IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Impact Factor: 6.252
Volume 2, Issue 2, February 2022

A Literature Survey on Real-Time Indian Sign Language Recognition System

Karmanya Verma[1], Mansi Bansode[2], Navaditya Kaushal[3], Mrs. Shubhangi Vairagar[4]
Students, Department of Computer Engineering[1,2,3]
Asst. Professor, Department of Computer Engineering[4]
Dr. D. Y. Patil Institute of Technology, Pune, Maharashtra, India
Affiliated to Savitribai Phule Pune University, Pune

Abstract: *Sign language is one of the visual means of communicating through hand signals, gestures, facial expressions, and visual communication. It's the main form of communication for people with the disability of hearing or speech. People with disabilities like autism spectrum disorder may also find sign language beneficial for communicating. The system will realize Indian Sign Language using a keypoint detection model. It will be used to make a sequence of keypoints. These keypoints can then be passed to an action detection model. The proposed system will be predicting Indian Sign Language signs using several frames and predicting what action is being demonstrated. The system will use Mediapipe Holistic to extract the keypoints. It allows extracting keypoints from the user's hands, from the user's body, and the user's face. We will then use Tensorflow and Keras to build an LSTM model to be able to predict the action from the live video feed. The system is going to train a deep neural network using LSTM layers to go on ahead and predict that temporal component so we will be able to predict action from several frames. Then we are going to put it all together, the Mediapipe Holistic and trained LSTM model using OpenCV and go on ahead and predict in real-time using our webcam.*

Keywords: Machine Learning, Feature Extraction, Mediapipe Holistic, LSTM, Sign Language Recognition.

I. Introduction

Sign language is the primary mode of communication between the hearing and vocally impaired population. For people with hearing impairment, Indian Sign Language is an important communication medium. People with disabilities including Autism, Down Syndrome may also find sign lang useful for communicating. There are more than 300 different sign languages, and they vary from nation to nation. Countries with the same spoken language do not have the same sign language. American, British, and Australian Sign Language are three different varieties of English.

Our application uses the Deep Learning model for its implementation. Deep Learning is a subfield of machine learning and is essentially a neural network with three or more layers. It's concerned with algorithms inspired by the structure and performance of the brain called artificial neural networks. These neural networks plan to simulate the behavior of the human brain allowing it to find out from large amounts of knowledge. Using the combination of data inputs, weights, and bias, Deep learning neural networks attempt to mimic the human brain. There are different types of neural networks to deal with specific problems or datasets. For example:

1. Convolutional Neural Network (CNN): A CNN contains one or more than one convolutional layer. It contains a three-dimensional arrangement of neurons. It takes an image as input, assigns significance to varied aspects/objects within the image, and differentiates one from the other. CNN can be used for image processing, computer vision, speech recognition, etc.
2. Recurrent Neural Network (RNN): The output of a particular layer in RNN is saved and fed back to the input. This helps predict the result of the layer. If the prediction is not right, the learning rate is employed to make small changes. The application of RNN can be found in the text to speech (TTS) conversion models. Other

applications of RNN are text processing like grammar checks, auto-suggest, etc., sentiment analysis, image tagger, translation, etc.

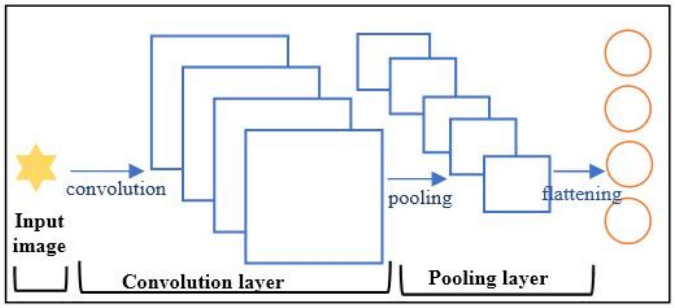


Figure 1: Convolutional Neural Network (CNN)

