Sign language to Text conversion

Bachelors of Technology
in
Information Technology

Submitted by
Amit Singh (IIT2019045)
Abhishek Raj (IIT2019058)
Sanidhya Gupta (IIT2019074)
Aastha Singh (IIT2019078)
Ritik Kumar (IIT2019088)

Under the supervision of
Dr. Anjali Gautam

Abstract

Sign language is one of the oldest forms of language for communication and it has been used since the origin of life but many people are not aware about sign language, that's why interpreters are very difficult. To handle it, we have built a CNN based model to convert the sign language to text form which is based on american sign language. We used the method in which the hand is first scanned through a filter and after the filter is applied it manipulates the only important feature like outer edge of our hand and ignore other property like color of hand and then the hand is passed through a classifier, after that on the basis of its class it predicts the hand gestures . In our project we have achieved 98.3% of accuracy for the 26 letters of the alphabet.

Contents

1. Introduction

2. Problem Statement

3. Literature Survey

4. Datasets

5. Prerequisites/Learning

6. Implementation

7. Results

8 . Future Scope

9. Conclusion

10. References

1. Introduction

Communication with hearing impaired people has become an issue in our nation. Proper communication is necessary for building a good and developed nation. From the world total population there are about 466 million people who have impaired hearing loss (approximately 5% of world population) out of 466million people 34 million of those are under 18 according to WHO. It is predicted that it would surpass 900 million by 2050. On the basis of inputs sign language recognition is categorized into 2 parts: data gloves-based and vision-based.In this report we will solve this problem by using vision-based.In last 7-8 years, vision based techniques have become more popular where we take input from cameras. To make hand detection easier Marcos and Sandjaja used color-coded gloves . Hybrid Architecture is a combination of both architectures .It is less constraining and more affordable than data gloves.It takes high computing power consumption and lower accuracy. When it comes to the angle of the hand ASL is also strict.If u try to detect to detect U it give result like D,T. To communicate and share ideas to other people Impaired people make use of their hands and our project helps in converting hand gestures to text and then text to voice message.

2. Problem Statement

Our group has decided to develop a project in which impaired people get more involved .Our Problem statements revolve around the idea of vision based sign language recognition that is used for impaired people for generating text using hand gestures . Our objective is to make it simple and intuitive.

3. Literature Survey

3.1 A breakthrough in translating speech into sign language using 3D Avatar Animator 2020 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)

This work is proposed by Thangarajah Akilan [16] in 2020. Here they have developed a robotics system avatar which converts the English Speech to Indian Sign Language. Here the input is the voice signal , which is then converted into ISL symbols in the output which is shown by the robot's hand gestures. Here first google cloud speech recognizer is used in the stage 1. For the conversion of english language to ISL language , all the grammar rules must be strictly followed which is already stored in the database. In the stage 1 , input speech is converted into text using google cloud speech API. Now these long sentences are splited into text words using semantic parsers ML and NLP tools. Now these parsed blocks are passed

into ISL grammar rule based parser, which then converts into ISL tokens. Now in stage 2 data preparation is done.After this the tokens are passed into ISL generator , which converts it into ISL words. Now , to generate its symbols (shown by the avatar), these are then converted into Hamburg Sign Language notation. These notations are used globally in all languages.The output of thai is then passed to the Avatar. Below is the output generated. Average accuracy obtained is 77% with a delay in response time of 0.87 sec.

