ISSN NO: 9726-001X
Volume 8 Issue 01 2020

JOURNAL OF CURRENT SCIENCE

The Use of Deep Learning Techniques for the Detection and Recognition of Indian Sign Language Symbols

Mr. S. Satheesh, Ms. S. Chandra Priyadharshini, Mrs. K. Lavanya, Mr. S. Sugavanam

Assistant Professor [1,3] Associate Professor [2,4]

Department of CSE,

ssatheesh@actechnology.in, chandrapriyadharshini.s@actechnology.in, klavanya@actechnology.in, sugavanam.s@actechnology.in

Arjun College of Technology, Thamaraikulam, Coimbatore-Pollachi Highway, Coimbatore, Tamilnadu-642 120

Abstract: It is always a huge problem to communicate with someone who has a hearing impairment. The problems in character categorization in Indian Sign Language (ISL) are the focus of the study reported in this article. Those who are deaf or hard of hearing cannot fully communicate by sign language alone. If you've never learned this language, you could find the disabled person's gestures confusing or disorganised. It is important for communication to be reciprocal. We provide an Indian Sign Language-based sign language recognition system in this article. For this analysis to work, the user's webcam must be able to record hand motions, and the system's prediction and display of the image's name are essential. A number of computer vision methods, including grayscale conversion, dilatation, and mask operation, are used to the collected picture throughout the processing phases. Our model is trained and the photographs are identified using a Convolutional Neural Network (CNN).

Keywords: Convolutional Neural Network (CNN), Hand Gesture, Deaf people, Sign Language, Sign language Recognition (SLR), ISL (Indian Sign Language).

ISSN NO: 9726-001X
Volume 8 Issue 01 2020

I. INTRODUCTION:

To stay alive in a social setting, communication is crucial. Sign language is a means of communication for the deaf and dumb, but it is challenging for others who are not deaf or dumb to comprehend. There has been a lot of research on ASL identification, however Indian sign language is quite different from ASL. While American Sign Language uses one hand for communication, International Sign Language uses two (20 out of 26). When utilising both hands, it's easy to hide details due to the overlapping of the hands. Additionally, ISL gesture identification studies have been hindered by a paucity of datasets and the fact that sign language differs based on location. The primary objective of this study is to begin the process of using Indian sign language to facilitate communication between hearing and non-hearing individuals. Not only will this project's expansion to words and frequent phrases facilitate communication for the deaf and dumb, but it may also contribute to the creation of autonomous systems that can comprehend and aid those who are hard of hearing. This paper's main objective is to learn the matching gesture for Indian Sign Language letter recognition. Indian Sign Language has lagged behind American Sign Language when it comes to research on gesture and sign language identification. We want to address this problem by using computer vision and machine learning methods to extract characteristics and categorise motions from webcam-accessible photos, rather than relying on high-end technology such as the Kinect or gloves.

II. LITERATURE REVIEW

Sign identification in visual media has been the subject of several efforts, using a wide range of approaches and algorithms, according to a literature review for the suggested framework. As stated in Jing-hao Sun In order to apply the CamShift algorithm to recognise hand motions in real-time, the human hand was isolated from its complicated environment. The next step is to identify 10 shared digits by training a convolutional neural network to detect the area of the hand motions captured in real-time. There are a total of 1600 photos in the training dataset, 4000 hand gestures, and 400 of each kind in the proposed system's dataset. About 98.3% of the time, this experiment gets it right. Hasan was able to identify gestures by comparing brightness factors with the usage of scaled normalisation. To segment the input photos, thresholding methods are used with a black backdrop. Any segmented picture has its coordinates adjusted at the X and Y axis origins so that they align with the hand unit's centroid. And the centre of mass of the picture is found. Wysoski et al. offered postures that are invariant to rotation by means of a boundary histogram. A camera was used to take the input picture, which was then processed using a skin colour detection