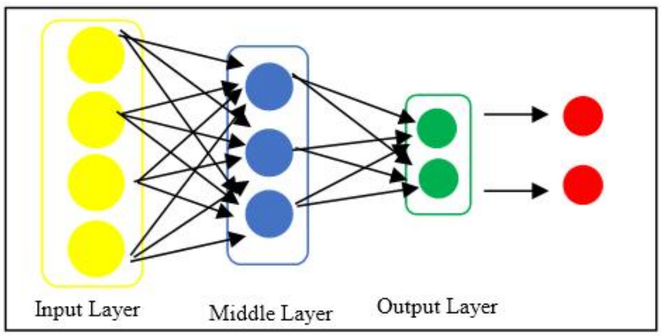


Figure 2: Recurrent Neural Network

II. Literature Survey

A. A Signer Independent Sign Language Recognition with Co-articulation Elimination from Live Videos: An Indian Scenario [1]

The system produced by the authors recognizes ISL gestures from mobile camera videos with none additional sensors to detect hand regions. Single-handed static and dynamic gestures and double-handed static gestures are identified in this system. The system uses ROI algorithm for feature extraction. After pre-processing the image, the centroid of the binary image is estimated. Using the gesture separation algorithm, the gesture is recognized, static as well as dynamic.

The advantage of this system is that it is economical and can be implemented with a mobile camera, making it very user-friendly. But the disadvantage is that it is not efficient under cluttered backgrounds and different illumination conditions.

B. A Deep Learning-based Indian Sign Language Recognition System [2]

This system was successfully trained on all ISL static alphabets with a training accuracy of 99.93% and with testing and validation accuracy of 98.64%. Facial expression and context analysis are the other part not included in this project.

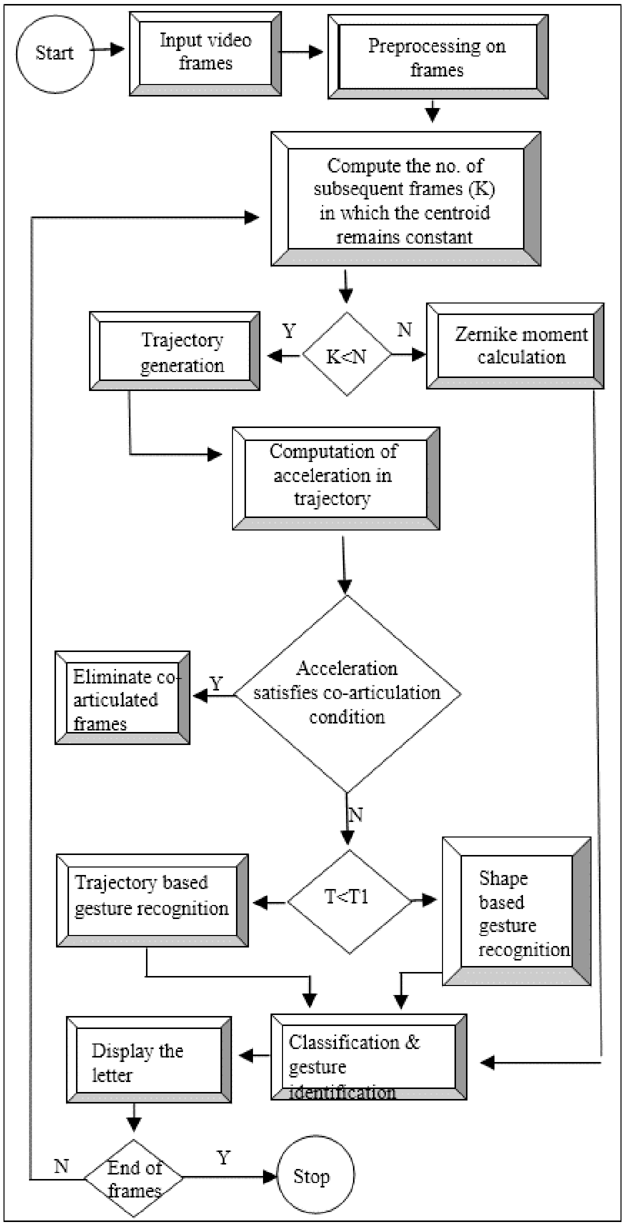
IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 6.252

