

Sign Language Translator to Text & Speech Using Machine Learning

Chirantan Dev¹, Nishu Sinha², Aditya Raj³, Ms. Sikha Singh⁴

^{1,2,3} Student, Department of ECE, Noida Institute of Engineering and Technology, Greater Noida, India

⁴ Assistant Professor, Department of ECE, Noida Institute of Engineering and Technology, Greater Noida, India

Abstract—Deaf and mute people can only express their needs and emotions through sign language. The majority of non-deaf-mute people do not comprehend sign language, which makes it difficult for deaf-mute persons to communicate in social situations. In recent years, there has been a lot of interest in sign language interpretation applications and systems. We examine research on machine learning, image processing, artificial intelligence, and animation tools for sign language detection and interpretation in this study. Illustrations are provided for the two reverse sign language interpretation procedures. Recent studies on translating sign language to speech and text using lip reading, hand gestures, and facial expression interpretation are covered in this paper. This work presents a real-time machine learning-based sign language translator that converts hand motions into audio and text using real-time machine learning. We have used computer vision techniques and Convolutional Neural Networks (CNNs) to recognize hand motions.

Keywords – CNN Algorithm, Gesture Recognition, OpenCV, Sign Language, YOLO

I. INTRODUCTION

The most expressive method for deaf and mute persons is sign language. Learning sign language is not ingrained in the culture of non-deaf-mute individuals, even if they wish to assist those with disabilities. Consequently, those who are deaf or dumb live apart from society. However, handicapped persons must interact with others in their social lives. Sign language recognition is therefore one of the most important and rapidly expanding fields of study. Numerous techniques for recognizing sign language have recently been developed to assist the deaf and dumb. A combination of hand movements, mouthing cues, and facial expressions make up sign language[1].

There are numerous sign languages in use worldwide.

Approximately 300 distinct sign languages are in use worldwide. The reason for this is because sign languages were organically developed by people from different ethnic groups[2].

Perhaps India does not have a common sign language.

There are dialects and lexical variations in Indian Sign Language in many areas of India. Nonetheless, there have been recent attempts to standardize Indian Sign Language (ISL) [4]. The system may be trained to understand gestures and convert them into voice and text. The system efficiently and precisely classifies hand movements to enable communication between vocal and deaf-mute people. The gesture name for the recognized sign is also spoken and shown. To identify the hand gesture made by the signer in real time, the project first uses a webcam to record hand motions. It then preprocesses and segments the acquired frame. Additionally, a vocal form is created from the detected sign, expanding the project's potential applications. We have two communication projects underway[5]. The two groups of people that currently translate sign language are voice-to-sign language translators and sign language-to-voice translators. Sign-to-voice systems usually use either sensor data or images, and sometimes they even use both. The idea of employing gloves with sensors built into them to recognize ASL signs has already been proven.

II. RELATED WORKS

A. Hand Gesture Recognition

New techniques, devices, models, and algorithms have emerged as the field of hand gesture detection has advanced quickly. A mobile computing device-based system that offers the technology for automatic

conversion of the Indian Sign Language system into English speech is the subject of this IEEE study by Harish, N. and Poonguzhali, S. et al. [3], which enables two-way communication between the general public and people with vocal impairments. Because this technology operates on the Gesture Recognition principle, it may soon be utilized to help people who do not comprehend sign language communicate with each other. Using a built-in mobile camera, the device detects and captures motions proposed by Lalithaanjana K. et al. [5]. The recorded gestures are then analyzed using algorithms such as Contour Extraction, LargeBlob Detection, Flood Fill, and the HSV model (Skin Color Detection). The traditional alphabet (A–Z) and numerical values can be indicated with one-handed signs (0–9) that the system can identify. According to S. S. Rautaray and A. Agrawal et al. [6] in "Vision-based hand gesture detection," A sign language translator is the best option for enabling deaf/mute people to use technology to communicate effectively in different languages. It consists of two gloves equipped with the necessary sensors and smartphone software that uses Lebanese Sign Language to translate hand gestures into spoken phrases. To generate precise findings, artificial neural networks are essential.