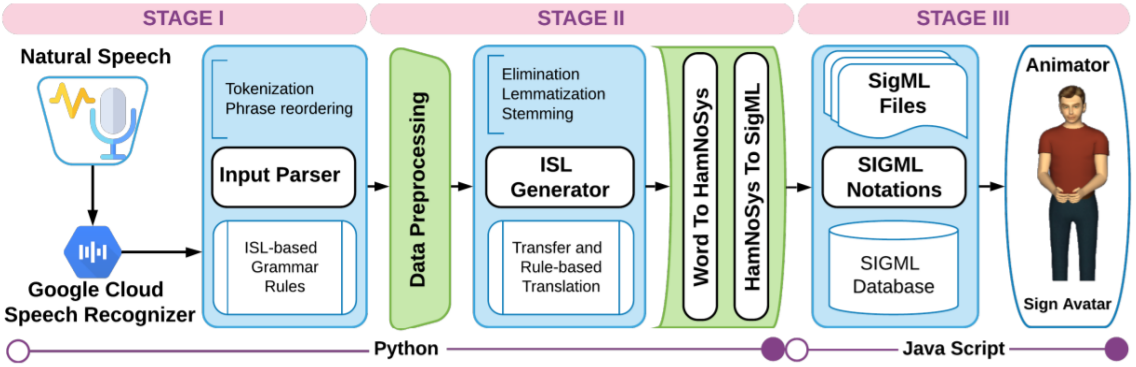
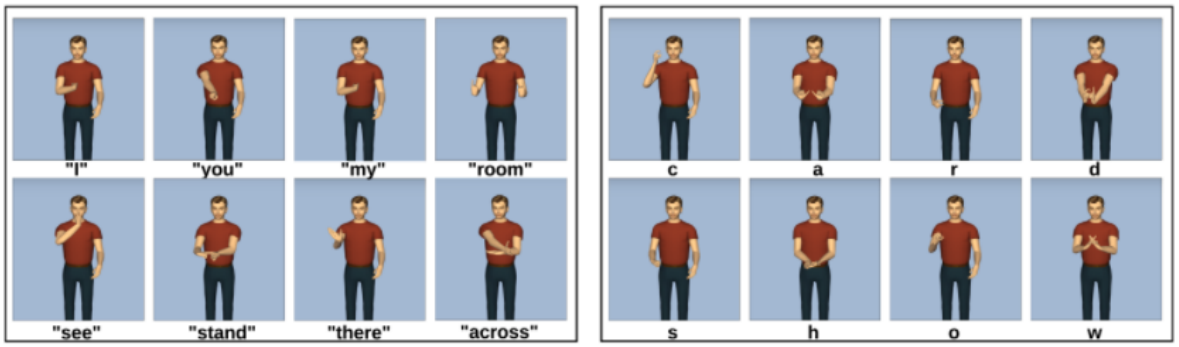


Fig. 1: Proposed ES2ISL. A Tri-stage Operational Flow: Speech to Text Conversion, Data Preprocessing, and Word to SIGML.



(a) Word-level parsing (b) Character-level parsing

3.2 Conversion of Sign Language to Text and Speech (2020) The Second International Conference on Inventive Research in Computing Applications (ICIRCA-2020) IEEE Xplore Part Number: CFP20N67-ART; ISBN: 978-1-7281-5374-2

This paper that was published in 2020 uses SVM algorithm to develop a real time American Sign Language to text converter software system. In this proposed work, 26 alphabets of american sign language are used with an additional symbol for “blank” is considered which is then captured by the smartphone in real time environment. In this project , One Plus smartphone having Android oreo version is used. Java -based OpenCV wrapper is used to develop the algorithm. Input images are of size 200*200 pixels in RGB format.

Support Vector Machine, a machine learning algorithm uses HOG descriptors as

characteristics of the image.SVM algo is used to train this model.The dataset used in this project is that of ASL which is publicly available on kaggle, consisting of 3000 input images of english alphabets.

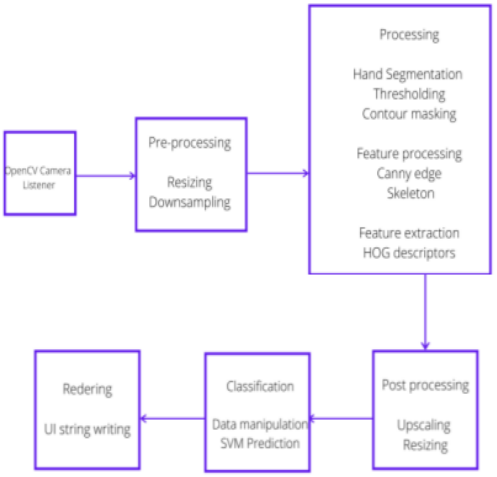


Figure 2: Architecture of the android application

Table 1· Performance comparison of all array parameters in SVM

Sl. No	Different combination of methods in SVM				
	Detection Method	***Kernel***	***Dimensionality Reduction***	***Average per Image processing time (MilliSecond)***	***Accuracy***
1	Contour Masking	Linear	None	18.2	97.45
2	Contour Masking	Linear	PCA	18.0	97.98
3	Contour Masking	RBF	None	18.4	98.12
4	Contour Masking	RBF	PCA	17.8	98.34
5	Skeleton	Linear	None	17.9	98.22
6	Skeleton	Linear	PCA	17.7	98.25
7	Skeleton	RBF	None	18.4	98.56
8	Skeleton	RBF	PCA	18.0	98.89
9	Canny Edges	Linear	None	18.1	98.52
10	Canny Edges	Linear	PCA	17.5	98.67
11	Canny Edges	RBF	None	18.3	98.74
12	Canny Edges	RBF	PCA	15.0	98.82