Volume 2, Issue 2, February 2022



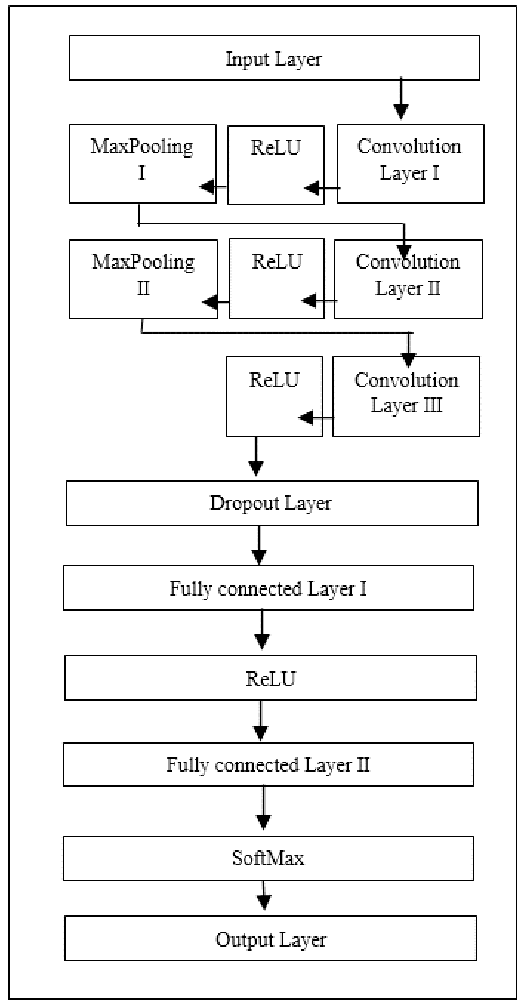


Figure 3: Flow diagram of the proposed system **Figure 4:** Diagrammatic Representation of Proposed Signet Architecture

C. An Efficient Binarized Neural Network for Recognizing Two Hands Indian Sign Language Gestures in Real-time Environment [3]

Considering the challenges of sign language recognition, on targeted embedded platforms, authors have proposed the novel architecture of a binarized neural network with binary values of weights and activations using bitwise operations. The advantage is using this architecture achieves an overall accuracy of 98.8% which is higher than other existing methods while the disadvantage is This system misclassifies some signs of M, N, E because of their similar kind of shapes, and also, the proposed BNN architecture is limited with small no of classes of gestures.

DOI: 10.48175/IJARSCT-2728

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Impact Factor: 6.252
Volume 2, Issue 2, February 2022

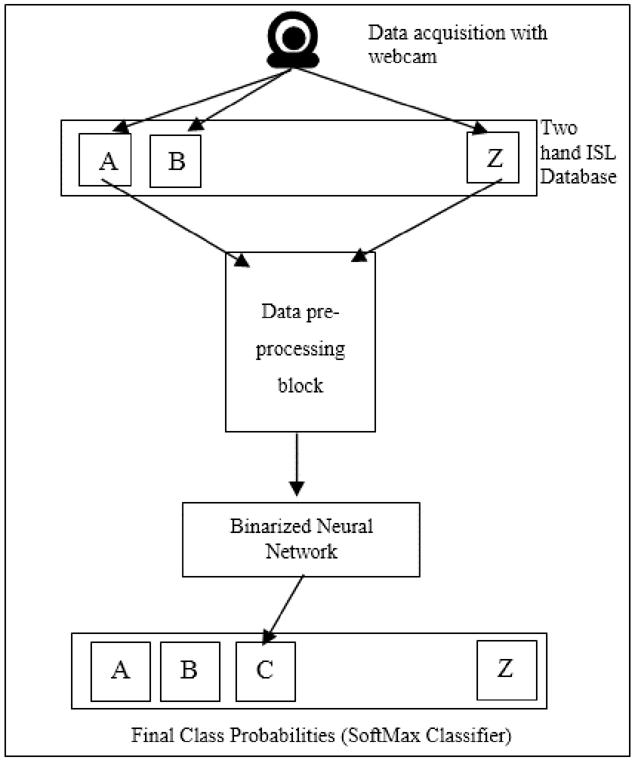


Figure 5: Architecture for the proposed system

D. Hand Gesture Recognition for Sign Language Using 3DCNN [4]

This study investigates the use of 3DCNN for hand gesture recognition. The approaches that were proposed by the authors were compared with six other state-of-the-art methods from the literature. They showed comparable performance to the two of these methods and outperformed four of them. However, it does not work for a live video feed.

