International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 02 | February - 2024 SJIF Rating: 8.176 ISSN: 2582-3930

Hand Sign Language Recognition and Conversion into Text Using Convolutional Neural Network

Prof. Jyotsna Nanajkar
Harsh Kshirsagar, Aniket Shelke, Nikita Shinde, Manoj Thombre
Zeal College of Engineering and Research

Abstract- *The current research looks into using convolutional neural networks, also known as CNNs, to develop a hand sign understanding and conversion system. The study addresses the urgent need for a precise and effective way to close the communication gap that exists between the hearing and the deaf communities. The study examines how well CNNs can identify and translate hand sign gestures into text, giving sign language users a real-time means of communication. We use three steps in our methodology: gathering datasets, training models, and thorough assessment. The outcomes show encouraging hand sign recognition accuracy, creating new opportunities for inclusive communication.*

Sign language is one of the oldest and most natural forms of language for communication, but since most people do not know sign language and interpreters are very difficult to come by, we have come up with a real-time method using neural networks for fingerspelling-based American sign language. In our method, the hand is first passed through a filter and after the filter is applied the hand is passed through a classifier which predicts the class of the hand gestures. Our method provides 95.7 % accuracy for the 26 letters of the alphabet.

Keywords-Sign Language, ASL, Hearing disability, Convolutional Neural Network (CNN), Computer Vision, Machine Learning, Gesture recognition, Sign language recognition, Hue Saturation Value algorithm.

INTRODUCTION

Based on the World Health Organization (WHO) statistics, there are over 430 million people with hearing loss disability (WHO 2023) which is 5% of the world population and it is It is estimated that by 2050 over 700 million people – or 1 in every 10 people – will have disabling hearing loss. According to Sign Solutions, there are more than 300 sign languages used around the world. Communication skills are most important for every human to share their thoughts and ideas. One of the biggest obstacles still facing the deaf and hearing communities is communication. We present a novel method using convolutional neural networks (CNN) for hand sign language recognition and conversion into text to address this problem. The purpose of this project is to determine whether it is feasible and accurate to translate hand sign gestures into text. The need for effective and inclusive communication,

where deep learning technology is essential, is what spurs this research.

American sign language is a predominant sign language Since the only disability D&M people have been communication-related and they cannot use spoken languages hence the only way for them to communicate is through sign language. Communication is the process of exchange of thoughts and messages in various ways such as speech, signals, behaviour and visuals.

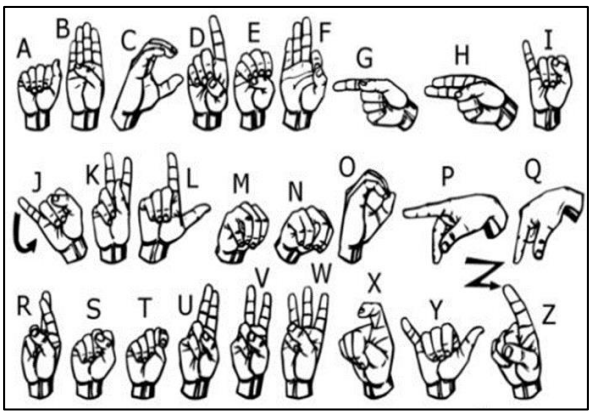

Fig 1. American Sign Language Gestures

Deaf and dumb(D&M) people make use of their hands to express different gestures to express their ideas to other people. Gestures are nonverbally exchanged messages and these gestures are understood with vision. This nonverbal communication of deaf and dumb people is called sign language.

In our project, we focus on producing a model which can recognise Fingerspelling-based hand gestures to form a complete word by combining each gesture.

LITERATURE SURVEY

Datasets like MNIST with accuracy 99.6% used for image recognition using CNN model, LSTM solves the vanishing gradient problem with accuracy 95.1%, Sports-1M dataset used for action recognition. The Jester dataset results in a change of loss, accuracy of 94.85%. The Nvidia dataset results with an average accuracy of 88.4%. [1]

King Saudi University saudi sign language(ksu-ssl) dataset which was recorded by 40 participants over five recording sessions with accuracy 98.75%. Two 3DCNN instances were used separately for learning the fine-grained features of the hand shape. SoftMax function was used for the classification. The proposed system outperformed state-of-the-art methods in terms of recognition rate, demonstrating its effectiveness.[2]

CNN algorithm is particularly used to improve the recognition accuracy under challenging conditions such as a change in scale, rotation and translation. Using a large dataset increases the accuracy of the result. [5]

using SURF with SVM and CNN" recognizes Indian sign language signs (A-Z) and (0–9) using the SVM with the accuracy of 99.14% on the test data and CNN with accuracy of 94% on the training set, it trained on all 36 ISL static alphabets and digits with an accuracy of 99%. [13]