3.3 An Efficient Approach for Interpretation of Indian Sign Language using Machine Learning 2021 3rd International Conference on Signal Processing and Communication (ICPSC) | 13 – 14 May 2021 | Coimbatore

The work proposed in this research paper is to apply many types of machine learning algorithms like Support Vector Machine, Convolutional Neural Network and Recurrent Neural Network and convert hand gestures into alphabets and classify them, then compare which algorithm among them is efficient to solve this problem. They used a dataset which had ISL hand gestures of English language alphabets(i.e A-Z), and also 7 English language words (STAND, OK, SWING, SMILE , HOLIDAY, FRIEND,and BOAT). So this dataset had different classes of images i.e. 42 (9 + 26 + 7). They captured 1200 different types of images for each class for this dataset. To remove noise and to mask unwanted areas in the image, pre-processing action was required for these images. They passed the input images into K-means clustering. Then they passed these images through the Bag of Visual words classifiers and finally into the SVM classifier. As described earlier in the dataset, there were 42 different classes of image so k=42 for the K Means clustering classifier. After applying the k-means algorithm to these images the visual words were accumulated for the train and test dataset. The SVM Model training used 40320 images out of the total 50391 in the dataset. To test the work of the classifier model they used the remaining images i.e. 10071 images. After testing, around 99.5% accuracy was achieved. Precision score, recall score, F1 score, and other performance metrics like these were also calculated. They modeled and

created a Convolutional Neural Network(CNN) using the Keras library in Python. The classifier model used 30,240 images for training which is 60% of the total images of the dataset. Several epochs were used to train the classifier. After testing around 88.89% maximum average accuracy was obtained. Using the Keras library Python, a Recurrent Neural Network was created . The classifier model was trained using around 30,240 images. A Different number of epochs were used to train the classifier. Around 82.3% maximum overall testing accuracy was obtained.

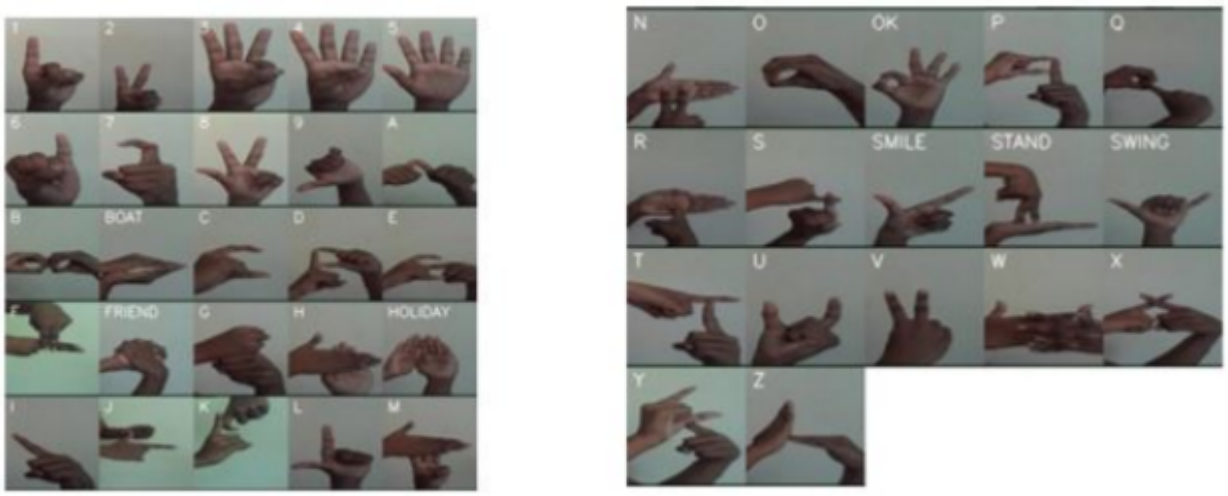


Fig. 1. ISL Hand Gestures

3.4 Tensor training decomposition application in S2VT model for sign language recognition (2021) National Engineering Laboratory for Speech and Language Information Processing, Institute of Statistical Signal Processing, University of Science and Technology of China, Hefei 230027, China