ISSN NO: 9726-001X
Volume 8 Issue 01 2020

filter. Next, a clustering approach was used to locate the category borders in the pooled image by means of a conventional contour tracking algorithm. From the image, grids were generated and their borders were standardised. An American Sign Language (ASL) symbol recognition system was created by Geethu Nath and Arun C.S. using the ARM CORTEX A8 CPU. The system can distinguish between letters and numbers by using the template matching algorithm and the Jarvis algorithm, respectively.Kumud Tripathi built a system to identify continuous ISl motions using PCA and several distance classifiers. We use Orientation Histogram to extract keyframe characteristics from our own data collection, and then we feed them into the device.Classifying sequential data using the Modified k-Nearest Neighbour (MKNN) method was suggested by Noor Tubaiz. Wearing data gloves allows for the detection of hand motions. Raw feature vectors from the past and the future are used to compute window-based statistical characteristics, which are added to the raw data. The suggested framework was built utilising innovative approaches that are based on current systems (ISL) in order to identify concepts in ISL.B. Bauer et al. propose a system for continuous sign language recognition; describe it. This system relies on HMM pictures, which are continuous hidden Markov models. It makes use of GSL, or German sign language. The gadget is given feature vectors that stand in for human indications.

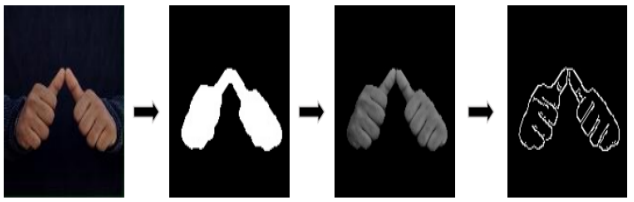
III.METHODOLOGY

Image Aquistion:

It is the action of extracting an image from a source, typically a hardware-based source, for process of image processing. WebCamera is the hardware-based source in our project. It is the first step in the workflow sequence because no processing can be done without an image. The picture that is obtained has not been processed in any way.

Segmentation:

The method of separating objects or signs from the context of a captured image is known as segmentation. Context subtracting, skin-color detection, and edge detection are all used in the segmentation process. The motion and location of the hand must be detected and segmented in order to recognise gestures.



a) Raw image b) Skin mask (gray scale) c) Skin (gray scale) d) Canny edge detection

Features Extraction:

The preprocessed pictures are then utilised for sign identification or classification using predefined characteristics such as shape, contour, geometry (position, angle, distance,

ISSN NO: 9726-001X
Volume 8 Issue 01 2020

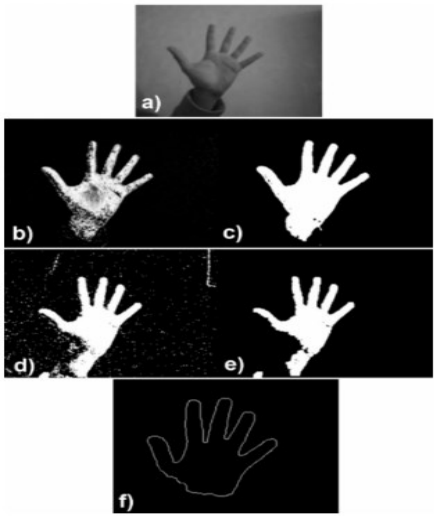
etc.), colour, and histogram. The dimensionality reduction method includes feature extraction, which organises and splits a massive dataset into smaller, more manageable groupings. Consequently, processing would become less complicated. Most importantly, these huge datasets include a great deal of variables. Massive computing resources are required to handle these variables. Hence, function extraction helps with optimum feature extraction from big data sets by minimising data size by variable selection and combination into functions. These properties correctly and uniquely describe the actual data collected while being easy to utilise.

Preprocessing:

Each picture frame is preprocessed to eliminate noise using a variety of filters including erosion, dilation, and Gaussian smoothing, among others. The size of an image is reduced when a colour image is transformed to grayscale. A common method for reducing the amount of data to be processed is to convert an image to grey scale.The phases of preprocessing are as follows:

Morphological Transform (Morphological Transform) :

Morphological operations use a structuring feature on an input image to create a similar-sized output image. It compares the corresponding pixel in the input image with its neighbours to determine the value of each pixel in the output image. There are two different kinds of morphological transformations Erosion and Dilation.



(a) original image with hand; (b) image of hand after skin color detection; (c) after morphological operations and binarization; (d) image of hand after background extraction ; (e) after binarization and morphological operations; (f) hand's contour made of image c and image e concatenation

i. Dilation:

The maximum value of all pixels in the neighbourhood is the value of the output pixel. A pixel in a binary image is set to 1 if all of its

neighbours have the value 1 Morphological dilation increases the visibility of artefacts and fills in small gaps.

ii. Erosion:

The o/p pixel's value is the minimum of all pixels in the neighbourhood. A pixel in a binary image is set to 0 if all of its neighbours have the value 0.small artefacts are eroded away by morphological erosion, leaving behind substantial objects.

