).