International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

COPY RIGHT

ELSEVIER
SSRN

IJIEMR Transactions, online available on 12th June 2022.

Link : https://ijiemr.org/downloads/Volume-11/Issue-06

Title: SIGN LANGUAGE TRANSLATION SYSTEM BASED ON CONVOLUTIONAL NEURAL NETWORKS

volume 11, Issue 06, Pages: 1492-1500

Paper Authors: **Mrs. Ch. SatyaKeerthi. N. V. L, M.Tech, Ch. Dharma Teja, D.Mayusha, P.BeenaJahanavi, B. Lavanya**

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

SIGN LANGUAGE TRANSLATION SYSTEM BASED ON CONVOLUTIONAL NEURAL NETWORKS

[1] Mrs. Ch. SatyaKeerthi. N. V. L, M.Tech,

[2] Ch. Dharma Teja, [3]D.Mayusha, [4]P.BeenaJahanavi, [5]B. Lavanya

[1]Assistant Professor, CSE, PSCMRCET, Vijayawada, Andhra Pradesh

[2,3,4,5]Student, CSE, PSCMRCET, Vijayawada, Andhra Pradesh

Abstract

Humans engage with one another in order to communicate their ideas, thoughts, and opinions to the people around them. However, this is not the case for deaf-mute people. Deaf-mute people can communicate with each other via sign language. Conversation is possible for a deaf-mute character using sign language rather than acoustic sounds. motif in the other time of these paintings is to detonate a tool to detect the language of the signal, which offers conversation among human beings with speech impairment and ordinary human beings, thereby decreasing the conversation hole among them. Hand gestures play an important function in comparison to critical gestures (arm, face, head, and body), because they represent the user's opinions in hundreds of times less time. In today's work, a flex sensor-based totally completely without a doubt sincerely gesture recognition module has been developed to recognize English alphabets and a few phrases, and a Text-to-Speech synthesizer based totally completely without a doubt totally on HMM has been built to transform the corresponding text.

Keywords: Sign language; Machine learning; Hand Recognition; Convolutional Neural Network (CNN); Image Processing.

I. INTRODUCTION

For hearing-impaired people, sign language is the final natural medium and means of expression. People who are not deaf do not need to understand sign language in order to interact with deaf people in their daily lives. Deaf people are isolated as a result of this. However, if the computer could be programmed in such a way that it could convert sign language into text, the gap between ordinary people and the deaf community could be bridged. Indian Sign Language (ISL) uses every finger to represent each alphabet and gestures. ISL alphabets are acquired from British Sign Language (BSL) and French Sign Language (FLS). Most researchers in this issue focus on American Sign Language recognition (ASL) because of the truth that

maximum signs and symptoms and signs and symptoms in ASLs are one-exceeded and therefore the complexity is less. Another thrilling function of the is that ASL already has a standard database available for use. In contrast with ASL, Indian Sign language is based totally mostly on every finger and therefore an ISL popularity system is extra complex. A few research projects were carried out with the help of researchers within the ISL recognition. Currently, additional researchers are conducting research in ISL. This proposed system can interpret the many alphabets of Indian Sign Language, which can help to reduce noise and provide correct results.

II. RELATED WORK

In [1] Abey Abraham, Rohini. V, the conversion of signs to speech is used to communicate between mute and non-mute people. It's a glove-based gadget that tracks hand movement gestures and their placements. Different sensors, such as flex sensors, are used in this model to consider the diverse gestures of our hands. They do this by using the Back propagation algorithm, which is used to swiftly calculate derivations. Backpropagation is a learning algorithm used by artificial neural networks to compute a gradient descent with respect to weights. Hand motions and gestures are captured using the flex sensor. These sensors are embedded in gloves that collect data from flex sensors via numerical data values. Here they using the four flex sensors to capture the hand signs and get the text which is converted into speech recognition. Here they use the four flex sensors to capture the hand signs and get the text which is converted into speech recognition.