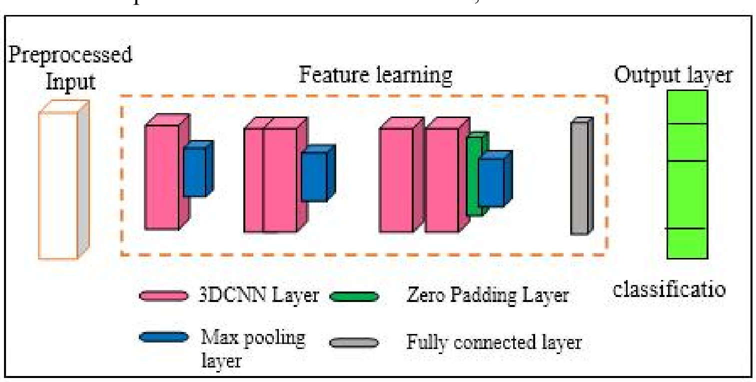


Figure 6: Single 3DCNN-based structure.

E. American Sign Language Recognition Using RF Sensing [5]

This composition provides an extensive, interdisciplinary perspective on ASL recognition using RF sensors. Results demonstrate the eventuality of RF sensing to give contactless ASL recognition capabilities in aid of ASL-sensitive smart surroundings while surviving effectively in the dark and guarding user privacy. A massive intermodal database of connected native signing would be required to make meaningful interpretations for technology and algorithm correlation.

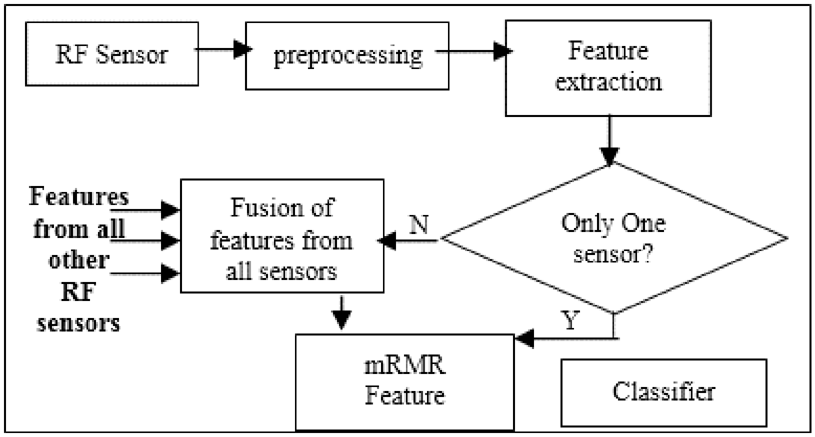


Figure 7: Block Diagram of ASL recognition using RF sensing

F. Video-based isolated hand sign language recognition using a deep cascaded model [6]

The authors used a cascaded architecture of SSD, CNN, and LSTM from RGB Videos to propose a deep-based model for systematic hand sign recognition. The accuracy and complexity of hand sign recognition were improved by this model. In case of an uncontrolled environment such as rapid hand motions, it provided fast processing. Using more data, the accuracy of detection can be improved.

