

A Robust Sign Language and Hand Gesture Recognition System Using Convolution Neural Network

K.Chandra Sekhar¹, Prakhya D², NSVK Reddy³, A Varaprasadh⁴, M Sri Manjari⁵

¹Asst.Professor, Dept of Computer Science Engineering, ANITS, Andhra Pradesh, India

²Student, Dept of Computer Science Engineering, ANITS, Andhra Pradesh, India

³Student, Dept of Computer Science Engineering, ANITS, Andhra Pradesh, India

⁴Student, Dept of Computer Science Engineering, ANITS, Andhra Pradesh, India

⁵Student, Dept of Computer Science Engineering, ANITS, Andhra Pradesh, India

Abstract - Sign Language is the only way of communication for the people who are not able to speak and hear anything. It is a boon for the physically challenged people to express their thoughts and emotions. But not everyone can understand their gestured language. Hence forming a communication gap. In this work, a novel scheme of sign language recognition has been proposed for identifying the alphabets, numbers and hand gestures in sign language. This serves as the bridge between verbally impaired people and normal ones. With the help of computer vision and neural networks we made an attempt to detect the signs and give the appropriate text as output, which can be read and understood by the normal people.

Key Words: Sign Language Recognition, Convolution Neural Network, Image Processing, Edge Detection, Hand Gesture Recognition.

1. INTRODUCTION

We capture the stream of frames from the hardware of the system and each frame goes through the prescribed process. The hand position in each frame can be obtained by various methods. Thus obtained image goes through various layers of neural network to obtain the most appropriate result as output. The system is properly trained using a large data set to make itself more robust.

1.1 Image Processing

Image processing is the process of performing some operations on an image to get an enhanced image or to extract some useful information from it. It is a type of signal processing in which input is an image and output may be image or characteristics/features of that image. Nowadays, image processing is one of the rapidly growing technologies.

Image processing basically includes the following 3 steps:

- Capturing the image through image acquisition tools.

- Performing various operations to extract or manipulate the image.
- Give the output which is either an alternative image or the report of analyzed results of the image.

1.2 Sign Language

It is a language that includes gestures made with the hands and other body parts, including facial expressions and postures of the body. It is used primarily by people who are deaf and dumb. There are many different sign languages as, British, Indian and American sign languages. British sign language (BSL) is not easily intelligible to users of American sign language (ASL) and vice versa.

A functioning signing recognition system could provide a chance for the inattentive to communicate with non-signing people without the necessity for an interpreter. It might be used to generate speech or text making the deaf more independent. Unfortunately there has not been any system with these capabilities thus far. During this project our aim is to develop a system which may classify signing accurately.

2. RELATED WORK

2.1 A Survey of Hand Gesture Recognition Methods in Sign Recognition.

Sign Language is that the sole technique utilized in conversation between the hearing-impaired community and normal community.

Signing Recognition gadget, that's required to renowned signing, has been extensively studied for years. The research are supported numerous sensors, gesture segmentation, extraction of capabilities and classification techniques. This paper aims to analyze and evaluate the techniques employed in the SLR systems, classification strategies that are used, and suggests the foremost promising technique for future research. Way to latest development in classification strategies, most of the current proposed works in particular contribute to the classification strategies, like hybrid method and Deep

Learning. This paper specializes in the classification methods used in earlier signing Recognition device. Supported our review, HMM-based approaches are explored drastically in earlier research, including its modifications. Sign Language is that the only technique applied in community among the hearing-impaired network and customary network. Signing Recognition (SLR) gadget, that's required to acknowledge sign languages, has been broadly studied for years. The studies are supported various input sensors, gesture segmentation, extraction of capabilities and classification techniques. This paper aims to investigate and examine the methods employed inside the SLR systems, classifications techniques which might be used, and suggests the most promising technique for future research. Thanks to current advancement in classification strategies, many of the current proposed works particularly contribute on the classification strategies, like hybrid method and Deep Learning. This paper specializes in the classification methods used in previous signing Recognition machine. Supported our review, HMM-primarily based methods are explored appreciably in previous research, including its modifications. Sign Language is that the sole technique applied in verbal exchange between the hearing-impaired community and standard community. Signing Recognition (SLR) system, that's required to renowned sign languages, has been widely studied for years. The research are supported various input sensors, gesture segmentation, extraction of functions and classification techniques. This paper aims to research and evaluate the techniques employed in the SLR systems, classifications techniques which can be used, and suggests the most promising approach for destiny research. thanks to current advancement in classification strategies, a number of the current proposed works mainly contribute at the classification strategies, like hybrid approach and Deep Learning. This paper specializes in the classification techniques applied in prior signing Recognition gadget. Supported our review, HMM-based methods are explored substantially in prior research, consisting of its modifications. Sign Language is that the sole technique applied in communicate between the hearing-impaired network and commonplace community. Signing Recognition (SLR) system, that's required to acknowledge signal languages, has been broadly studied for years. The studies are supported numerous input sensors, gesture segmentation, extraction of functions and classification techniques. This paper aims to research and compare the strategies employed in the SLR systems, classifications techniques that are used, and suggests the most promising approach for future research. way to recent advancement in classification techniques, most of the latest proposed works specially contribute at the classification techniques, like hybrid technique and Deep Learning. This paper makes a speciality of the classification methods used in earlier signing Recognition gadget. Supported our review, HMM-based totally

