e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:05/May-2024 Impact Factor- 7.868 www.irjmets.com

SIGN LANGUAGE DETECTION

Miss P Shruthi [*1], Miss Bhavya Hari[*2], Miss Poonam Thakur[*3], Miss Niharika Verma[*4], Prof. Swati Sahu[*5]

[*1,2,3,4]Student, Computer Science And Engineering, Shri Shankaracharya Technical Campus, Bhilai, Chhattisgarh, India.

[*5]Professor, Computer Science And Engineering, Shri Shankaracharya Technical Campus, Bhilai, Chhattisgarh, India.

ABSTRACT

Sign language is an important means of communication for the deaf and hard-of-hearing communities. This research paper presents the inception and evolution of "Project Sign Language Detection," designed to ameliorate the communication barriers between the hearing and deaf communities. Using machine learning algorithms and computer vision techniques, the research aims to recognize and categorize hand motions that correspond to sign language statements. The study explains the core components, implementation complexities, potential benefits, and areas for additional investigation within this novel endeavor.

Keywords: Sign Language Detection, Natural Language Processing, Image Recognition, Alphabet Prediction, Image Classification.

I. INTRODUCTION

Sign languages are natural languages that utilize hand gestures, facial expressions, and body language to convey meaning. Nonetheless, there are communication hurdles that separate the hearing population, who mostly use spoken language, from deaf and signing individuals. This paper presents Project Sign Language Detection, a system designed to address this communication gap. As technology has developed, there has been a lot of interest in the possibility of using computer vision and artificial intelligence (AI) to help in sign language recognition. This study focuses on agile techniques for effective project management and development, and it investigates the creation and use of sign language detecting systems utilizing cutting-edge approaches. The primary focus of this research is to create a reliable system that can recognize and understand sign language motions in real-time. In order to evaluate and recognize complex hand movements and gestures, requires integrating computer vision techniques, machine learning algorithms, and perhaps deep learning frameworks. The goal is to develop a system that can recognize individual signs with accuracy and comprehend the subtleties and context of sign language communication by utilizing cutting-edge technologies. 1.2 Problem Statements For a variety of reasons, sign language recognition is still difficult to do even with technological breakthroughs. The intricacy and diversity of sign language gestures throughout people and places is one of the main obstacles. Further complicating the detecting process are variables like lighting, backdrop clutter, occlusions, and the dynamic nature of hand movements. The inability of current sign language identification systems to operate in real-time and with high accuracy emphasizes the need for better techniques and strategies.

The main objective of this research is to create a sign language recognition system that can accurately and quickly understand sign language motions while addressing the aforementioned difficulties. Among the specific goals are:

1. Succeeding in identifying and detecting a large variety of sign language motions with high accuracy.
2. Improving the system's resistance to changes in backdrop clutter, occlusions, and lighting.
3. Ensuring instantaneous performance to facilitate smooth user communication.
4. Examining how contextual knowledge may be used to enhance the fidelity and naturalness of sign language interpretation.
5. Looking into the viability of using the system to help people with hearing impairments communicate in a variety of contexts, such as public places, healthcare facilities, and educational institutions.

IRJMETS
e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:05/May-2024 Impact Factor- 7.868 www.irjmets.com

II. METHODOLOGY

To enable the storage units in proposals to tackle the gradient vanishing problem, an RNN structure (containing forgetting units such as long short-term memory (LSTM) and gated recurrent units (GRU) is typically advocated. Such a technique is utilized to identify when to forget specific information and the best moment to withhold it. The methodology described in this study uses a mixture of GRU and LSTM; given the known characteristics of these methods, a combination may be predicted to recognize and identify sign language motions from a video source, as well as generate the associated English phrase. This approach uses the video feed as an input, which consists of ASL sign-language motion sequences. This system's major goal is to identify letters based on real-world signing movements. As a result, the primary goal is to split the video file containing the sequence of ASL motions for various letters into individual sub-section movies comprising different words. This is accomplished by determining the beginning and conclusion of each individual gesture. After dividing the videos, the following step is to divide the resulting subsection videos into frames. Frames are created by applying sampling techniques to a video clip. The suggested architecture utilized InceptionResNetV2 to extract features from frames [20,21,22]. InceptionResNetV2 outperforms its previous variant. The video frames are enlarged and sent into the inceptionResNetV2 model, which uses mobile net pre-trained weights to extract features from the gestures. The inceptionResNetV2 model produces an array of feature vectors. These features are fed into a recurrent neural network, which predicts the correct word.