Now, Sequence-to-sequence video to text (S2VT) models are being used widely in the field of our project. However, it consumes a large amount of memory and computer resources as a result So this brings S2VT models(more than 20,000,000 different trained parameters) at a disadvantage of using them in mobile phones and similar devices. To solve this problem, in S2VT models to reduce the parameters using tensor-train have been proposed in this paper. Firstly, they checked the results of this decomposition using tensor-train on the working of the model. Then, for Chinese sign language recognition, they built 6 tensor-train S2VT models; i.e. tensor-train decomposition in different layers of a S2VT model was applied.

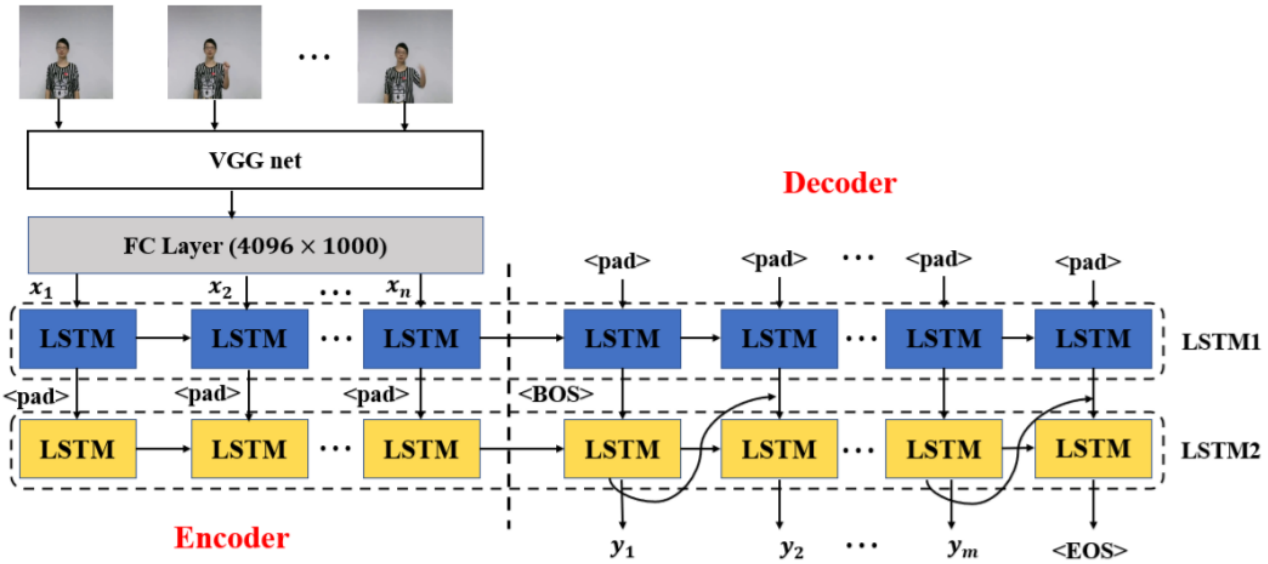


FIGURE 1. **The framework of S2VT, which has two LSTMs. The first LSTM layer (colored blue) models visual features inputs of sign language videos. And the second LSTM layer (colored yellow) models the language given the text input and the hidden representation of the video sequence.**

The image is from a public dataset-USTC-SLR [2]. 50 short phrases were chosen, each of which contains 6-8 words that we frequently use in our lives. Each sentence has 100 samples as 20 signers play each word 5 times and there were 5,000 samples in the dataset.These samples were divided into 2 parts, one was used for training and other was used for testing. In the training part, 16 signers from the 20 signers were randomly chosen from the dataset. The remaining of the data set was used for testing.

TABLE 3. Experimental Results on USTC-SLR Dataset. We report i) the number of parameters, ii) the memory of models, iii) the decrease of parameters iv) the accuracy.