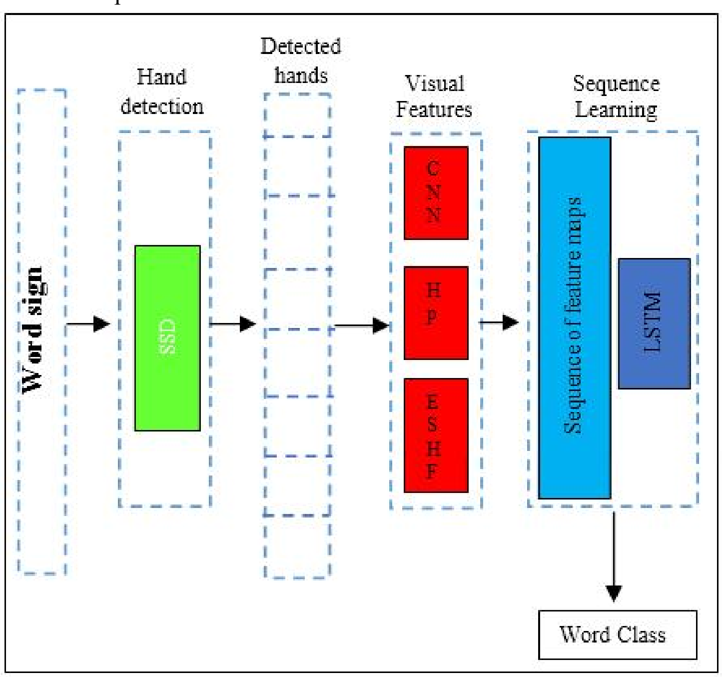


Figure 8: The Proposed Model

G. A Modified-LSTM Model for Continuous Sign Language Recognition using Leap motion [7]

In this paper, they have presented a novel framework for continuous-SLR using the Leap motion sensor. A modified LSTM architecture has also been proposed for the recognition of sign words and sentences. Average accuracies of 72.3% and 89.5% have been recorded on the signed sentences and isolated sign words, respectively. The recognition performance can be improved by increasing more training data for better model learning.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Impact Factor: 6.252
Volume 2, Issue 2, February 2022

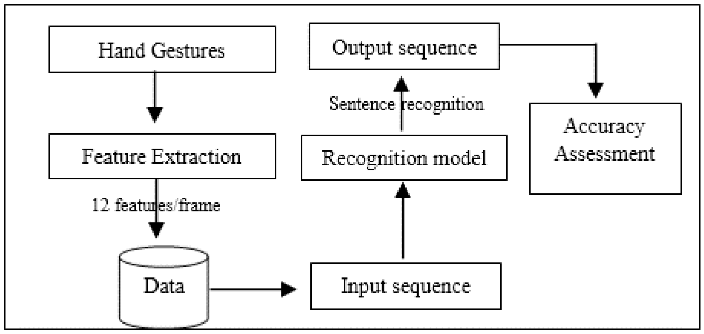


Figure 9: A proposed framework for continuous SLR using Leap Motion sensor

H. A depth-based Indian Sign Language recognition using Microsoft Kinect [8]

Overlapping signs, double hand signs, and signs unique to ISL were recognized successfully using this method. The advantage of this project is that the average recognition accuracy was improved up to 71.85% with this method. The system achieved 100% accuracy for a few of the signs but the system doesn't consider the environment of gestures, leading to incorrect translations on many occasions.