B. A Sign Language Interpreter Employing Neural Networks and Sensor Gloves

Smart gloves use a low-power printed circuit board and multiple sensors to translate different sign language gestures into the relevant text. Punsara KKT et al. [7] provide an example of how every letter in American Sign Language has a distinct combination that is transmitted over Bluetooth to a smartphone or computer screen. A particularly constructed solution, the circuit board is limited to using the 26 alphabets. Sensor gloves, which feature two sensors to detect the tilt and rotation of the hand and five sensors on each finger to measure the tension between the knuckle and the finger's initial joint, are used in the training method to create a neural network. In Sood Anchal et al. [8], the neural network that was built is used to map the data to the proper character. The system recognizes the twenty-four alphabets and two punctuation marks of the English language. In real-time, the technology translates hand gestures into the corresponding letters or numbers.

C. Identifying Emotions in Speech through Deep Learning Methods.

According to S. Padmavathi and colleagues [11] IEEE Access An essential yet difficult aspect of HCI (human-computer interaction) is the identification of emotions from speech signals. Several methods, such as popular speech analysis and classification algorithms, have been used in the literature on speech emotion recognition (SER) to extract emotions from signals. As an alternative to conventional machine learning methods, deep learning approaches have just lately been introduced. This study by Navuluri Sainath, A. et al. [12] provides an overview of Deep Learning approaches for speech-based emotion recognition and examines some current studies on the topic. The evaluation covers the databases that were used, the emotions that were retrieved, the advancements in speech emotion identification, and its limits. To identify the emotions, we are using a variety of categorization techniques. The audio features MFCC, MEL, chroma, and Tonnetz were employed in addition to Support Vector Machine and Multilayer Perception. The emotions that these 13 models have been trained to display include: serenity, neutrality, surprise, delight, sadness, anger, fear, and disgust. The accuracy was 86.5% when we tested it using the input audio.

III. PROPOSED METHODOLOGY

The main objective of this project is to recognize the movements and display the relevant words. The gesture is initially recorded using a camera and a posture estimation library. The Tensor-Flow tool uses a posture estimation technique to evaluate the image captured by the camera. It shows how the posture estimate library created the skeleton superimposed on the image and how the webcam is understanding the image. The values from [2] the produced skeleton are used to create the data set, which is a collection of the coordinate values of the skeleton's endpoints. These data are added to the computer and given the proper labels to forecast when the input will be taken.

A.) Workflow of the proposed methodology

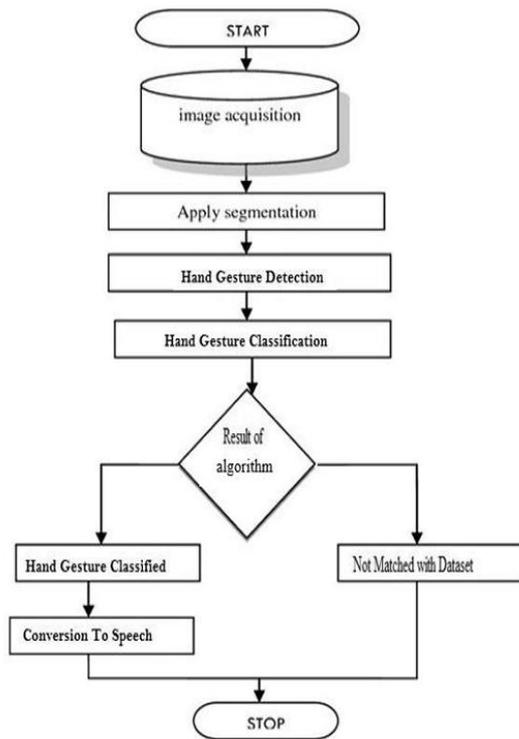


Fig 1: Methodology for Converting Sign Language to Text and Speech Model

The model architecture is shown here. This includes learning, identifying, and categorizing hand motions. This [1] will illustrate the procedure of taking an image and converting it into a voice. The relationship between picture data and the model is depicted in the architecture diagram below.

