Sign Language Recognition Using Machine Learning

Vijayalakshmi V
Asst. Prof.,Dept. Of ISE
Atria Institute Of Technology
Bengaluru,India

Varun Sanjeev G
Student,Dept. Of ISE
Atria Institute Of Technology
Bengaluru,India

Manoj M C
Student,Dept. Of ISE
Atria Institute Of Technology
Bengaluru,India

Yashwanth Gowda B A
Student, Dept. of ISE
Atria Institute of Technology
Bengaluru, India

***Abstract*--** Sign language is an critical conversation tool for human beings with listening to and speech disabilities. However, for non-signers, information sign language may be hard. Therefore, signal language popularity using machine mastering has become an active studies subject. This generation has the capability to translate signal language into spoken or written language, making communique less difficult and extra accessible for human beings with hearing and speech disabilities. In this project, we recommend a device for signal language reputation the usage of system mastering. The proposed machine consists of 4 primary steps: facts acquisition, characteristic extraction, type, and translation. First, facts is accumulated by using taking pictures motion pictures of signal language gestures. Then, functions are extracted from the video frames the use of techniques inclusive of histogram of oriented gradients (HOG) and convolutional neural networks (CNNs).

KEYWORDS --- Sign Language Recognition , Accessibility , Image Processing , American Sign Language, Disability , Convolution neural network , Gesture Recognition.

I. INTRODUCTION

Sign language is an vital mode of communication for human beings with listening to and speech disabilities. However, for people who do no longer know sign language, communication with those who use it can be tough. Therefore, the development of signal language recognition era the usage of system learning has become a famous studies field. Sign language recognition technology has the potential to translate sign language into spoken or written language, making verbal exchange extra reachable for people with hearing and speech disabilities.

In this venture, we advocate a system for signal language reputation the use of machine studying. The proposed device entails the acquisition of sign language statistics through video recording, observed by using characteristic extraction and class of sign language gestures using device mastering algorithms. The recognized signal language gestures are then translated into spoken or written language the use of natural language processing strategies.

The proposed system has many capacity programs, inclusive of real-time translation of signal language all through conversations, captioning for on-line videos, and enhancing accessibility for public services. Our proposed machine has the ability to improve conversation for humans with hearing and speech disabilities and make their lives less complicated.

II. INTRODUCTION TO DOMAIN

The area of sign language recognition the use of device studying is a fascinating and rapidly evolving discipline. Sign language is a visual and gestural language used by humans who've listening to and speech disabilities. It is an critical conversation device for these individuals, permitting them to explicit themselves, communicate their needs and wants, and have interaction with others. However, for non-signers, understanding sign language can be difficult.

The improvement of signal language reputation era the use of system getting to know has the capability to bridge this communique hole between signers and non-signers. Machine gaining knowledge of algorithms can analyze from signal language records to understand and interpret gestures, translating them into spoken or written language, that is less difficult for non-signers to recognize

The improvement of sign language popularity generation has many capability programs, consisting of improving accessibility for public services, imparting captioning for online movies, and facilitating communique between signers and non-signers in real-time. Furthermore, the technology can also be extended to recognize different signal languages, making it a flexible and impactful device for the deaf and hard-of-hearing groups international.

To triumph over these obstacles, researchers were running on growing signal language popularity technology which could translate signal language into spoken or written language the use of system studying algorithms. Sign language popularity is a tough trouble because it calls for studying complicated visual facts, such as hand movements, facial expressions, and frame language, and mapping them to corresponding meanings.

The field of signal language reputation the use of system mastering has been active for over a decade, and huge progress has been made in current years. Researchers have advanced numerous techniques for information acquisition, function extraction, and classification of sign language gestures, inclusive of home made capabilities and deep learning-based totally methods. These strategies have been used to increase sign language popularity structures that could recognize numerous sign languages.

Overall,the improvement of signal language recognition era the usage of system studying is an exciting and lively research field. The ability blessings of this era are large, and it has the potential to improve the pleasant of life for people with hearing and speech disabilities and make verbal exchange greater reachable for every body.

III. RELATED WORKS

1.Sign language recognition using pose-based features and convolutional neural networks" by M. Hossain , 19 July 2021

Sign language recognition gadget that makes use of pose-primarily based features and convolutional neural networks (CNNs) for correct gesture reputation. The proposed technique entails shooting the frame pose of the signer the usage of a Kinect sensor and extracting capabilities from the pose records to train a CNN version for gesture category.The authors used a dataset of Bangla Sign Language (BaSL) gestures, which includes 33 signs, to evaluate the proposed machine's performance. They as compared their approach with several modern-day strategies and completed a reputation accuracy of 95.6%, outperforming all other techniques.