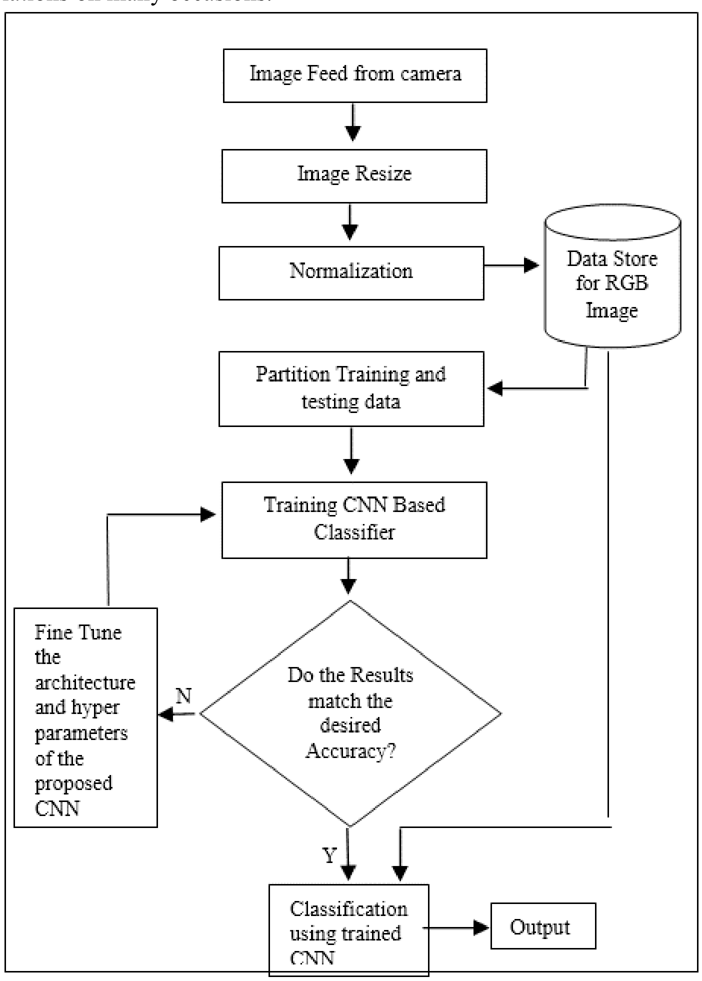


Figure 10: Proposed System

DOI: 10.48175/IJARSCT-2728

I. Deep learning-based sign language recognition system for static signs [9]

The authors achieved the highest training accuracy of 99.17% and validation accuracy of 98.80%, concerning changes in the number of layers and filters. To refine the recognition method, there is a need to collect more datasets.

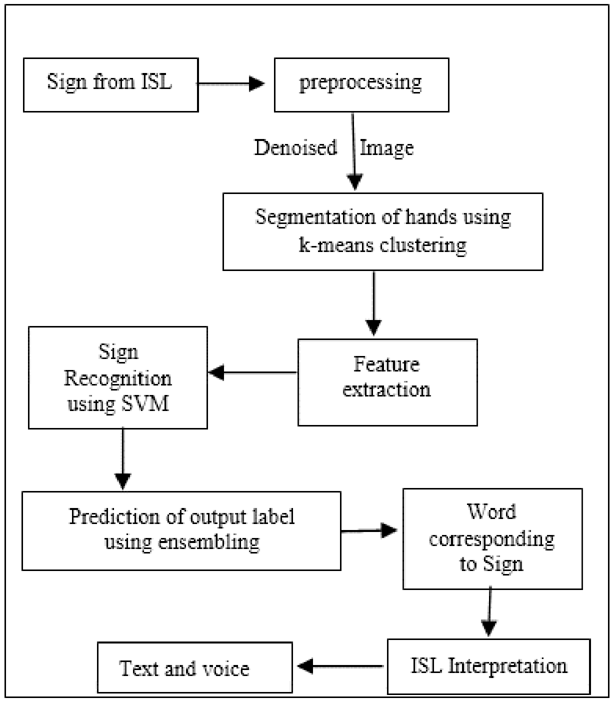


Figure 11: System Flowchart

J. Deep Convolutional Neural Networks for Sign Language Recognition [10]

For classifying selfie sign language gestures, the authors proposed a CNN architecture. A less amount of training and validation loss is observed with the proposed CNN architecture, but the database is not available publicly.