	Parameters	Memory	Decrease	Accuracy
Baseline[6]	24,393,144	97.6MB	-	98.4%
Scenario-1	20,298.440	81.2MB	16.8%	98.0%
Scenario-2	16,395,224	65.6MB	32.8%	98.2%
Scenario-3	12,395,464	49.6MB	49.2%	91.9%
Scenario-4	4,397,544	17.6MB	81.9%	83.9%
Scenario-5	**12,300,520**	**49.2MB**	**49.5%**	**95.6%**
Scenario-6	302,840	1.2MB	98.8%	29.3%

Proposals were made to decrease the parameters here. After checking different cons and pros of parameters on the model , they discovered that it works best when rank is 3. The output tensors are from large to small ,and the input tensors are vice-versa, and without increasing or spreading the distribution. For Chinese sign language recognition, factorization(tensor-train) was used in many layers of the S2VT model. 6 scenario outcomes displayed through experiments that parameters of these models can be decreased further.

3.5 Sign Languages to Speech Conversion Prototype using the SVM Classifier In TENCON 2Ol9-2Ol9 IEEE Region lO Conference (TENCON) (pp. 1803-1807). IEEE.

This work proposed an example of giving output as speech for the Sign Language gestures to help speech impaired people in this paper. This prototype has a part that consists of accelerometers, gyroscopes and flex sensors, embedded in it. These sensors capture the real time gestures. The data is collected from these sensors and sent to the PC via Bluetooth using an Arduino Nano microcontroller. The data is processed by Personal Computerand then it sends it to the Arduino and classifies the Sign Language gestures and uses a Machine Learning Algorithm to give the output by predicting the words related to it. For classification they use Support Vector Machine (SVM). This prototype is very useful and identifies both ISL and ASL. It helps mute people by giving speeches. By making gestures, the proposed system allows people who use it to speak some Indian languages and English also. They trained and tested 11 gestures in Indian sign language and 22 gestures in American sign language successfully. 100% accuracy is achieved for database of Indian sign language with 75% training data and 25% test data.

98.91% accuracy is achieved for database of American sign language with 75% training data and 25% test data.

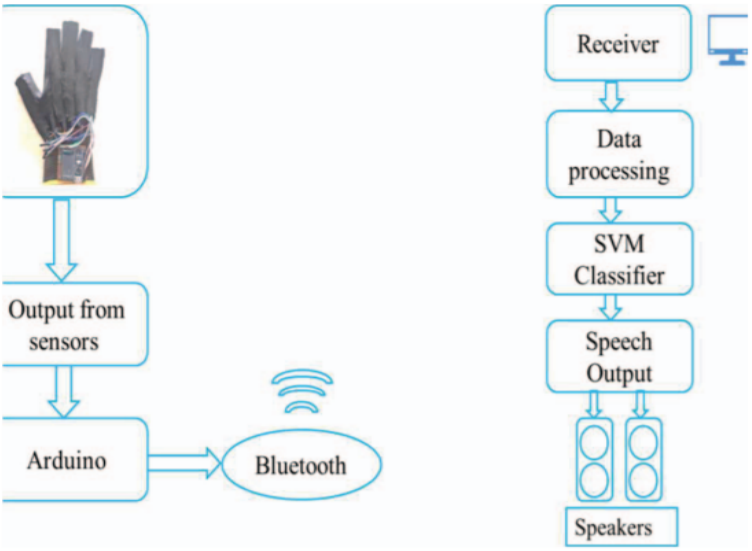


Fig. 1 Block diagram of Proposed System

3.6 Machine-learning-assisted stretchable sensor arrays Sign-to-speech translation(2020)

This project is made by Zhaou Zo.[4]. It consists of a wearable hand glove which consists of PCB, electrical wires , sensors , etc which is used to take electrical signals from the outside. These electrical signals are generated by the device which is worn in the hand. Signals frequency , length and all its properties highly depend on the orientation of the gloves.This device takes signals from the thumb, index finger, little finger, ring finger and the middle finger, then these are amplified. Then these amplified signals are then passed to low pass filters to reduce noise . Then these are converted for the feature extraction part. Now , a multi class SVM algorithm is applied on it to classify the dataset. During dataset generation time, a total of 660 ASL hand gestures were taken into consideration and were analysed by PCA for the feature extraction part. For the hand gesture recognition part a multi class SVM is then applied on this.