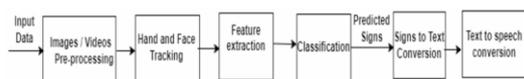


Fig 2: The process of translating speech to sign language

Creation of Neural Network: A deep artificial neural network known as a multilayer perceptron (MLP) uses the input layer as the first layer to receive signals, and the output layer as the last layer to forecast the input's class. The actual computational engine [7] of the MLP is comprised of an arbitrary number of hidden layers situated between these two levels. Each of the six hidden layers in the suggested system has 100 perceptrons. The neural network is created by adjusting the weights and biases using the Backpropagation Algorithm, which lowers error. Since there are several classes, 96 perceptrons are added to the output layer to perform multiclass classification. The softmax function is used by this output layer to increase accuracy. A vector of logits

or feature variables is mapped to a posterior probability distribution using the softmax function. As a result, all of the output perceptrons add up to one. When the classes are mutually exclusive, this function is helpful. The formula for the softmax function is provided by

$$y_c = \frac{e^{z_c}}{\sum_{d=1}^n e^{z_d}} \quad \text{for } c = 1, 2, \dots, 96$$

where $n = 96$ and z_c is the value of the c th perceptron in the output layer.

Hand Detection using YOLO: The region of interest (ROI) associated with the hand gesture in each video frame is found using the YOLO technique. YOLO formulates object detection as a single regression problem that predicts bounding box coordinates (x,y,w,h) , objectness score P_{obj} , and class probabilities $P(\text{class} | \text{object})$ directly from the input image. For a given class, the total confidence is provided by:

$$P(\text{class}_i) \cdot P_{obj} = P(\text{class}_i | \text{object}) \cdot P_{obj}$$

Integration of CNN: After Hand detection it is passed through a convolution neural network and the convolution operation in each layer is defined as

$$y = \sigma(W * X + b)$$

where X is the input feature map, b is the bias term, σ is a non-linear activation function, and W stands for the learned filter weights.

The CNN's last layer outputs a probability distribution across the collection of gesture classes using a softmax activation:

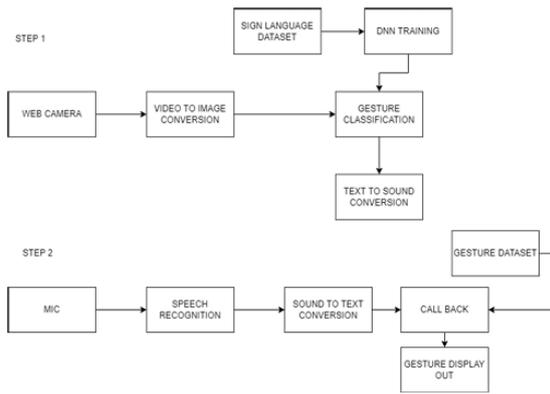
$$\text{Softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}$$

Where K is the total number of gesture classes.

These equations work together to create a strong pipeline for translating sign language, which allows the system to efficiently recognize, categorize, and translate hand signs into their equivalent written representations.

We must use these techniques as we are converting both text-to-speech and speech-to-text.

HAND GESTURE TO VOICE: One helpful piece of contemporary technology that is used for both human-computer interaction and impaired individuals is a hand gesture recognition system [13]. The system has used a variety of methods and technology to achieve its goal.



VOICE TO HAND GESTURE: Using a Microphone for Text Translation and Speech Input Allows Ambient Noise Adjustment [19] Due to the fluctuating nature of ambient noise, we must allow the software some time to adjust the energy threshold of the recording to match the amount of external noise. Spoken speech can be converted into text using Google Speech Recognition.

IV. IMPLEMENTATION

The suggested study's author employed the SVM technique, but Python SVM is inaccurate at recognizing hand motions. For this reason, we are training hand gesture photographs using a deep-learning Convolution Neural Network. The taught hand motions from the webcam can then be predicted by the trained model[12].

Step 1: The skin section of the image must be separated from the rest of the image because the latter could be interpreted as noise in connection to the character classification problem.

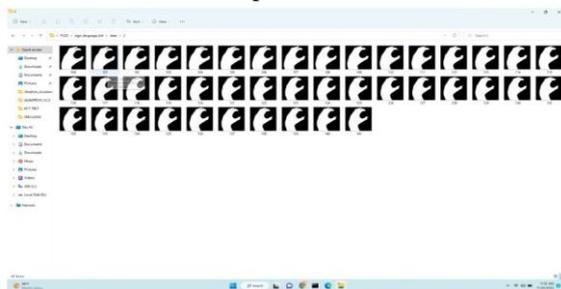


Fig 3: Pre-processed Image

Step 2: The second step is to extract relevant features from the skin-segmented images that could be crucial for the following stages of learning and classification.