"Sign language recognition system for communicating to people with disabilities" using ASL Hand Sign Dataset (Grayscaled Thresholded) which contains 24 classes applied a gaussian blur filter. 3 dataset classes were taken from Nikhil Gupta, namely the datasets for classes J, Z, and 0 (blank) with training and validation accuracy of 89.1% and 98.6%,

respectively. Using confusion matrix in two layers Convolutional Neural Network (CNN) with accuracy 96.3%. [14]

PROPOSED WORK

Individuals who are deaf or hard of hearing try to communicate with the general public in American Sign Language, yet they are unaware of this sentence construction in the language. In the end, crucial to develop an intelligent and programmable arbitrator to obtain them. The suggested system operates by adhering to the specified methods for identifying and translating the hand motions into written communication and vice versa.

A. Convert ASL into Speech and Text

1. Take a picture or a video and feed it into the system.
2. The most basic stage of processing digital images is image acquisition and enhancement.
3. Every pertinent feature is retrieved, and every redundant and unnecessary detail is disregarded.
4. The input is analysed to identify the correct character.
5. Text and speech are produced from the identified ASL characters.

B. Transcribing Text or Speech into ASL

1. Record the user's voice or speech and feed it into the system as input.
2. The input that is received will be examined and transformed into its corresponding text
3. The corresponding text is found
4. The text is thereafter transformed into the matching ASL character.

SYSTEM FLOWCHART

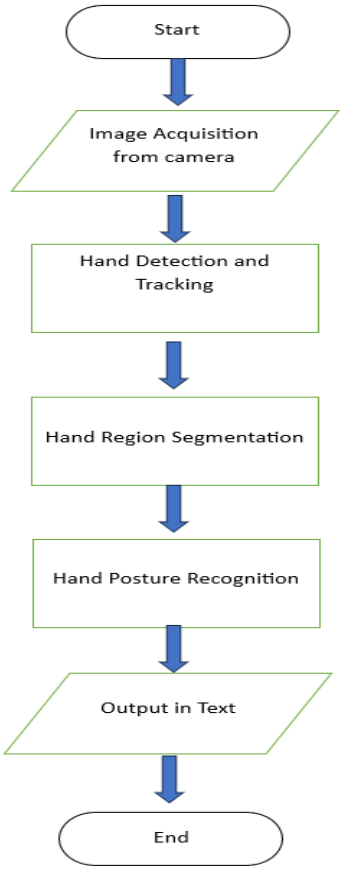

fig 2. System flowchart for image recognition

SETUP FOR AN EXPERIMENT

1. There are 16,120 entries in our data set overall, and 26 labels are used for training (30%).
2. The dataset is being divided at random by divided into 20% testing and 80% training.
3. We need processors with Intel I3 or above.
4. Memory: 80 GB, RAM: 4 GB
5. It will result in improved GPU performance.
6. The camera is positioned appropriately for taking a picture of the input.
7. The Eclipse/NetBeans IDE/ Visual Studio Code
8. Python for the dataset's testing and training with CNN