Reference No.	Year of publication	Dataset Used	Algorithm/ Notation Used	Accuracy Achieved	Classification
Abraham E. et.al. [7]	2019	40 samples each for 26 ISL Gestures containing readings of accelerometer, gyroscope and flex sensors	LSTM (Long Short-Term Memory)	Good Accuracy of 98% using LSTM, low cost hardware	Gesture based, Sensor Based, Sign-to-Speech
Chandra M.M. et.al. [3]	2019	22 ASL(American) Gestures and 11 ISL(Indian) Gestures	Signal Processing, SVM (support vector machine)	Can recognize both ISL and ASL (98.3% approx)	Gesture Based, Sensor-based, Sign-to-Speech
Zhou Z. et.al. [4]	2020	11 ASL Symbols (C, A, B, I, L, Y, 3,2,1, 8, I Love You)	PCA (Principal Component Analysis), multi-class SVM algorithm	Effective wearable, machine learning assisted translator	Gesture Based, Sensor Based, Sign-to-Speech6
Thangarajah Akilan[16]	2020	ISL T2SLT consisting of 3268 english words	ML and NLP tools were used	Achieved avg accuracy of 77%	Sign to speech
Kohsheen Tiku[5]	2020	26 ASL alphabets, ASL Kaggle dataset	SVM algorithm used	Achieved avg accuracy of 97%	Finger spelling , sign to text
Krishnaa Hari K B [6]	2021	ISL dataset generated (26 symbols, 7 words,and 9 numbers 1-9 , total 42 classes)	SVM, CNN and RNN were compared	SVM- 97.5% testing accuracy CNN-88.9% RNN-82.3%	Finger spelling , sign to text
BIAO XU [17]	2021	Chinese sign language (USTC-SLR)	LSTM	Avg accuracy is 86%	Gesture Based, Sensor Based, Sign-to-text

4. Datasets

4.1 American Sign Language Kaggle Dataset

A publicly available data set which we used in our project is ASL Kaggle dataset.
Link : https://www.kaggle.com/grassknoted/asl-alphabet

- This dataset size is 1.12 GB
- It contains colored images of each different symbols
- There are two folders in it , the first one contains a training dataset and the second one contains testing data.
- Training Folder contains 30 folders for each alphabet along with del , space and nothing symbol.
- Each folder contains 3000 colored images.
- Total images in training folder :- 81000
- Testing folder contains 28 colored images of all alphabets along with nothing and blank symbol images in it.
- The dataset consists of multiple grayscale 28x28 image pixels.

4.2 Generation Of our Own Dataset

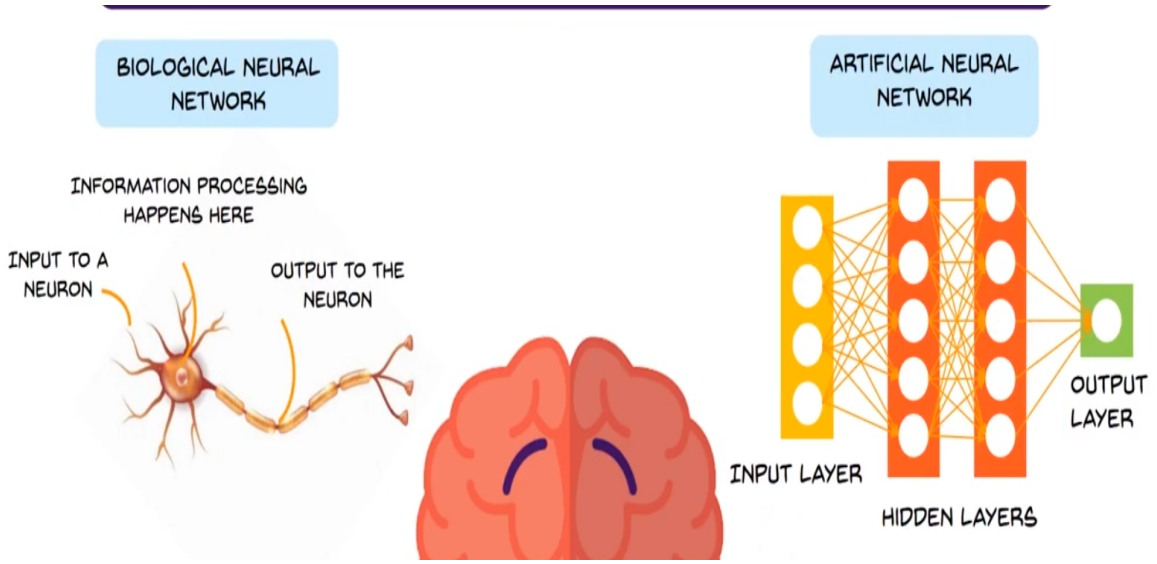
- We have captured different images for training and testing the model.
- For the training part we have captured images around 450-500 images of each character
- For the testing part we have captured images around 150-200 images of each character