Step 3: As previously mentioned, this step entails feeding the training algorithm the extracted features. Classification is then done using the learned models.

Implementation models: 1Sign Language in India FRCNN Images of double-handed gestures in Indian Sign Language, comprising both letters and numerals, were utilized as our data set. The FRCNN model was given data. The Fast R-CNN detector generates region suggestions using an Edge Box-like algorithm, just like the R-CNN detector [13]. Unlike the R-CNN detector, which resizes and reduces area suggestions, the Fast R-CNN detector processes the full image. While an R CNN detector must categorize each region, Fast R-CNN pools CNN characteristics related to each area proposition. Since computations for overlapping regions are pooled, as supplied by the object identification process in YOLO, which is conducted as a regression problem, the Fast R-CNN detector is more efficient than R CNN. The YOLO approach uses CNN to quickly identify objects. As the name suggests, the method uses a neural network to recognize objects with just one forward propagation. This illustrates that all of the image's predictions are produced by a single algorithm run. The CNN predicts bounding boxes and several class probabilities at the same time. The YOLOv3 version is used in this project.

2. CNN in American Sign Language We acquired the American Sign Language data set from Kaggle, where each hand represented the English letters and numbers. Convolutional neural networks are one Deep Learning technique that can recognize and differentiate between different objects and elements in an input picture. Compared to other classification methods, a ConvNet requires far less preparation. While filters are manually designed using rudimentary techniques, ConvNet can learn these filters and features. The convolution operation's goal is to extract high-level features, such as edges, from the input picture.

3. For American Sign Language, use YOLO. We used the YOLO model to translate sign language to achieve the project's goal and improve the outcomes. YOLO is an acronym that stands for "You Only Look Once." This technique finds and identifies many things in a picture. The class probabilities of the discovered images are ascertained by regression analysis. CNN is used in the YOLO technique to identify objects rapidly. As the name implies, the technique detects objects with just one forward propagation across a neural network. This illustrates that a single algorithm run is used to do prediction throughout the full image. The CNN simultaneously

predicts several class probabilities and bounding boxes.

4 Personalized Media Pipe for Sign Language The classes are specified individually for each hand gesture, and the training data set that is fed into the Media pipe structure in this model contains a range of words that people use in everyday conversations. These tasks may now be completed quickly, accurately, and uniquely with MediaPipe, an open-source framework designed specifically for complex perception pipelines employing accelerated inference (such as GPU or CPU). This is a particularly difficult problem that requires the simultaneous inference of several interdependent neural networks to integrate them all in real time into a semantically consistent end-to-end solution. To address this issue, MediaPipe Holistic now offers a new, state-of-the-art human pose topology that opens up new applications.\

Capturing the Hand Gesture: One method for deploying the camera to record hand motions in real time is computer vision. Greyscale was applied to the photograph. The grayscale technique known as the Canny Edge Detection Algorithm is used here to help transform RGB photos into a format that a machine can understand. Computer vision is the process of giving machines intelligence so they can view pictures or movies and extract information from them in the same way that people can. With the use of the Open CV framework, a Python package that facilitates the acquisition of the system's camera input image.

Classifying the Hand Gestures: Ultimately, a trained model is used to classify the webcam image, and the gesture is translated into the appropriate text. The model file, which includes the pre-processed pictures of the motions taken from the Kaggle website, is compared to the image produced using Open CV. To categorize the differences between the acquired image and the dataset images, the CNN algorithm is utilized. The resulting image is tested against the model file with the aid of this algorithm.

Audio to Hand Sign Conversion: Gtts are used to convert text to speech. This speech recognition service was created by Google as an operating system program for screen reading purposes. It offers programs for reading text aloud from the screen in a variety of languages with a wide range of support. For ease of comprehension, the displayed sentence has been changed to voice.

Converting Audio to Hand Sign: Using a Python speech recognition library,[19] computer software can recognize words and phrases in spoken language and translate them into text that is readable by humans. Afterward, a Python script is used to map the text to the appropriate gestures. And the outcome of converting a speech that goes, "Hello, how are you?" into a graphic is displayed below.