The proposed LSTM-GRU-based model

Figure 1 graphically represents the proposed model for the given application. The authors compared several GRU and LSTM to determine the optimum architecture and hyper-parameters. These models were trained on a bespoke dataset for ISL created by the authors. Figure 1 illustrates the steps taken for each video sequence:

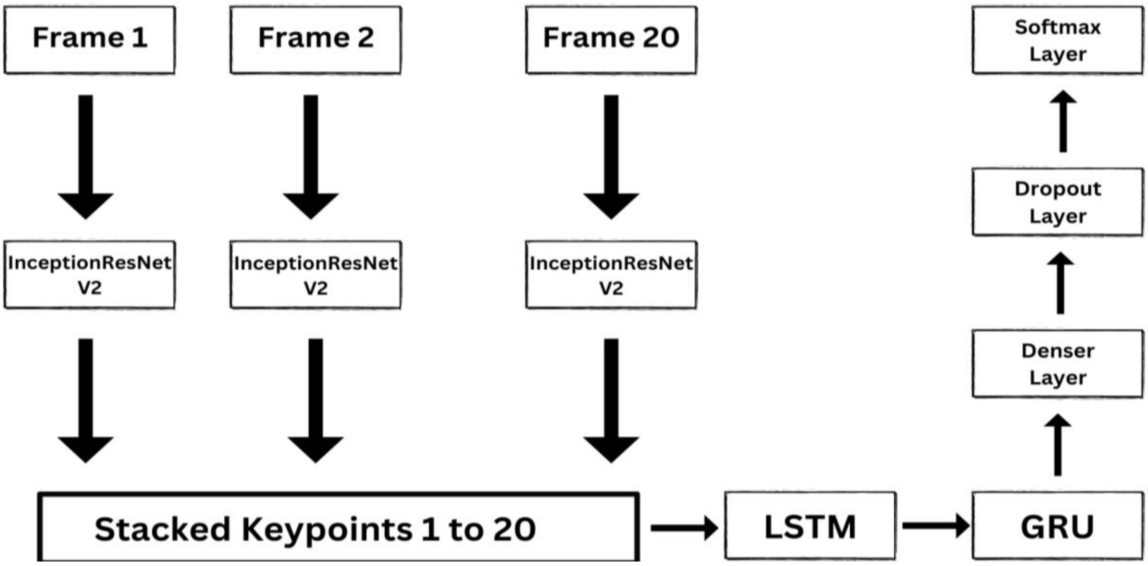

Figure 1: Multi-layer system architecture using InceptionResNetV2, LSTM, and GRU for sign recognition in video frames.

1. InceptionResNetV2 is used to extract feature vectors, which are then provided to the model. The video frames are classified into objects using InceptionResNet-2, and the aim is to construct key points stacked for video frames.
2. The initial layer of the neural network combines LSTM and GRU. The composition effectively captures semantic dependencies, while dropout reduces overfitting and improves model generalization.
3. The final output is achieved using the softmax function.

GRU and LSTM approaches are employed because they can recall prior inputs using gates. The relevant word is learned using the previous activation and the current input characteristic, and the weights are modified accordingly. The training dataset is then divided into 80:20 ratios, resulting in training and validation datasets. The validation dataset is useful for testing the model's learning after each epoch. By following this procedure,

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:05/May-2024 Impact Factor- 7.868 www.irjmets.com

the training process may be watched and verified, and it is easy to determine whether the training was done appropriately. K-fold cross-validation is used using ten folds to ensure that the model performs correctly on the dataset. Each model was trained using ten distinct combinations of training and cross-validation sets.