- Total images for testing part :- 4268
- Total images for training part:- 12845
- 128*128 image pixels

5. Prerequisites/Learning

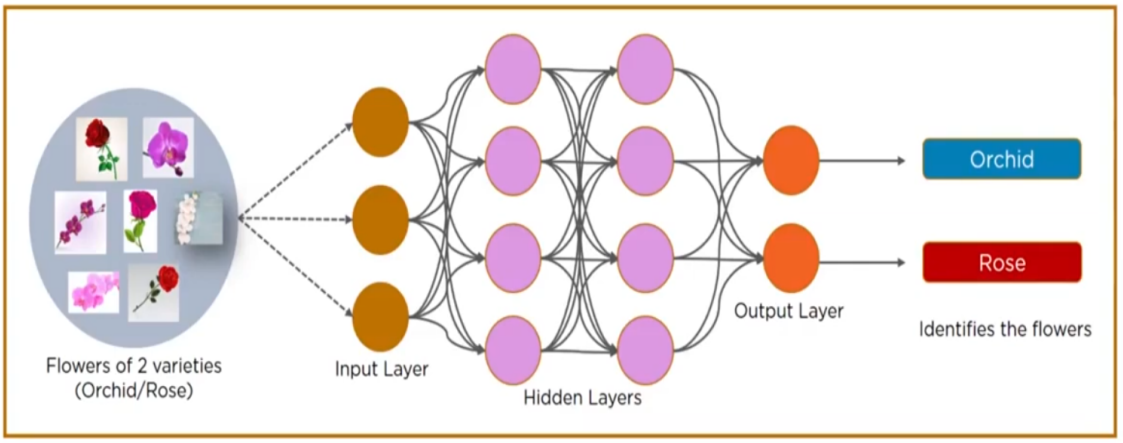
5.1 Artificial Neural Network :-

An artificial neural network is a system of hardware and software patterned that replicates the structure of the human brain.The first layer that processes input and after that it goes to next layer of neurons that is hidden layer where processing of information is done through multiple hidden layers and after that final information is passed to final output layer.