Blurring:

Adding a low-pass filter to an image is an example of blurring. The word "low-pass filter" refers to eliminating noise from an image while leaving the rest of the image intact in computer vision. A blur is a simple operation that must be completed before other tasks such as edge detection.

Recognition:

We'll use classifiers in this case. Classifiers are the methods or algorithms that are used to interpret the signals. Popular classifiers that identify or understand sign language include the Hidden Markov Model (HMM), KNearest Neighbor classifiers, Support Vector Machine (SVM), Artificial Neural Network (ANN), and Principle Component Analysis (PCA), among others. However, in this project, the classifier will be CNN. Because of its high precision, CNNs are used for image classification and recognition. The CNN uses a hierarchical model that builds a network, similar to a funnel, and then outputs a fully-connected layer in which all neurons are connected to each other and the output is processed.

Text output:

Understanding human behaviour and identifying various postures and body movements, as well as translating them into text.

IV.IMPLEMENTATION

Creating the sign language recognition dataset:

Two directories, "train" and "test," have ten folders containing pictures recorded using the create gesture knowledge. Any frame that recognises a hand inside the ROI (region of interest) produced may be moved to one of these directories.do anything using Python. We use OpenCV to push the live cam stream and construct a ROI, which is the specific region of the frame where we want to locate the hand for the movements; this is all we need to create the dataset. To separate the background from the foreground, we typically compute the total weighted average of the background and then decrypt it from the frames that include an item in the foreground.To do this, we need to calculate the total weight for each frame and then take the average of all frames in the context.In order to find any item that obscures the backdrop, we usually deduct the cumulative

ISSN NO: 9726-001X
Volume 8 Issue 01 2020

average for the background from each frame that we read after sixty frames.