V. EXPERIMENT RESULT

The proposed technology reduces hardware expenses by utilizing software-based artificial intelligence techniques. An artificial intelligence model first records a hand signal using the webcam categorizes the word that corresponds with the sign, and then converts the audio. Similarly, the proposed system may convert voice to sign using Python sound and the OpenCV framework.

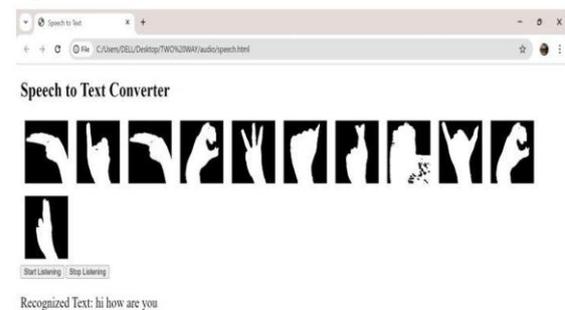


Fig 4: Speech to Hand Gesture

This innovative SER technique has significant promise for speech recognition of nuanced emotional expressions due to its capacity to bridge the gap between audio and visual data representations. It addresses the need for feature-rich representations and advanced modeling techniques to increase the accuracy of emotion recognition.

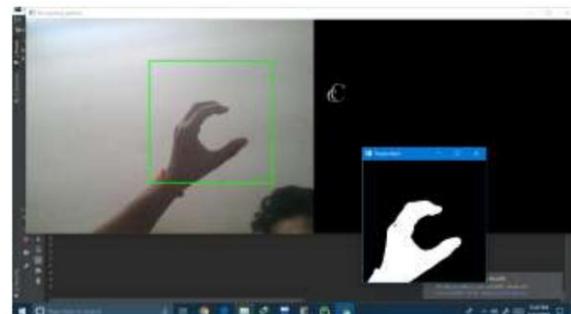


Fig 5: Detecting sign C

VI. COMPARATIVE ANALYSIS

To establish a connection between research projects and practical implementations, we are showcasing

new studies and real-world applications in the sign language recognition sector. Additionally, we present a proposal for integrated functionalities that encompass all distinct focal directions in recent research. Communication for the deaf and mute will be substantially improved by covering these integrated research directions in real application(s) used in our daily lives. Lastly, we offer a list of unexplored study avenues for further investigation.

S. No.	Study	Dataset(s)	Technique	Skey Performance Metrics
1.	Tariq MU (2020) D-talk	1000 photos make up the dataset. Each sign features fifty pictures in various lighting conditions and orientations.	Hidden Markov models	Accurac y= 60%
2.	Lalithanjanana K (2021)	Dataset of static words and letter hand signals. Each sign had 20 photos taken in various settings, including various lighting conditions, backdrops, and viewpoints.	Deep Learning	Accurac y= 85%
3.	Prathap BR (2020) Hand kinesics	Dataset of Indian words with their corresponding animation	NLP, SVM, polarity detection method Google API	Accurac y= 80%
4.	Meka PK (2023)IC ISC	An ASL dataset including 8000 images	OpenCV, CNN	Accurac y= 90.56%
5.	Proposed Work	1820 letter pictures and 700 number images make up the dataset. Various letter and number variations were displayed.	CNN, YOLO, Media Pipe, SER	Accurac y for letters and numbers recognition= 97.04 and 98.9%

A comparison of the aforementioned studies on the translation of speech and text to sign language is shown in the above Table. These studies' methods, datasets, sign languages, and results are shown. Machine learning techniques, text pre-processing, and natural language processing were applied. Various animation tools were employed. These

studies employed datasets for the following sign languages: American, Indian, German, Filipino, Urdu, Sindhi, Arabic, Punjabi, Gujarati, and Bangla. The highest accuracy of 80-90% was attained by [16] on this dataset, out of all the previously listed speech/text-to-sign language translation research. The only research that focused on signs for alphabets, numbers, words, and sentences was in [14–16].