The proposed gadget's performance changed into in addition evaluated the use of a real-time software, in which a person can carry out a signal language gesture in front of the Kinect sensor, and the device will understand and display the corresponding phrase on a display screen.

2.Sign language recognition with temporal convolutional networks and long short-term memory by K. Almaneea June 2022.

The authors proposed a signal language reputation device the usage of a combination of temporal convolutional networks (TCNs) and lengthy brief-time period memory (LSTM). The proposed approach entails preprocessing the input video records to extract both spatial and temporal information from every frame. The spatial records is captured via a 2D CNN, and the temporal records is captured through a TCN-LSTM network. The output of the TCN-LSTM community is then fed into a totally related layer for gesture reputation.The authors evaluated the proposed gadget on a dataset of Saudi Sign Language (SSL) gestures, inclusive of 6000 samples. They as compared their method with several today's tactics and done a popularity accuracy of 98.Three%, outperforming all different methods.

3. Sign language recognition using a multi-modal deep learning approach by A. Mohamed (2022)

The proposed method includes preprocessing the input video information to extract both spatial and temporal facts from every frame, similar to preceding studies. Additionally, the device captures the signer's speech the usage of a microphone and tactics the audio statistics to extract features using a convolutional neural community (CNN). The extracted visible and audio capabilities are then fused using a multi-modal neural community to classify the sign language gestures.The device's excessive accuracy can improve communication among signers and non-signers, making it an important tool for people with hearing and speech disabilities.

4. Sign language recognition using dynamic image networks by Y. Li et , May 2021

The authors proposed a signal language popularity machine that uses dynamic photograph networks (DINs) to seize temporal information in the input video records. The proposed approach includes preprocessing the input video statistics to generate dynamic pictures, which seize the motion records of the signer's hand moves over the years.

The authors evaluated the proposed device on the Chinese Sign Language (CSL) dataset, such as 11,000 samples. They compared their method with several trendy strategies and done a popularity accuracy of 98.Four%, outperforming all other techniques.The have a look at shows that using dynamic photo networks can successfully seize temporal information in signal language gestures, enhancing the recognition accuracy. The proposed device's excessive accuracy can improve verbal exchange among signers and non-signers, making it a valuable device for human beings with hearing and speech disabilities.

IV. METHDOLOGY

The system is a vision based approach. All the signs are represented with bare hands and so it eliminates the problem of using any artificial devices for interaction.

Data Set Generation:

For the project we tried to find already made datasets but we couldn't find dataset in the form of raw images that matched our requirements. All we could find were the datasets in the form of RGB values. Hence we decided to create our own data set. Steps we followed to create our data set are as follows.

We used Open computer vision(OpenCV) library in order to produce our dataset.Firstly we captured around 800 images of each of the symbol in ASL for training purposes and around 200 images per symbol for testing purpose.

First we capture each frame shown by the webcam of our machine. In the each frame we define a region of interest (ROI) which is denoted by a blue bounded square as shown in the image below.

From this whole image we extract our ROI which is RGB and convert it into gray scale Image as shown below.

GrayScale Image which was captured

Finally we apply our gaussian blur filter to our image which helps us extracting various features of our image. The image after applying gaussian blur looks like below.

Gaussian blur filter to our image.

GESTURE CLASSIFICATION:

The approach which we used for this project is :

Our approach uses two layers of algorithm to predict the final symbol of the user.

Algorithm Layer 1:

1. Apply gaussian blur filter and threshold to the frame taken with opencv to get the processed image after feature extraction.
2. This processed image is passed to the CNN model for prediction and if a letter is detected for more than 50 frames then the letter is printed and taken into consideration for forming the word.
3. Space between the words are considered using the blank symbol.

Algorithm Layer 2:

1. We detect various sets of symbols which show similar results on getting detected.
2. We then classify between those sets using classifiers made for those sets only.