SEQUENCE DIAGRAM

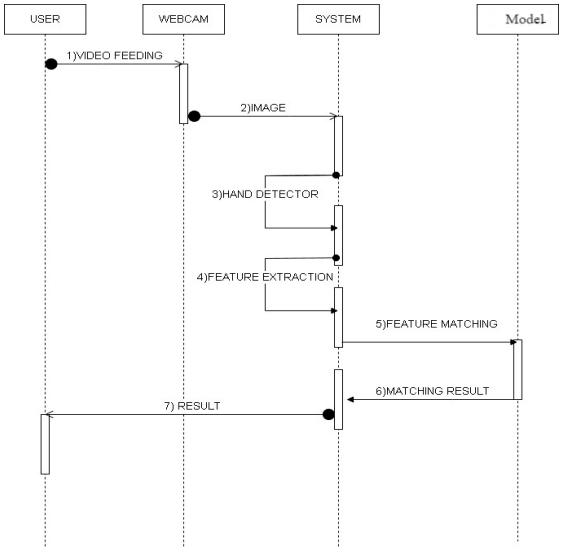

Fig 3. Sequence diagram for image recognition model

CONCLUSION

Based on a survey of prior literature, we explained how varied accuracies result from the range of sign language datasets, which comprise different motions. This study demonstrated that several datasets have been utilised while SLR systems are being trained and tested. It illustrated the benefits and drawbacks of both the vison-based and glove-based approaches both, clarified the distinction between signer-dependent and signer-independent, and covered the fundamental skin detector, picture segmentation, hand tracking, feature extraction, and hand gesture classification are examples of preprocessing procedures. The report also demonstrates that deep learning outperforms standard machine learning by contrasting a few ML techniques with the most popular deep learning algorithm, CNN. Due to the precise signals that researchers obtain during feature extraction, some gloves-based systems perform better than deep learning algorithms. In contrast, the features that researchers obtain during model training with deep learning are not as accurate as those obtained with gloves-based systems. This previous problem states that we must remove any electronic devices or obstructions that might prevent users from interacting with the system, such as gloves, sensors, and leap gadgets. There had been numerous attempts, but the accuracy was lower.

REFERENCES

[1] Wenjin Zhang, Jiacun Wang, Senior Member, IEEE, and Fangping Lan, “Dynamic Hand Gesture Recognition Based on Short-Term Sampling Neural Networks,” IEEE/CAA JOURNAL OF AUTOMATICA SINICA, VOL. 8, NO. 1, JANUARY 2021.

[2] Muneer Al-Hammadi, Ghulam Muhammad, Senior Member, IEEE, Wadood Abdul, Member, IEEE, Mansour Alsulaiman, Mohamed A. Bencherif, and Mohamed Amine Mekhtiche, “Hand Gesture Recognition for Sign Language Using 3DCNN,” DOI 10.1109/ACCESS.2020.2990434, IEEE Access.

[3] Mehreen Hurroo, Mohammad Elham Walizad, “Sign Language Recognition System using Convolutional Neural Network and Computer Vision,” International Journal of Engineering Research & Technology (IJERT) http://www.ijert.org ISSN: 2278-0181 IJERTV9IS120029 Published by: www.ijert.org Vol. 9 Issue 12, December-2020.

[4] R.S. Sabeenian, S. Sai Bharathwaj, M. Mohamed Aadhil “Sign Language Recognition Using Deep Learning and Computer Vision,” Jour of Adv

Research in Dynamical & Control Systems, Vol. 12, 05-Special Issue, 2020

[5] Arvind Sreenivas, Mudit Maheshwari, Saiyam Jain, Shalini Choudhary, Dr. G. Vadivu, "Indian Sign Language Communicator Using Convolutional Neural Network," International Journal of Advanced Science and Technology Vol. 29, No. 3, (2020), pp. 11015 - 11031.

[6] Ahmed Sultan, Walied Makram, Mohammed Kayed, Abdelmaged Amin Ali, "Sign language identification and recognition: A comparative study," https://doi.org/10.1515/comp-2022-0240 received April 29, 2021; accepted April 12, 2022.

[7] Sayali Gore, Namrata Salvi, Swati Singh, "Conversion of Sign Language into Text Using Machine Learning Technique," International Journal of Research in Engineering, Science and Management Volume 4, Issue 5, May 2021 https://www.ijresm.com | ISSN (Online): 2581-5792.

[8] Rachana Patil1, Vivek Patil1, Abhishek Bahuguna1, and Mr. Gaurav Datkhile, "Indian Sign Language Recognition using Convolutional Neural Network," ITM Web of Conferences 40, 03004 (2021) ICACC-2021.

[9] G. Poon, K. C. Kwan, and W.-M. Pang, "Occlusion-robust bimanual gesture recognition by fusing multi-views," Multimed. Tools Appl., vol. 78, no. 16, pp. 23469–23488, Aug. 2019.

[10] N. K. Bhagat, Y. Vishnusai and G. N. Rathna (2019). Indian Sign Language Gesture Recognition using Image Processing and Deep Learning. Digital Image Computing: Techniques and Applications (DICTA), Perth, Australia. pp. 1-8, doi: 10.1109/DICTA47822.2019.8945850.

[11] M. Al-Hammadi, G. Muhammad, W. Abdul, M. Alsulaiman, and M. S. Hossain, ''Hand gesture recognition using 3D-CNN model,'' IEEE Consum. Electron. Mag., vol. 9, no. 1, pp. 95–101, Jan. 2020.

[12] G. Muhammad, M. F. Alhamid, and X. Long, ''Computing and processing on the edge: Smart pathology detection for connected healthcare,'' IEEE Netw., vol. 33, no. 6, pp. 44–49, Nov./Dec. 2019.

[13] Shagun Katoch, Varsha Singh, Uma Shanker Tiwary "Indian Sign Language recognition system using SURF with SVM and CNN"