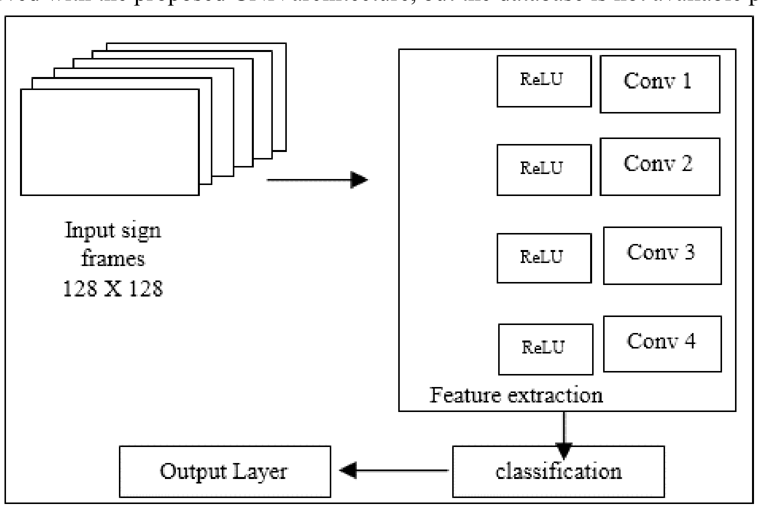


Figure 12: Proposed Deep CNN Structure

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Volume 2, Issue 2, February 2022

IJARSCT

Impact Factor: 6.252

III. Comparative Analysis

No.	Title of the Paper	Author and Year	Advantage	Disadvantage
1	A Signer Independent Sign Language Recognition with Co-articulation Elimination from Live Videos: an Indian Scenario	P.K. Athira, C.J. Sruthi, A. Lijiya (2019)	Economical can be implemented with a mobile camera which makes it very user-friendly	Not efficient under cluttered backgrounds and different illumination conditions
2	A Deep Learning based Indian Sign Language Recognition System	Sruthi C. J and Lijiya A(2019)	Training accuracy of 99.93% and with testing and validation accuracy of 98.64%.	Facial expression and context analysis are the other part not included
3	An Efficient Binarized Neural Network for Recognizing Two Hands Indian Sign Language Gestures in Real-time Environment	Mohita Jaiswal, Vaidehi Sharma, Abhishek Sharma, Sandeep Saini, Raghuvir Tomar (2020)	Using this architecture achieves an overall accuracy of 98.8% which is higher than other existing methods	This system misclassifies some signs of M, N, E because of their similar kind of shapes, and also, the proposed BNN architecture is limited to a small no of classes of gestures.
4	Hand Gesture Recognition for Sign Language Using 3DCNN	Muneer Al-Hammadi, Ghulam Muhammad, Wadood Abdul, Mansour Alsulaiman, Mohamed A. Bencherif, And Mohamed Amine Mekhtiche (2020)	The proposed approaches were compared with six other state-of-the-art methods from the literature. They outperformed four of these methods and showed comparable performance to the other two.	Does not work for a live video feed.
5	American Sign Language Recognition Using RF Sensing	Sevgi Z. Gurbuz, Ali Cafer Gurbuz, Evie A. Malaia, Darrin J. Griffin, Chris S. Crawford, Mohammad Mahbubur Rahman (2020)	Results demonstrate the eventuality of RF sensing to give contactless ASL recognition capabilities in aid of ASL-sensitive smart surroundings while surviving effectively in the dark and guarding user privacy.	A massive intermodal database of connected native signing would be required to make meaningful interpretations for technology and algorithm correlation.

6	Video-based isolated hand sign language recognition using a deep cascaded model	Razieh Rastgoo1 Kourosh Kiani Sergio Escalera (2020)	The accuracy and complexity of hand sign recognition were improved by this model. In case of an uncontrolled environment such as rapid hand motions, it provided fast processing.	Using more data, the accuracy of detection can be improved.
7	A Modified-LSTM Model for Continuous Sign Language Recognition using Leap motion	Anshul Mittal, Pradeep Kumar, Partha Pratim Roy, Raman Balasubramanian, and Bidyut B. Chaudhuri (2019)	Average accuracies of 72.3% and 89.5% have been recorded on the signed sentences and isolated sign words, respectively	The recognition performance can be improved by increasing more training data for better model learning.
8	A depth-based Indian Sign Language recognition using Microsoft Kinect	T Raghuveera, R Deepthi, R Mangalashri, and R Akshaya (2019)	The average recognition accuracy was improved up to 71.85% with this method. The system achieved 100% accuracy for a few signs.	The system doesn't consider the environment of gestures, leading to incorrect translations on many occasions.
9	Deep learning-based sign language recognition system for static signs	Ankita Wadhawan, Parteek Kumar (2020)	The authors achieved the highest training accuracy of 99.17% and validation accuracy of 98.80%, concerning changes in the number of layers and filters.	To refine the recognition method, there is a need to collect more datasets.
10	Deep Convolutional Neural Networks for Sign Language Recognition	G. Anantha Rao, K. Syamala, P.V.V. Kishore, A.S.C.S. Sastry (2018)	A less amount of training and validation loss is observed with the proposed CNN architecture.	The database is not available publicly.