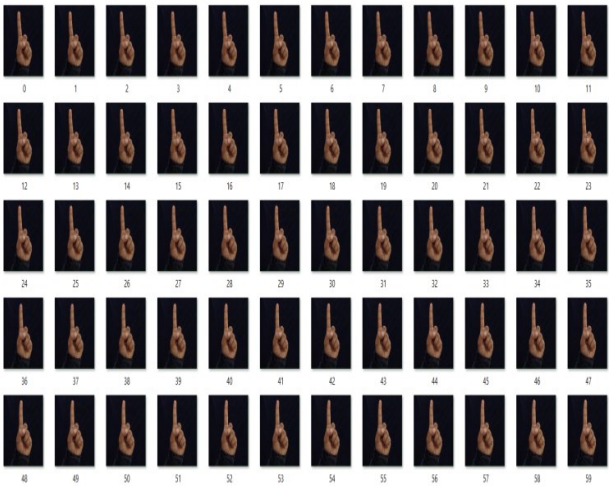


FIG: DATASET

V RESULTS

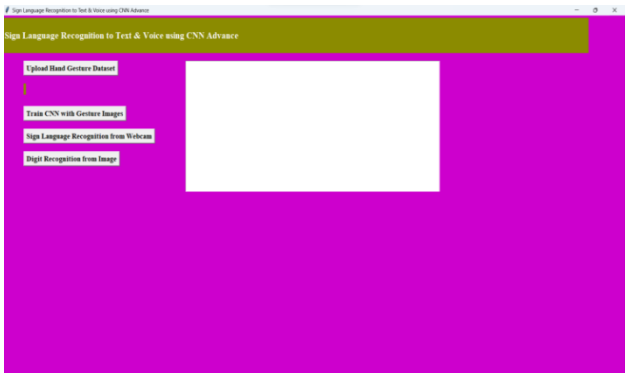


FIG :HOME SCREEN

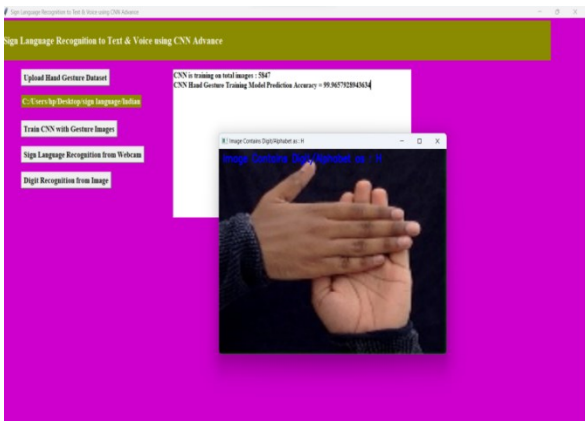


FIG : DIGIT RECOGNITION FROM IMAGE

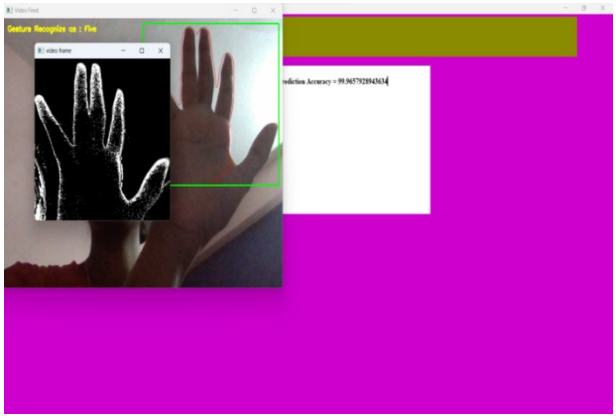


FIG: SIGN LANGUAGE RECOGNITION FROM WEBCA

VI CONCLUSION

When given the right context, the Sign Language Recognition (SLR) system can recognise a set of created signs and convert them into text or voice. When it comes to creating efficient interactions between humans and machines, gesture recognition is crucial. For this assignment, we looked at constructing a model using a CNN. A validation accuracy of around 95% is the end outcome of this. Future work should improve the Image Processing component to make the system bidirectionally interactive, meaning it can translate between spoken and sign language.

VII.REFERENCES

[1] Jing-Hao Sun,Ting-Ting Ji, Shu-Bin Zhang, Jia-Kui Yang, Guang-Rong Ji "Research on the Hand Gesture Recognition Based on Deep Learning",07 February 2019 .

[2] Mokhtar M. Hasan, Pramoud K. Misra,

(2011). “Brightness Factor Matching For Gesture Recognition System Using Scaled Normalization”, International Journal of Computer Science Information Technology (IJCSIT), Vol. 3(2).

[3] Simei G. Wysoski, Marcus V. Lamar, Susumu Kuroyanagi, Akira Iwata, (2002). “A Rotation Invariant Approach On Static-Gesture Recognition Using Boundary Histograms And Neural International Journal of Artificial Intelligence Applications (IJAIA), Vol.3, No.4, July 2012.

[4] Stergiopoulou, N. Papamarkos. (2009). “Hand gesture recognition using a neural network shape fitting technique,” Elsevier Engineering Applications of Artificial Intelligence, vol. 22(8), pp. 1141– 1158, doi: 10.1016/j.engappai.2009.03.008.

[5] V. S. Kulkarni, S.D.Lokhande, (2010) “Appearance Based Recognition of American Sign Language Using Gesture Segmentation”, International Journal on Computer Science and Engineering (IJCSE), Vol. 2(3), pp. 560-565.

[6] Geethu G Nath and Arun C S, "Real Time Sign Language Interpreter," 2017 International Conference on Electrical, Instrumentation, and Communication Engineering (ICEICE2017)

[7] Kumud Tripathi, Neha Baranwal and G. C. Nandi, "Continuous Indian Sign Language Gesture Recognition and Sentence Formation", Eleventh International MultiConference on Information Processing2015 (IMCIP-2015), Procedia Computer Science 54 (2015) 523 – 531

[8] Noor Tubaiz, Tamer Shanableh, and Khaled Assaleh, "Glove-Based Continuous Arabic Sign Language Recognition in User-Dependent Mode," IEEE Transactions on Human-Machine Systems, Vol. 45, NO. 4, August 2015.

[9] B. Bauer,H. Hienz "Relevant features for video-based continuous sign language recognition", IEEE International Conference on Automatic Face and Gesture Recognition, 2002.

[10] Pigou, L., Dieleman, S., Kindermans, P.-J., Schrauwen, B. (2014). Sign Language Recognition Using Convolutional Neural Netwoks