In [2] Victoria A. Adewale,Dr. Adejoke O. Olamiti, Humans mostly communicate with one another through natural language channels such as speech and writing. However, persons who are unable to communicate cannot use the natural language channel. The deployment of sign to speech conversion is recommended for this purpose, and it is expected to have a significant impact on deaf and stupid individuals. Kinect sensors, picture segmentation, feature recognition and extraction from ROI, K-Nearest Neighbour (KNN) algorithm, and text-to-speech (TTS) conversion are used to create this system. The image input is taken, and the image is segmented, making the image more meaningful and easier to analyze. The Kinect sensor is employed as a tool in this project. The photos are converted to grayscale images using MATLAB. Then elements such as movement and hand gestures are retrieved. Now we match the image with the dataset which is in the database, if there is a match the text will appear and that text will be converted through speech by using text- to -speech (TTS) conversion method.

In [3] Manikandan ,AyushPatidar, PallavWalia, Aneek Barman Roy, It is very hard to communicate with people who are hearing impaired and dumb people. In this

They proposed the solution by using modern technology methodologies which are based on hand gestures, and through images. This method is implemented by using contour analysis and feature extraction. In this they use equipment like gloves , which is used to point out the hand movement and gestures which is useful for communication. The image is converted into a monochromatic image by using the open CV library. OpenCV is used to determine the edge of the object, which is defined as the contour of the object. And then it creates the binary of an image by using the background which makes the recognition process simpler. Next it finds the outline of the image by using contour, then it will calculate the Hull and defects of the image , meaning the area which does not belong to the image, but it will present inside the image. And then the identification of the letters is done , then we will get the text. The writing is transformed into language by using the text-to-speech conversion methodology.

In [4] Aarthi M. Vijayalakshmi P, Deaf and mute humans want to talk with regular humans for his or her daily routine. They use their signal language to talk with different humans. It is simplest possible via which humans go through unique schooling to recognize signal language. It may be processed in methods which can be imaginative and prescient - primarily based totally gesture popularity and via sensor - primarily based totally. In imaginative and prescient primarily based totally gesture the cameras are used to seize the hand movements. In this the photograph processing strategies also are used to extract the functions and to apprehend the gestures and deliver the identified hand gesture as output. In the sensor - primarily based totally gadget they use Tactile sensors and a few flex sensors. Tactile sensors are used to degree the pressure implemented on one finger

through some other finger. Deaf - mute humans face many issues speaking with others, so to conquer those issues this technique could be very useful.

In [5] S. Rajaganapaty ,Aravind.B, Keerthana.B, Sivagami.B, those who are unable to speak, communicate with others through unusual techniques and actions. In today's world, there are a plethora of ways to communicate with various people using sign language. They employed Microsoft Kinect to capture movement and movements in this project, which was advanced with the help of Microsoft. There are several distinct classes in sign language, such as ISL (Indian Sign Language), ASL (American Sign Language), and BSL (British Sign Language) (British Sign Language). Kinect is capable of capturing the surrounding world in 3-D by integrating information from intensity sensors and a standard RGB camera.

III. PROPOSED SYSTEM

The application is then set up to extract used to predict alphabets and digits. Finally, the user input image is sent to the CNN model for prediction, which compares the input footage with the store

within the CNN model. The CNN model creates output in text and audio forms based on a comparison. Our model's fundamental concepts are intended to classify the sign language alphabet, which is backed by human hand gestures. The quick model's operational methodology.technique, also known as the binary image model, is then compared to the teaches model. The initiative is designed to encourage a customer to take a photo with a mobile device camera. This type of captured The model creates a file of type h5 that contains an overview of the work input frame must be forced to be removed. In addition, the hand component of the image is discovered, as well as the hand footage is saved in the input folder. The CNN work model is then often supported by hand gesture footage. Hand gesture frames must then be forced to be converted from RGB to Grayscale. To improve the precision of the entered gesture, the noise contained within the stream, which various image processing algorithms are then applied to. The input frames from the camera's time-lapse video technique as part of the work output. According to the model, this file will be

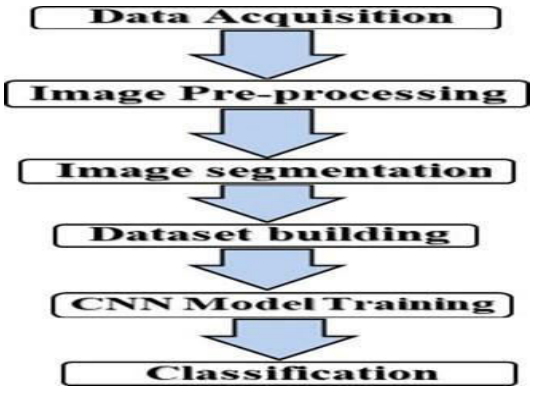

Fig 1: Working process of the suggested model.