VII. CONCLUSION

In conclusion, the MFCC-based CNN approach for Speech Emotion Recognition (SER) proposed in the abstract provides a novel and exciting avenue for human-computer interaction and emotional intelligence applications. This method uses Mel-frequency cepstral waves to combine deep learning with traditional audio analysis techniques MFCC to extract features and transform them into MFCC images that can be used by Convolutional Neural Networks (CNNs). This combination of visual and auditory data representation addresses the need for more complex modeling techniques and richer feature sets in speech recognition (SER) by offering the potential to capture nuanced emotional expressions in speech. As a result, the proposed approach holds great promise for real-world applications like virtual assistants, mental health monitoring, and sentiment analysis, where it is essential to identify and respond to human emotions communicated through speech. Future efforts to expand the dataset and improve the customized CNN architecture are anticipated to further develop and validate this innovative SER methodology, which could result in notable improvements in the utility and accuracy of emotion detection.

When the suggested system was put into practice using k-NN, neural networks, and decision tree classifiers, it was discovered that k-NN had the best accuracy, 97%. The accuracy of the Decision Tree and Neural Network was 76%.

When dealing with limited data and several classes, the Decision Tree Classifier performs poorly. Overfitting occurs when the decision tree is constructed because of noise, which introduces several inconsistencies within the branches. The algorithm keeps lowering the training set error at the expense of higher test set errors, which leads to the overfitting issue [. Consequently, this method has demonstrated a reduced accuracy rate. Overfitting

occurs when neural networks are built with insufficient training data. The resulting neural network can categorize the training dataset with very high accuracy, but it cannot classify any test records other than the training records due to the lack of a large and varied training dataset. The system does not generate a generic solution; instead, it is taught to function for a particular dataset. As a result, when the dataset is less, the accuracy decreases.

The best results have been obtained with k-NN out of the three approaches.

Technique	Accuracy
Neural Network	76.0274873413
Decision Tree Classifier	76.3760110475
KNN	97.0474123759

VIII. FUTURE SCOPE

The majority of current research on the recognition of sign language uses letters and numbers. Few research employed a static restricted collection of words or sentence signals. Studies on sign language recognition systems that use a large number of words and phrases from language dictionaries are nonexistent. Additionally, recent studies have not focused on internet services for the deaf and mute. There are no current studies that have suggested ways to help deaf silent persons use online services. For instance, there is no research devoted to providing deaf-mute persons with internet shopping, banking, and other online services. Furthermore, establishing social networks to assist deaf-mute people is not a priority. One significant unresolved issue is the introduction of research solutions for creating online social societies that are appropriate for deaf-mute people.

Proposals for additional self-education and online reading of books, papers, and other materials for deaf-mute people do not now exist.

This technology may eventually be released as an Android app for smartphones, allowing deaf and mute people to interact and communicate with hearing people.

ACKNOWLEDGEMENT

The NIET Institute of Innovation and Technology is deeply appreciated by the authors for their assistance and resource provision, which made this research project much easier to finish. We also want to express

our sincere gratitude to everyone who made the "Sign Language Conversion to Text and Speech" project a reality. The diligent research and development team made significant contributions to the design and implementation of the unique system. We are grateful to the contributors who provided a range of sign language datasets, which enabled us to build a strong model training foundation. Sincere appreciation is given to the advisors and mentors for their crucial guidance throughout the project's development. We are particularly grateful to the open-source community for their contributions to the project's success. Our financing partners' financial support also played a significant role in the project's execution. We are appreciative of everyone and everything that helped make this collaborative effort, which has produced a novel method for sign language conversion.

REFERENCES

- [1] Fatma M. Najib "Sign language interpretation using machine learning and artificial intelligence" In: Neural Computing and Applications (2025), DOI: 10.1007/s00521-024-10395-9
- [2] Medhini Prabhakar, Prasad Hundekar, Sai Deepthi B P, Shivam Tiwari "Sign Language Conversion to Text and Speech".In:JournalofEmergingTechnologiesandInnovativeResearch(JETIR)(2022),DOI:d626/JETIR220738 0
- [3] Harish, N. and Poonguzhali, S., 2015, May. Design and development of hand gesture recognition system for speech impaired people. In 2015 International Conference on Industrial Instrumentation and Control (ICIC) (pp. 1129-1133). IEEE.
- [4] Utkarsh Jagtap, Vinayak Nangnurkar , Suhas Chalwadi , Neelam Jadhav " Indian Sign Language Recognition: A Survey" In: International Journal of Scientific Research in Engineering and Management (IJSREM) (2023), DOI: 10.55041/IJSREM20838
- [5] Sanmitra PR, Sowmya VS, Lalithanjana K (2021) Machine learning based real time sign language detection. Int J Res Eng Sci Manage 4(6):137–141
- [6] S. S. Rautaray, and A. Agrawal, "Vision based hand gesture recognition or human computer interaction: a survey," Artificial Intelligence Review, vol. 43, no. 1, pp. 1-54, 2015.