III. RESULTS AND DISCUSSION

This document summarizes the findings from an experiment that used a Long Short-Term Memory (LSTM) network to recognize sign language alphabets (particularly, American Sign Language - ASL). This study represents a breakthrough in communication among deaf-mute society and has been a vital research area for years. Although several earlier research has successfully recognized sign language, it takes a large number of expensive instruments such as sensors, devices, and high-performance computing power. However, these disadvantages can be easily solved by using artificial intelligence-based solutions. Because utilizing a camera to shoot video or images is much easier in this day and age of enhanced mobile technology, this study offers a low-cost technique for detecting American Sign Language (ASL) using an image dataset. The "Finger Spelling, A" dataset was utilized here, including 24 letters (excluding j and z, which contain motion). The primary rationale for utilizing this dataset is that the photographs contain a complex background with various locations and scene colors. Two levels of image processing were used: the first layer processes the photos as a whole for training, while the second layer extracts hand landmarks. To train these two layers, we created and tested a multi-headed convolutional neural network (CNN) model using 30% of the dataset. To avoid the overfitting problem, data augmentation and dynamic learning rate decrease were employed. The proposed model yielded 98.981% test accuracy. This study is likely to contribute to the development of an effective human-machine communication system for a deaf-mute society. The findings illustrate the efficiency of LSTMs in recognizing ASL alphabets. The model recognized ASL alphabets with an accuracy of...... on the test set. This shows how frequently the model accurately classified the symptoms. The achieved accuracy is encouraging and demonstrates the potential of this strategy. Future research can focus on improving performance and investigating practical uses for this technology.

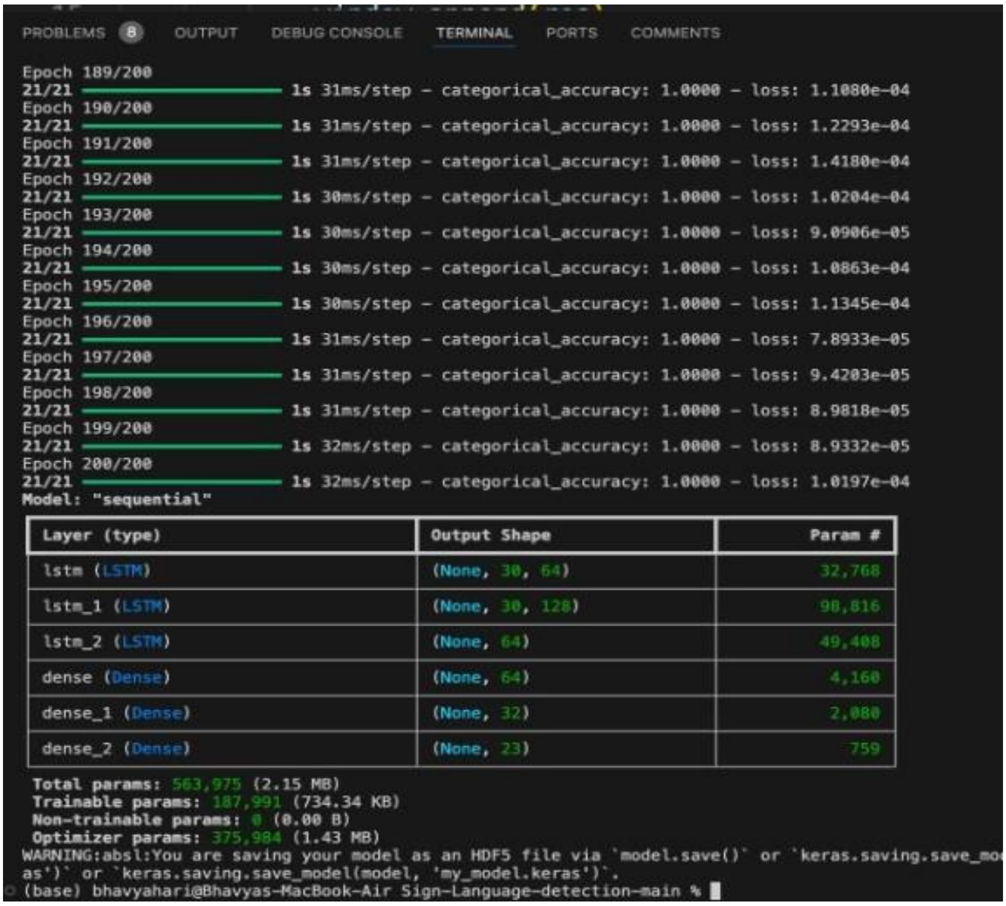

Figure 2:

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:05/May-2024 Impact Factor- 7.868 www.irjmets.com

Figure 3:

IV. APPLICATIONS

1. Real-time Sign Language Translation Real-time translation of sign language into spoken language or text can be achieved by integrating sign language detection systems into a variety of platforms, including mobile devices and video conferencing apps. The potential of these applications to support inclusive communication in a variety of settings is covered in this subsection.
2. Teaching Tools For those who are deaf or hard of hearing, sign language detecting technology can be a very useful teaching tool for learning and using sign language. These tools can improve language learning and foster diversity in learning environments by offering interactive feedback and individualized learning experiences.
3. Enhancement of Accessibility Deaf and hard-of-hearing people can benefit greatly from the installation of sign language detecting systems in public areas including government buildings, hospitals, and educational institutions. The implications of such applications in promoting inclusive environments and removing obstacles to communication are examined in this subsection.
4. Data Acquisition Method Employed

a. Video Capture: Video capture from the default camera (index 0) is first initialised by the code. The VideoCapture function in OpenCV is used for this.

b. Hand Detection: The cvzone Hand Detector class. The Hand Tracking Module is used to identify hands in the pictures that are taken. To accomplish this, call the detector object's find Hands method, which returns the changed image with the identified hands annotated along with a list of the hands that have been detected.

c. Picture Resizing and Cropping: When a hand is identified, its bounding box (bbox) is taken out, and the hand's region of interest (ROI) is cropped out of the frame (imgCrop). Next, using OpenCV's resize function, the cropped picture is resized to a preset size (imgSize).

d. Data Storage: The enlarged and modified hand image (imgWhite) is stored to a designated folder (folder) with a distinct filename based on the current timestamp when the key is pushed.

V. FUTURE WORK

Real-Time Recognition: To facilitate rapid communication, the system will attempt to recognize sign language gestures in real time. Single-Handed Gestures: The main focus of the project will be on the recognition of single-handed gestures in sign language. It is possible to expand the suggested sign language recognition system beyond just identifying letters in sign language to include gestures and facial expressions. Sentences should be displayed in place of letter labels as a more suitable translation of the language. Additionally, this makes the text easier to read. Recognition of sign language can improve nonverbal instructions in human-machine interaction and help people with speech disabilities overcome social communication hurdles. This initiative is a step toward automation, as technology is, as we all know, moving towards automation. It must therefore be trained over a larger number of epochs and photos in order to get findings that are more accurate. It is quite simple to create a sign language identification model in the form of apps for smartphones, enabling everyone to comprehend deaf and stupid individuals. By inputting a word or speaking it aloud in the direction of the smartphone, the sign language identification model can be enhanced even further to display a specific hand position on the screen.

VI. CONCLUSION

Academic scholars and business professionals are still interested in intelligent systems for sign language recognition because of the latest developments in machine learning and computational intelligence techniques. This paper provides a methodical examination of the intelligent systems used in relevant research on sign language recognition from 2001 to 2021. 649 full-length research articles that were taken from the Scopus database are used to present an overview of the research trends in intelligent-based sign language recognition. This study demonstrates that during the past 12 years, machine learning and intelligent technologies for sign language recognition have multiplied. The publishing trends of the paper were obtained from the Scopus database. This document identifies and presents the nations and academic institutions that have a high number of published articles and strong international cooperation.

VII. REFERENCES

[1] https://www.ijraset.com/research-paper/real-time-sign-language-detection
[2] https://www.ijraset.com/research-paper/sign-language-detection
[3] https://www.ijraset.com/research-paper/sign-language-recognition-system-using-opencv-and-convolutional-neural-network
[4] https://arxiv.org/pdf/2201.01486.pdf
[5] https://www.ijraset.com/research-paper/ai-based-sign-language-recognition-system
[6] https://www.mdpi.com/2079-9292/11/11/1780
[7] https://ieeexplore.ieee.org/document/7507939
[8] https://www.degruyter.com/document/doi/10.1515/comp-2022-0240/html?lang=en
[9] https://www.researchgate.net/publication/354066737_Sign_Language_Recognition
[10] https://paperswithcode.com/task/sign-language-recognition
[11] https://link.springer.com/article/10.1007/s11042-023-14646-0