methods are explored drastically in previous research, along with its modifications. This observation is primarily based on various input sensors, gesture segmentation, extraction of features and classification strategies. This paper aims to research and examine the strategies employed within the SLR systems, classifications methods which are used, and suggests the most reliable approach for future research. One way to current advancement in classification strategies, a few of the these days proposed works mainly contribute to the classification strategies, like hybrid method and Deep Learning. Supported our review, HMM-primarily based approaches are explored extensively in earlier research, together with its modifications. Hybrid CNN-HMM and completely Deep Learning approaches have proven promising effects and offer possibilities for in addition exploration. Hybrid CNN-HMM and absolutely Deep Learning strategies have proven promising consequences and offer possibilities for further exploration .Hybrid CNN-HMM and absolutely Deep Learning approaches have shown promising consequences and offer possibilities for similarly exploration

2.2 Communication among Deaf-Dumb People and Normal People

Chat packages became a sturdy media that assist humans to talk in several languages with each other. There are many chat applications which can be used different people in numerous languages but there are not such an interview utility that has facilitate to speak with signal languages. The developed device is primarily based on Sinhala signing. The system has included 4 main components as textual content messages are converted to sign messages, voice messages are transformed to sign messages, sign messages are transformed to textual content messages and signal messages are converted to voice messages. Google voice reputation API has used to expand speech man or woman popularity for voice messages. The device has been trained for the speech and text patterns by using the usage of a few text parameters and symptoms of Sinhala signing is displayed through emoji. Those emoji and signs which can be included during this method will bring the regular humans more on the point of the disabled human beings. that is regularly frequently often frequently a 2 way communicate device however it uses pattern of gesture popularity which isn't always reliable in getting suitable output.

2.3 A System for Recognition of Indian signing for Deaf People the use of Otsu's Algorithm

Trough this paper we proposed a few techniques, through which the popularity of the signs and symptoms will become clean for peoples while verbal exchange. . And therefore the effects of those symbols symptoms are becoming to be converted into the text. at some stage in this project, we're taking pictures hand gestures via webcam and convert this photo into grey scale picture.

The segmentation of grey scale picture of a hand gesture is performed using Otsu thresholding algorithm.. Total photograph stage is split into two lessons one is hand and other is history. The highest quality threshold cost is about via computing the ratio between class variance and total magnificence variance. to hunt out the boundary of hand gesture in image Canny facet detection approach is used. In Canny part detection we used aspect primarily based segmentation and threshold based segmentation. Then Otsu’s set of rules is employed due to its simple calculation and stability. This set of rules fails, when the global distribution of the goal and background vary broadly.

3. ARCHITECTURE

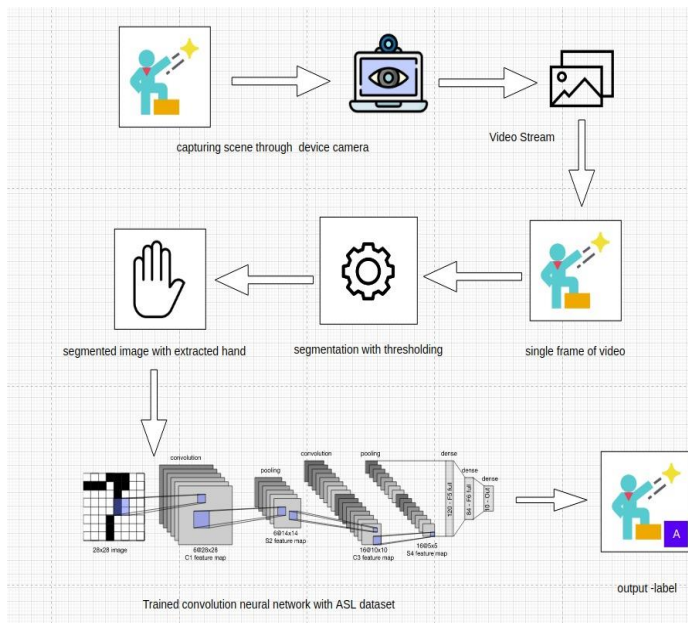


Fig-1: Architecture of the system

4. SEGMENTATION

Segmentation is the system of partitioning a digital image into couple of segments. The aim of segmentation is to simplify and/or exchange the illustration of an photograph into something this is more significant and less difficult to analyze. Image segmentation is commonly used to locate objects and boundaries in images. More precisely, photograph segmentation is the procedure of assigning a label to every pixel in an photo such that pixel is with the identical label share positive characteristics.