A. Data Acquisition

To begin, record a hand gesture video on a handheld laptop, such as a tablet, and collect the video frames for extra calculations.

B. Pre-Processing stage

Now that the ROI has been isolated from the body, it can be recognized on a border basis and printed with the use of a single body marker. The value of the grayscale depth is exceptional, and the site should be divided on the pel depth. The pixels inside the body are usually assigned intensity values of zero (foreground) or one (for heritage). When it recognizes a position as a hand or a heritage, it ignores the rest of the video body that isn't being used and resizes it to a certain resolution.

Fig 2: ROI

C. Division and element extraction stage

For pre-taking care of the acquired images, we have one of each of the two ways. Equal Mode is used for the basic setup to convert the image to grayscale, and any spot Out is used to convert the dim image to matching photos. Any spot with a value of zero (for setup) or one is used to convert the image to grayscale (for closer view).However, establishment derivation is used for the compound establishment to provoke a lot of a picture that it processes the number shuffling of the intro page between the pad store and hence the establishment picture, or if nothing else, the decent piece of the gathering or, extra consistently, something that may be shown because the establishment concerning the properties of the scene being pondered. More racket clearing techniques, such as Gaussian fog and Erosion, are used in everything. Morphological filtering is essential for applying morphological isolating to segmental images in order to achieve a cleaner, more closed, and more structured movement. This is traditionally accomplished by developing scratched spot efforts while the invariant segmental film is turned.

D. Customized Dataset

For American Sign Language, HGR screenshots for twenty-six letter signs were obtained. Each sign and individual has a total of 2000 images. Frames are being used as photos in coordinators along similar lines. The coordinator's name is becoming familiar with stamping photographs that correspond to the type of signs found inside the packing (for instance mark A for letter set A, and subsequently on for sure classes of signs). Every representation of the language will be employed; nevertheless, the comprehension anticipated is for the Yankee language Alphabet, as shown in the fig 3.

Fig 3: ASL dataset Images

E. Include determination & Training stage

For the time being, the deleted attributes have been chosen for characterization. The vector part of this supported model of element extraction comes from the casing of a video sequence used by the CNN. When the picture is extracted into the record, some of the removed properties of the picture are restored. Strong AI calculations have been acclimated to separate attributes. CNN is one of the most well-established methods for deep learning. CNN employs a variety of various visual alterations. CNN can collect potential highlights for the order model from a wide range of images. There are four secret crucial layers, the information aspect is explicit. The300x300x1 goal of the information framework in this study is 300x300x1.

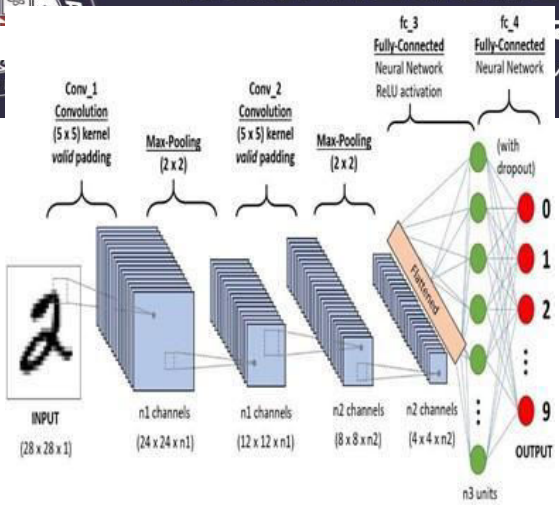

Fig 4: CNN architecture

CNN produces a partner degree actuation type, that is utilized as a partner degree input picture for each layer. . the technique proceeds with layer upon layer. Without a doubt, their region unit essentially many layers of material for the extraction of picture highlights all through CNN. inside this brief model, stowed away layers of region units expected on the grounds that the extraction of properties. By these layers, the modest picture properties of the layer's region unit acquired at the organization began. More profound layers of the organization strategy these fundamental highlights and union the primer component to build picture elements of the upper level. These regions unit all properties at the powerful region unit all around intended for arrangement exercises. Since more profound layers of the organization integrate these highlights of simpleness into a higher picture epitome.

