RESEARCH ARTICLE OPEN ACCESS

SIGN LANGUAGE TRANSLATOR WITH TENSOR FLOW OBJECT DETECTION AND PYTHON

Abhishek Nachankar*, Rahul khapre, Yash Gour, Ritik Warkade, Rahul Kothe**
*(Assistant Professor,*Dept. Of Computer Technology, KDK College of Engineering, Nagpur*
Email: abhishek.nachankar@kdkce.edu.in)
** (*Dept. Of Computer Technology, KDK College of Engineering, Nagpur*
Email: khaprerahul312@gmail.com, Yashgour252@gmail.com, Ritikwarkade987@gmail.com, rahulkothe510@gmail.com)

Abstract:

Sign language, which is considered the dominant language uses manual contact and communication for deaf and hard of hearing. The language of the body to communicate expressions and to build an identity plays a major role. Recognition of sign language is currently an emerging area of study to promote contact with the deaf community. Recognition of hand gesture serves as a key for overcoming several barriers and creating comfort for existence for people. Machines capacity to comprehend human beings in a vast array of practices and their sense can be used in applications. Sign language Recognition is one particular area of interest. This article offers an in-depth study of state-of-the-art techniques used in recent hand movements and signs. Study into language comprehension. A system that will serve as a learning tool for sign language starters that involves hand recognition has been developed.

Keywords — **Python ,CNN ,Image processing, Datasets**

I. INTRODUCTION

Hand movements are used as a means for individuals to express thoughts and emotions, helping to enhance the knowledge given in our everyday conversation. Sign language is a standardized type of hand gestures which are used as a communication device, including visual movements and signs. Sign language is a valuable method for everyday communication for the deaf and speech-impaired community. Sign language requires the use of various parts of the body to convey information, including fingers, hand, arm, head, body and facial expression. Sign language is not popular in the hearing population, however, and fewer are able to understand it. This creates a genuine obstacle to contact between the deaf community and the rest of society, an issue that has yet to be completely addressed to this day. Object detection, especially if there are several objects in the image, is a crucial step for excellent object recognition. Standalone objects with well-defined borders and centers, as opposed to history detection of Objects Techniques localize object instances in an image or a frame Video. A typical technique for object detection is to formulate it as a problem of classification, where either convolutions or sliding windows are used to identify small regions over feature maps or the image itself. The human being's effective communication is through gestures. People use gestures to express their thoughts and feelings, to draw attention, and to emphasize communication.

The gestures are different than the SL. This makes the SL difficult to learn by the general public. The solutions have, however, been provided by active research in computer vision and machine learning.

II. LITERATURE SURVEY

- For images or videos, many appliances may make use of fast and lightweight automated object detection. The technology industry has been continually introducing computational and hardware solutions, such as computers with impressive processing and storage capacities, over the past five years. Object detection methods, however, typically require either high processing power or large processing power, Availability of storage, which makes it difficult for resource-restricted devices to perform real-time detection without a strong server connection. In template matching, before a match is found, we slide a template image over a source image. But, since it has significant drawbacks, it is not the best tool for object recognition.

- That technique isn't really durable. The following variables make the matching prototype a poor option for object detection.
  1. Rotation makes this technique inefficient.
  2. Size (known as scaling) also impacts this.
  3. Photo-metric alterations (e.g. brightness, contrast, hue etc.)
  4. Changes in the view point of the distortion form (Affine).
- Human Based Translation of Neural Sign Language Estimation Key point:
  - There have been several ways of understanding hand languages that are used with a single hand to describe letters of the alphabet. As each letter of the alphabet simply corresponds to a particular hand shape, it is comparatively simpler than understanding sign languages. To identify the English alphabet with 92 percent recognition precision, the authors used depth cameras and the random forest algorithm. For understanding the American Sign Language, a pose estimation method of the upper body represented by seven main points was suggested (ASL).
  - We also note an approach by Kim et al. to identify the language of the Korean hand by examining latent hand picture characteristics. There are other ways by which signals are captured by using motion sensors like electromyography, for example (EMG) Sensors, RGB cameras, sensors from kinect, and Controller of leap motion or their combinations.

- Hand gesture segmentation is actually a segmentation of the hand Done by using the technique of hand skin threshold. The device due to lighting, it does not yield successful outcomes .
- Conditions, interference with skin tone, and dynamic backgrounds increased noise. There are three kinds of skin color that are International Machine Learning and Computing Journal.
- There are three kinds of skin color that are Detection: the method of explicit range the nonparametric method process and the method of parametrics.
- The explicit variety The method divides the pixel class into skin- and skin- and from the allocated set of colors, non-skin-based

forms. This approach is primarily used because of its non-complex existence and reasonable computing pace.

III. PROPOSED WORK

- **Methodology**

- The system will be implemented via a desktop with a desktop 1080P Web Camera in Full-HD. camera is going to catch the Pictures of the hands that are being fed into the device. Remember that the signer will change the frame size to allow the frame to be changed. The device is capable of recording the orientation of the signer's Hand. When the camera has captured the gesture already from the camera the framework classifies the test sample from the consumer and compares it in the gestures in a dictionary that are stored to the corresponding .
- The performance is shown to the user on the screen. A. Collection of training data, augmentation of images, and Cropping Procedures The collection of static SLR datasets was achieved with the use of continuous image capture using python pictures automatically cropped and transformed to 50 ?? 50 pixels A

sample in black and white. There were 1,200 images in each class, considering the left-handed signers, these were then flipped horizontally.