The result of photo segmentation is a set of segments that collectively cowl the entire picture, or a fixed of contours extracted from the image. Each of the pixels in a area is comparable with admire to some characteristic or computed property, which includes color, intensity, or texture. Adjacent regions are significantly different based on the same characteristic. Modern image segmentation techniques are powered by deep learning technology. In our project we are using skin

colour values to get the pixels that belong to face and hand. Then we use Harcascode classifier to identify the face and block it using white square . Then we take the hand image cluster and apply preprocessing techniques.

5. PRE-PROCESSING

There are 3 important steps in image preprocessing. They are:

5.1 Uniform aspect ratio:

An aspect ratio is a proportional relationship between an image's width and height. Essentially, it describes an image's shape. Since images don't got to have an equivalent dimensions to possess the same aspect ratio, it's better to crop them to a selected ratio than to undertake to match their exact dimensions.

5.2 Image Scaling:

Once we've ensured that each one images are square (or have some predetermined aspect ratio), it's time to scale each image appropriately. We've decided to possess images with width and height of 100 pixels. We'll got to scale the width and height of every image by an element of 0.4 (100/250). There are a good sort of up-scaling and down-scaling techniques and that we usually use a library function to try this for us.

5.3 Data augmentation:

Another common pre-processing technique involves augmenting the prevailing data-set with perturbed versions of the prevailing images. Scaling, rotations and other affine transformations are typical. This is often done to show the neural network to a good sort of variations. This makes it less likely that the neural network recognizes unwanted characteristics within the data-set.

This step of preprocessing the image is completed to bring the obtained image to appropriate size. Since we are training the system with images of certain height and width it's important for us to resize the obtained image . This makes it easier for us to match the pictures and find the match.

6. CONVOLUTION NEURAL NETWORK

CNN is a special architecture of ANN, proposed by Yann Le Cun in 1988. Convolution Neural Network uses some features of the visual cortex. One of the most popular architecture image classification is one of the most popular architecture. For example Amazon for generating product recommendations and Google for search through among users' photos.

The image is seen as an array of pixels by the computer. For example, if image size is 400 x 400. In this case, the size of the array will be 400x400x4. Where 400 is width, next 400 is height and 4 is RGB channel values. The computer is assigned a value from 0 to 255 to each of

these numbers. This value describes the intensity of the pixel at each and every point.

To solve this problem the computer looks for the characteristics of the base level. In human understanding such characteristics are for example are the trunk or large ears. For the computer, these characteristics are the boundaries or curvatures. Then, it is passed through the groups of convolutional layers the computer constructs more abstract concepts. In more detail: the image is passed through a series of convolutional layers (nonlinear, pooling, fully, connected layers).

RESULTS

Our system worked well to provide appropriate results. We trained our system using the American Sign Language gestures. The samples of the obtained results are shown below.



Fig-2: Result of Alphabet A



Fig-3: Result of Alphabet W

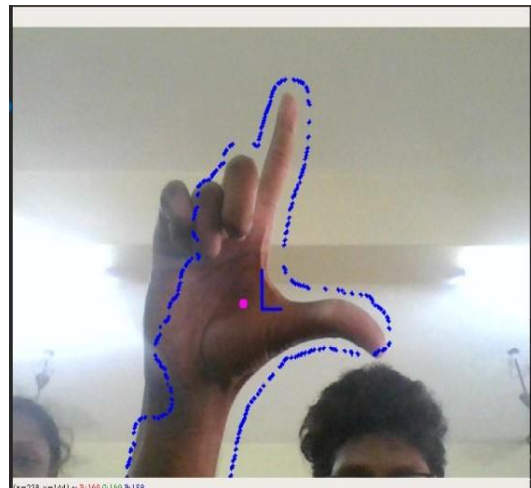


Fig-4: Result of Alphabet L

CONCLUSION

We obtained letters as output of the signs . The project can be further extended such that we conjugate all the letters to obtain statements. Thus making the output more readable.

Nowadays, applications need several kinds of images as sources of information for elucidation and analysis. Several features are to be extracted so as to perform various applications. The aim of this project is to create an offline sign language recognition system that is more robust in order to get the most appropriate output. In our work we trained the system with a large data set of images. Then we used cluster segmentation with respect to colour and edges. Image then undergoes feature extraction using various methods to make the image more readable by the computer. The closest match from the database is the obtained. Its label is displayed as output. Sign language recognition system is a powerful tool to prepare an expert knowledge, edge detect and the combination of inaccurate information from different sources. The intend of convolution neural network is to get the appropriate classification and accurate output.

REFERENCES

- [1] <https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148>
- [2] <https://adeshpande3.github.io/adeshpande3.github.io/A-Beginner's-Guide-To-Understanding-Convolutional-Neural-Networks/>
- [3] <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.734.8389&rep=rep1&type=pdf>