F. Classification stage

At last, the SoftMax work is utilized to evaluate each sequential sign all through a definitive piece of this encouraged model. That is a speculation of supply relapse in light of the fact that it will be maintained to steady information (instead of matched portrayal) which could represent a couple of decision limits. It manages multinomial naming systems. SoftMax refers to the work that we will most likely evaluate in the classifier's outcome layer. The SoftMax start-up work allows for the movement of chance between correspondingly specific outcome characterizations. In many circumstances, SoftMax is employed as a learning device for the diagram of determined relapse and elements. SoftMax is a classifier that uses a regulated learning approach to categorize data into many categories. SoftMax's main functional operation is to categorize the example data input into a few distinct types. One class is segregated from the other in SoftMax, and an official conclusion is reached for this approach by selecting the most extreme SoftMax esteem result.

IV. RESULT

The guided deep learning technique was applied using Python and the Keras and TensorFlow backend packages. The steered model is assessed using a 26-symbol interconnected data set. Every sign is made up of 2000 images of each individual. The data is divided into two groups. The primary branched assortment contains eighty percent of the training photographs, while the other contains the remaining

twenty percent of the testing pictures.CNNis utilized in the extraction of a feature. When using CNN to extract features from pictures, we take into account the number of teaching features and the number of examining features. All of these qualities are beneficial in distinguishing the classes within each unique sign. Figure 5 shows that model accuracy is around 98 percent on average. Each of the indicators exhibited by the SoftMax has been detected using these more extensive training and testing features. The classification's quality, which is 96.68 percent, has been accepted.

V. CONCLUSION

In terms of the delicacy and responsibility associated with it, hand sign recognition might be a significant strike in real-life executions. The input gestures square measure recorded utilizing a mobile device camera in this study, which introduces the hand sign identification in ASL while not touching. A still-hand shot obtained from a time-lapse videotape stream frame and used by CNN to hunt for new perceptive qualities.SoftMax, in the end, discovered the ABC symbol. Our designed dataset is ensure the validity of the model that is being developed used in accordance with ASL norms to signing recognition with impaired persons as a product of HCI result in a bracket delicacy of 96.68. The motion must be made in front of the camera, so the event is provided in the form of a textbook or audio.

REFERENCES:

1. Abey Abraham, Rohini.V. Real time conversion of sign language to speech andprediction of gestures using Artificial Neural Network. Procedia Computer Science(2018),143(),587–594. doi:10.1016/j.procs.2018.10.435
2. Victoria AdebimpeAdewale,Adejoke OOlamiti. Conversion of sign language to textand speech using machinelearning techniques(2018).DOI:10.36108/jrrslasu/8102/50(0170).
3. K. Manikandan, AyushPatidar, PallavWalia, Aneek Barman Roy. Hand Gesture Detection and Conversion to Speech and Text(2018).
4. Vijayalakshmi P, Aarthi M. Sign language to speech conversion(2016). 1–6. doi:10.1109/ICRTIT.2016.7569545.
5. Mallikarjuna Reddy, A., Venkata Krishna, V. and Sumalatha, L. Face recognition approaches: A survey. International Journal of Engineering andTechnology (UAE), 4.6,6(7)(2018) 117-121. doi: 10.14419/ijet.v7i4.6.20446.
6. M. R. Ayaluri, K. Sudheer Reddy, S. R. Konda, and S. R. Chidirala, “Efficient steganalysis using convolutional auto encoder network to ensure original image quality,” PeerJ Computer Science, vol. 7, p. e356, 2021.