- [7] Swapna Johnnya, S Jaya Nirmalab “Sign Language Translator Using Machine Learning” In: ResearchGate publication” DOI: 10.1007/s42979-021-00896-y
- [8] PunsaraK.K.T,PremachandraH.H.R.C,ChanakaA.W.A.D, NimsiriAbhayasinghe “IoT Based Sign Language Recognition System” In: ResearchGate Publication (2020) DOI:10.1109/ICAC51239.2020.9357267
- [9] Sood, Anchal, and Anju Mishra. "AAWAAZ: A communication system for deaf and dumb." Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO), 2016 5th International Conference on. IEEE, 2016.
- [10] Ms. KAVITHA R, Ms. Varsha S, Mr. Bharaneedharan C “Two-Way Sign Language Translator for Visually Impaired People” In: International Journal of Scientific Research in Engineering and Management (IJSREM) (2024),DOI: 10.55041/IJSREM29475
- [11] Shangeetha, R. K., V. Valliammai, and S. Padmavathi. "Computer vision based approach for Indian Sign Language character recognition." Machine Vision and Image Processing (MVIP), 2012 International Conference on. IEEE, 2012.
- [12] Kotikalapudi Vamsi Krishna,Navuluri Sainath,A. Mary Psonia,"Speech Emotion Recognition using Machine Learning",2022 6th International Conference on Computing Methodologies and Communication (ICCMC)
- [13] Shinde, Shweta S., Rajesh M. Autee, and Vitthal K. Bhosale. "Real time two way communication approach for hearing impaired and dumb person based on image processing." Computational Intelligence and Computing Research (ICCIC), 2016 IEEE International Conference on. IEEE, 2016.
- [14] Saleh BM, Al-Beshr RI, Tariq MU (2020) D-talk: sign language recognition system for people with disability using machine learning and image processing. Int J Adv Trends Comput Sci Eng 9(4)
- [15] Amal H, Reny RA, Prathap BR (2020) Hand kinesics in indian sign language using NLP techniques with SVM based polarity. Int J Eng Adv Technol IJEAT 9(4):2249–8958
- [16] Meka PK, Parusu YR, Vaddi R. Real-Time Sign Language Detection Using OpenCV. InInventive Systems and Control: Proceedings of ICISC 2023 2023 Jun 15 (pp. 523–534). Singapore: Springer Nature Singapore.
- [17] S. H. Bani Hani and M. M. Ahmad, “Machine learning Algorithms for Ischemic Heart Disease Prediction: A Systematic Review,” Curr Cardiol Rev, vol. 19, no. 1, Jun. 2022, doi: 10.2174/1573403X18666220609123053.
- [18] Chen Jie,"Speech emotion recognition based on convolutional neural network",2021 International Conference on Networking Communications and Information Technology (NetCIT)
- [19] Ahire, Prashant G., et al. "Two Way Communicator between Deaf and Dumb People and Normal People." Computing Communication Control and Automation (ICCUBEA), 2015 International Conference on. IEEE, 2015.
- [20] Rishi K,Prarthana A,Pravena K S,S. Sasikala,S. Arunkumar,Two-Way Sign Language Conversion for Assisting Deaf-Mutes Using Neural Network,2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS)
- [21] Kumari, Sonal, and Suman K. Mitra. "Human action recognition using DFT." Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), 2011 Third National Conference on. IEEE, 2011.
- [22] A. Y. Satpute, A. D. Bhoi, and T. Engineering, "ELECTRONIC SPEAKING SYSTEM FOR DUMB," vol. 6, no. 3, pp. 1132-1139, 2013.
- [23] Dhivyasri S, Krishnaa Hari K B, Akash M, Sona M, Divyapriya S, Dr. Krishnaveni V An Efficient Approach for Interpretation of Indian Sign Language using Machine Learning , 2021 3rd International Conference on Signal Processing and Communication | 13 – 14 May 2021 | Coimbatore.