IV. Conclusion

After studying the above-mentioned papers in detail, what we noticed was there is no system that uses real-time video sequence for predicting words of Indian Sign language or a system that works for a live video feed. Most of the systems work for static signs only or capture images and then compare them with a trained database or public datasets. So, there is a need for a system that predicts words and not just the alphabets of Indian Sign Language for a live video feed.

Acknowledgment

We acknowledge with thanks the assistance and timely guidance which we have received from our Guide Mrs. Shubhangi Vairagar Ma'am. We are very much thankful to Dr. S.V. Chobe Sir, Head of Department of Computer Engineering, and Dr. Pramod Patil Sir, Principal, Dr. D Y Patil Institute of Technology for their constant support and encouragement at various stages of our project. We would wish to express our deep sense of gratitude to Prof. Abhijit Jadhav Sir and Prof. Snehal Ma'am whose guidance, encouragement, suggestion, and very constructive criticism have contributed immensely to the evolution of our ideas on the project. Lastly, we would like to express our warm appreciation and sense of gratitude to every member of this group for their valuable contribution to making this project a success.

References

[1]. P. K. Athira, C. J. Sruthi, A. Lijiya, "A Signer Independent Sign Language Recognition with Co-articulation Elimination from Live Videos: An Indian Scenario",2019.https://www.sciencedirect.com/science /article/pii/S131915781831228X

[2]. M. Jaiswal, V. Sharma, A. Sharma, S. Saini, R. Tomar, "An Efficient Binarized Neural Network for Recognizing Two Hands Indian Sign Language Gestures in Real-time Environment",2020https://ieeexplore.ieee.org/ abstract/document/9342454

[3]. C. J. Sruthi, A. Lijiya, "Signet: A Deep Learning based Indian Sign Language Recognition System", 2019.https://ieeexplore.ieee.org/abstract/document/8698006

[4]. T. Raghuveera, R. Deepthi, R. Mangalashri, R. Akshaya, "A depth-based Indian Sign Language recognition using Microsoft Kinect", 2020.https://link.springer.com/article/10.1007/s12046-019-1250-6.

[5]. A. Wadhawan, P. Kumar, "Deep learning-based sign language recognition system for static signs", 2020.https:// link.springer.com/article/10.1007/s00521-019-04691-y

[6]. S. Z. Gurbuz, A. C. Gurbuz, E. A. Malaia, D. J. Griffin, C. S. Crawford, M. M. Rahman, E. Kurtoglu, R. Aksu, T. Macks, R. Mdrafi, "American Sign Language Recognition Using RF Sensing", 2020.https://ieeexplore.ieee. org/abstract/document/9187644

[7]. M. Al-Hammadi, G. Muhammad, w. Abdul, M. Alsulaiman, M. A. Bencherif, M. A. Mekhtiche, "Hand Gesture Recognition for Sign Language Using 3DCNN", 2020.https://ieeexplore.ieee.org/abstract/document/9078786

[8]. R. Rastgoo, K. Kiani, S. Escalera, "Video-based isolated hand sign language recognition using a deep cascaded model", 2020.https://link.springer.com/article/10.1007%2Fs11042-020-09048-5

[9]. G. A. Rao, K. Syamala, P. V. V. Kishore, A. S. C. S. Sastry, "Deep Convolutional Neural Networks for Sign Language Recognition", 2018. https://ieeexplore.ieee.org/abstract/document/8316344.

[10]. A. Mittal, P. Kumar, P. R. Roy, R. Balasubramanian, B. B. Chaudhuri, "A Modified-LSTM Model for Continuous Sign Language Recognition using Leap motion",2019. https://ieeexplore.ieee.org/abstract/ document/8684245