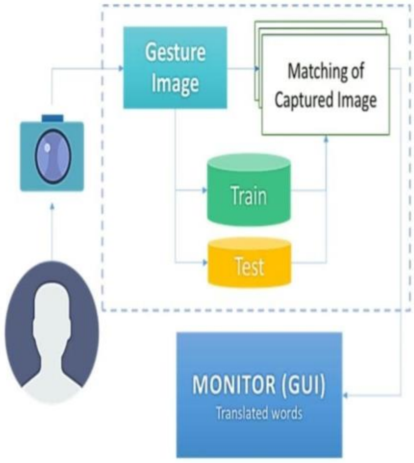
- Collection of training data, augmentation of images and cropping procedures- the processing of datasets for static SLR was carried out using Python to continuously capture images. Pictures were automatically cropped and transformed to a black and white sample of 50 X 50 pixels. Every class contained 1,200 images that, given the left-handed signers, were then flipped horizontally.

- **Pre-Processing Image**:

- These images were preprocessed with augmentation techniques in the pre-processing stage by adding pixels and colors, top and black hat, morphological transformation, blurring, saturation,
- and sharpening to enhance the variations in the datasets.
- The arbitrary size of the image in the datasets cannot be fed into the convolution neural network immediately.
- At the moment of reading images, the images were rescaled to 64 64 3 pixels by the deep learning library called TensorFlow during the model training.
- During the preparation, image augmentations such as width and height change, rotation and shearing, and horizontal flip were also performed, which were not part of the previous image augmentation.

- **Recognizing pictures:**

- Using TensorFlow as the backend, OpenCV to read video frames, Visual Studio Code, and Python as the editor and programming language respectively, the qualified model was loaded onto a laptop.
- OpenCV captures hand-shaped video frames from the signer in realtime and resizes them to 64 64 3 pixels.
- The model detects and forecasts sign digits successfully. camera, the machine classifies the



test sample and compares it with the gestures stored in thedictionary, and the corresponding output is shown to the user on the screen.

Fig 1: Conceptual framework diagram

Fig 2. Recognizing gestures

IV. CONCLUSIONS

In order to familiarize users with the basics of sign language, the main purpose of the project was to create a framework that could translate static sign language into its corresponding word counterpart that includes letters, numbers, and simple static signs. An evaluation plan was designed by the researchers and a series of experiments were carried out to ensure the importance of the functionality of the system intended for non-signers. Testing results have been impressive in terms of the system's usability and learning effect. This research was carried out with adequate assistance and consultation from specialists in sign language. Recognition of sign language is a difficult issue if we consider all the potential variations of signs that a device of this type requires to comprehend and interpret. That being said, the best way to solve this problem is possibly to break it into simpler issues, and a potential solution to one of them will lead to the system described here. The system did not perform very well, but it has been shown that only cameras and convolution neural networks can be used to construct a first-person sign language translation system. It was noted that the model appears to confuse many signs, such as U and W, with each other. But thinking a bit about it, because using an orthography corrector or a word predictor will improve the accuracy of the translation, maybe it doesn't need to have a perfect score.

The next step is to evaluate the solution and discuss ways the method can be improved. By gathering more quality data, attempting more convolution neural network architectures, or redesigning the vision system, some changes could be carried. In order to familiarize users with the basics of sign language, the main purpose of the project was to create a framework that could translate static sign language into its corresponding word counterpart that includes letters, numbers, and simple static signs. An evaluation plan was designed by the researchers and a series of experiments were carried out to ensure the importance of the functionality of the system intended for non-signers. Testing results have been impressive in terms of the system's usability and learning effect. This research was carried out with adequate assistance and consultation from specialists in sign language.

REFERENCES

[1] Lean Karlo S. Tolentino, Ronnie O. Serfa Juan, August C. Thio-ac, Maria Abigail B. Pamahoy, Joni Rose R. Forteza, and Xavier Jet O. Garcia, *"Static Sign Language Recognition Using Deep Learning"*, International Journal of Machine Learning and Computing, Vol. 9, No. 6, December 2019.

[2] Ming Jin Cheok Zaid Omar1 · Mohamed Hisham Jaward2 , *"A review of hand gesture and sign language recognition techniques"*, Int. J. Mach. Learn. & Cyber.

[3] L. G. Zhang, Y. Chen, G. Fang, X. Chen, and W. Gao, *"A vision-based sign language recognition system using tiedmixture density HMM,"* in Proc. the 6th International Conference on Multimodal Interfaces, 2004, pp. 198204.

[4] Q. Wang, X. Chen, L. G. Zhang, C. Wang, and W. Gao, *"Viewpoint invariant sign language recognition,"* Computer Vision and Image Understanding, vol. 108, no. 1-2, pp. 87–97, 2007.

[5] T. Chouhan, A. Panse, A. K. Voona, and S. M. Sameer, *"Smart glove with gesture recognition ability for the hearing and speech impaired,"* in Proc. 2014 IEEE Global Humanitarian Technology Conference-South Asia Satellite (GHTC-SAS), 2014,pp.105110.

[6] M. Süzgün, H. Özdemir, N. Camgöz, A. Kındıroğlu, D. Başaran, C. Togay, and L. Akarun, *"HospiSign: An interactive sign language platform for hearing impaired,"* Journal of Naval Sciences and Engineering, vol. 11, no. 3, pp. 75-92,2015.

[7] J. A. Deja, P. Arceo, D. G. David, P. Lawrence, and R. C. Roque, *"MyoSL: A Framework for measuring usability of two-arm gestural electromyography for sign language,"* in Proc. International Conference on Universal Access in Human-Computer Interaction, 2018, pp. 146-159