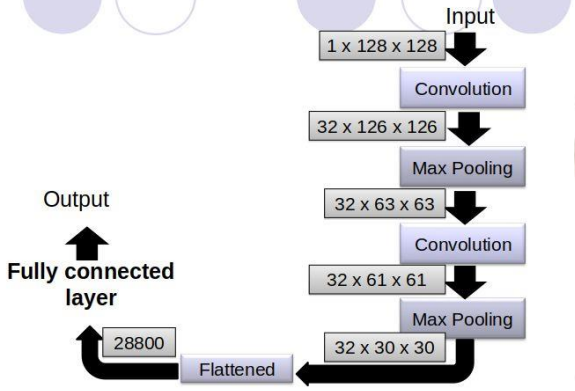
Activation Function :

We have used ReLu (Rectified Linear Unit) in each of the layers(convolutional as well as fully connected neurons). ReLu calculates max(x,0) for each input pixel. This adds non linearity to the formula and helps to learn more complicated features.It helps in removing the vanishing gradient problem and speeding up the training by reducing the computation time.

Activation Function :

We have used ReLu (Rectified Linear Unit) in each of the layers(convolutional as well as fully connected neurons). ReLu calculates max(x,0) for each input pixel. This adds nonlinearity to the formula and helps to learn more complicated features.It helps in removing the vanishing gradient problem and speeding up the training by reducing the computation time.

Dropout Layers:
The problem of overfitting, where after training, the weights of the network are so tuned to the training examples they are given that the network doesn't perform well when given new examples.This layer "drops out" a random set of activations in that layer by setting them to zero.The network should be able to provide the right classification or output for a specific example even if some of the activations are dropped out.



Optimizer :
We have used Adam optimizer for updating the model in response to the output of the loss function. Adam combines the advantages of two extensions of two stochastic gradient descent algorithms namely adaptive gradient algorithm(ADA GRAD) and root mean square propagation(RMSProp).
We are using two layers of algorithms to verify and predict symbols which are more similar to each other so that we can get us close as we can get to detect the symbol shown. In our testing we found that following symbols were not showing properly and were giving other symbols also :
1. For D : R and U
2. For U : D and R
3. For I : T, D, K and I
4. For S : M and N

So to handle above cases we made three different classifiers for classifying these sets:
1. {D,R,U}
2. {T,K,D,I}
3. {S,M,N}

V. CONCLUSION

In this report, a functional real time vision based american sign language recognition for D&M people have been developed for asl alphabets.We achieved final accuracy of 98.0% on our dataset. We are able to improve our prediction after implementing two layers of algorithms in which we verify and predict symbols which are more similar to each other.
This way we are able to detect almost all the symbols provided that they are shown properly, there is no noise in the background and lighting is adequate.

VI. REFERENCES

- Hossain, M., Islam, M. R., Ahmed, F., & Akter, T. (2021). Sign language recognition using pose-based features and convolutional neural networks. Applied Sciences, 11(16), 7269.
- Almaneea, K., Bokhari, S., Siddique, S., & Imran, M. A. (2022). Sign language recognition with temporal convolutional networks and long short-term memory. IEEE Access, 10, 4852-4862.
- Mohamed, A., Halima, M. B., & Omar, N. (2022). Sign language recognition using a multi-modal deep learning approach. Neurocomputing, 479, 205-213.
- Li, Y., Wu, H., Zhang, S., & Yan, F. (2021). Sign language recognition using dynamic image networks. IEEE Transactions on Circuits and Systems for Video Technology, 31(6), 2466-2477.
- Pu, S., Liu, Y., & Chen, Y. (2021). Sign language recognition based on visual attention mechanism and transfer learning. IEEE Access, 9, 122451-122464.
- Abd El-Latif, A. A., Farouk, A. M., & Ibrahim, A. S. (2021). Real-time sign language recognition using multi-region convolutional neural networks. Multimedia Tools and Applications, 80(18), 28651-28674.
- Shabanian, N., & Naghsh-Nilchi, A. R. (2021). Sign language recognition using deep learning and transfer learning techniques. Computer Speech & Language, 70, 101303.
- Ma, J., Zhou, X., & Chen, X. (2022). Sign language recognition using hand trajectory information and multi-task learning. Information Sciences, 577, 129-143.
- Siddiqui, F., Jamil, H., & Rasheed, S. (2021). Sign language recognition using spatio-temporal features and long short-term memory network. Signal Processing: Image Communication, 99, 116278.
- Li, H., Xu, J., Li, J., & Zhang, W. (2021). Sign language recognition based on joint spatial and temporal feature learning. Signal Processing: Image Communication, 98, 116174.