7. Rajaganapathy S, Aravind B, Keerthana B &Sivagami M. "Conversation of Sign Language to Speech with Human Gestures." Process.
8. J. F. and H. H. Jonathan Lazar, Research Methods in Human Computer Interaction. Morgan Kaufmann, 2017.
9. Mallikarjuna Reddy, A., RupaKinnera, G., Chandrasekhara Reddy, T., Vishnu Murthy, G., et al., (2019), "Generating cancelable fingerprint template using triangular structures", Journal of Computational and Theoretical Nanoscience, Volume 16, Numbers 5-6, pp. 1951-1955(5), doi: https://doi.org/10.1166/jctn.2019.7830.
10. A. Mantri and M. Ingle, A Comparative Study of Various Techniques Used in Current HGRSs.

Springer Singapore,2018.
11. C Ramakrishna, G Kiran Kumar, A Mallikarjuna Reddy, and P. Ravi, "A Survey on various IoT Attacks and its Countermeasures," Int. J. Eng. Res. Comput. Sci. Eng., vol. 5, no. 4, pp. 2394–2320, 2018, [Online]. Available: http://ijercse.com/specissue/april-2018/27.pdf.
12. Swarajya Lakshmi V Papineni, SnigdhaYarlagadda, HaritaAkkineni, A. Mallikarjuna Reddy. Big Data Analytics Applying the Fusion Approach of Multicriteria Decision Making with Deep Learning Algorithms International Journal of Engineering Trends and Technology, 69(6), 24-28. doi: 10.14445/22315381/IJETT-V69I6P204.
13. R. A. Bhuiyan, A. K. Tushar, A. Ashiquzzaman, J. Shin, and R. Islam, "Reduction of Gesture Feature Dimensionfor Improving the Hand Gesture Recognition Performance of Numerical Sign Language," pp. 22–24, 2017.

14. B. A. Krizhevsky, I. Sutskever, and G.E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” 2012.
15. Mallikarjuna Reddy A, RupaKinnera G, Chandrasekhara Reddy T, Vishnu Murthy G. 2019. Generating cancelable fingerprint template using triangular structures. Journal of Computational and Theoretical Nanoscience 16(5–6):1951-1955
16. A Mallikarjuna Reddy, VakulabharanamVenkata Krishna, LingamguntaSumalatha and AvukuObulesh, Age Classification Using Motif and Statistical Feature Derived On Gradient Facial Images, Recent Advances in Computer Science and Communications (2020) 13:965. https://doi.org/10.2174/2213275912666190417151247.
17. A. Z. Karen Simonyan, “VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION,” pp. 1–14, 2015.
18. J. Donahue et al., “DeCAF : A Deep Convolutional Activation Feature for Generic Visual Recognition,” vol. 32, 2014.
19. U. G. NurettinÇag˘rıKılıboz, “A hand gesture recognition technique for human–computer interaction,” vol. 28, pp. 97–104, 2015, doi: 10.1016/j.jvcir.2015.01.015.
20. G. Modanwal and K. Sarawadekar, Trans. ICT, no. June, 2019, doi: 10.1007/s40012-019- 00251-w. and Y. J. , Hongbin Ma, Baokui Li, Static Hand impaired for computer interaction,”CS Gesture Recognition for Human Robot Interaction. Springer International Publishing, 2019.
21. A. Haria, A. Subramanian, N. Asokkumar, S. Poddar, and J. S ayak, “Hand Gesture Recognition for Human Computer Interaction,” ProcediaComput. Sci., vol. 115, pp. 367–374, 2017, doi: 10.1016/j.procs.2017.09.092.
22. P. P. M. C. Simran Shah, Ami Kotia, KaushaNisar, AneriUdeshi, “A Vision Based Hand Gesture Recognition System using
23. Convolutional Neural Networks,” Int.Res. J. Eng. Technol., vol. 463, no. 6, pp. 2570–2575, 2019, doi: 10.1007/978-981-10- 6571- 2_132.
24. N. Umadevi and I. R. Divyasree, “Development of an Efficient Hand Gesture Recognition system for human computer interaction,” no. September, 2018, doi: 10.18535/Ijecs/v4i12.5