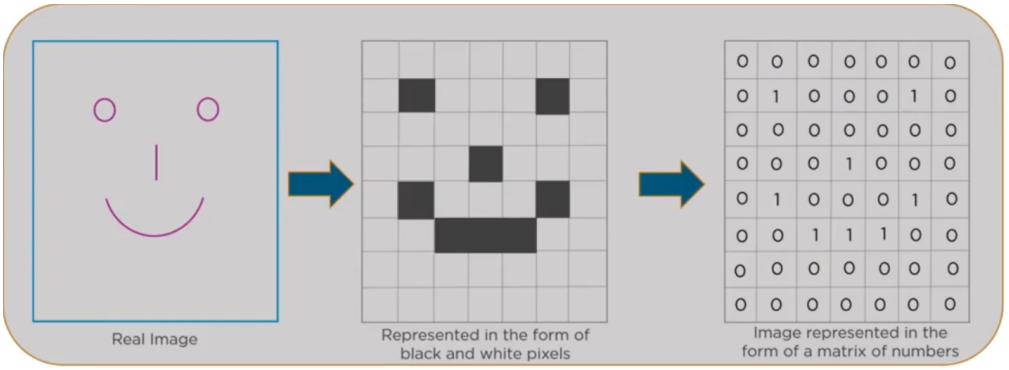


5.2 Convolution Neural Network(CNN)

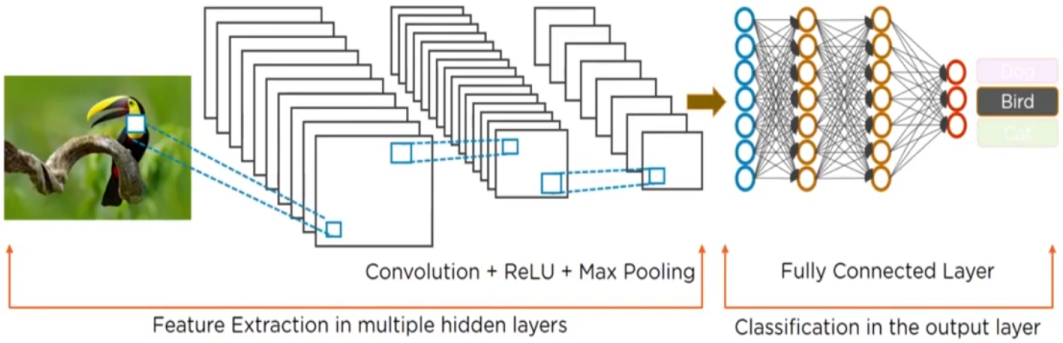
The CNN layers are arranged in 3 dimensions:,depth , height, and width.The neuron in a layer of CNN is attached to a small area of the layer(window size).Using CNN, the full image will be converted into a single vector of class scores.



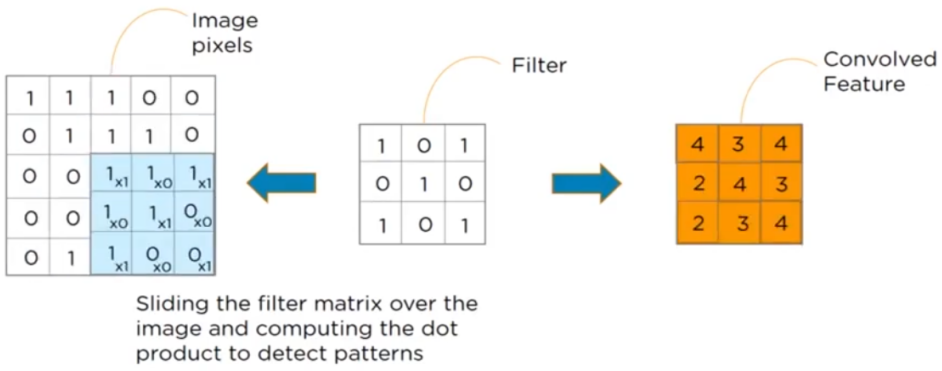
How CNN recognizes image



Fully connected Layer



1.Convolution Layer:- In convolution layer we use window size(small)(of 5*5) that increases the depth of input matrix. During each iteration we slide the window by stride size and filter entries in the form of dot product is used and input values as computation.After computation 2-D activation matrix is created which gives the results of that matrix at every spatial position. These networks get activated when they detect some kind of visual features such as the side of some orientation.



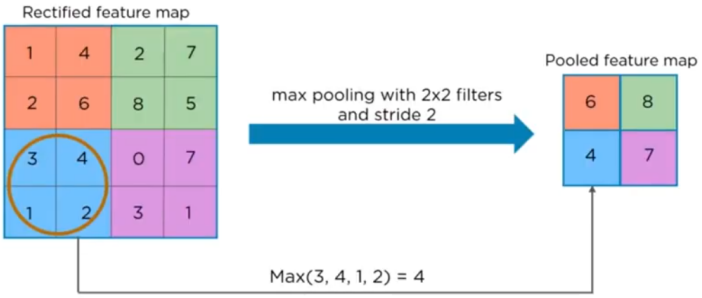
Simplilearn. All rights reserved.

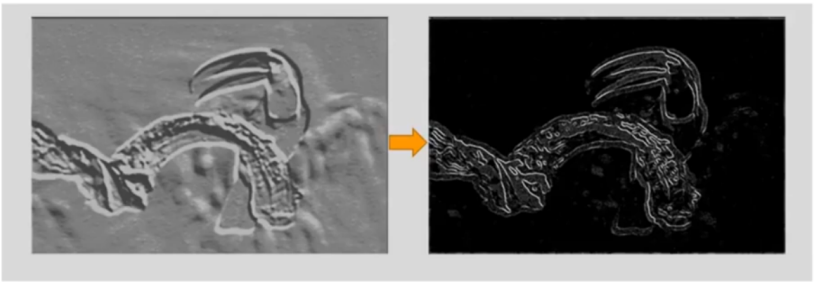
2.Relu Layer:-

It performs element wise operation and reset all negative pixels to zero and then introduces non-linearity to the network and finally the output is a rectified feature map

3.Pooling Layer:- Pooling layer processed rectified feature map.We use Pooling to reduce the dimension of the feature map . It has different filters that are used to identify different parts of the image like edges,corners,body etc.

we take the maximum of all values in **Max pooling** while we take the average of all values in **average pooling.**

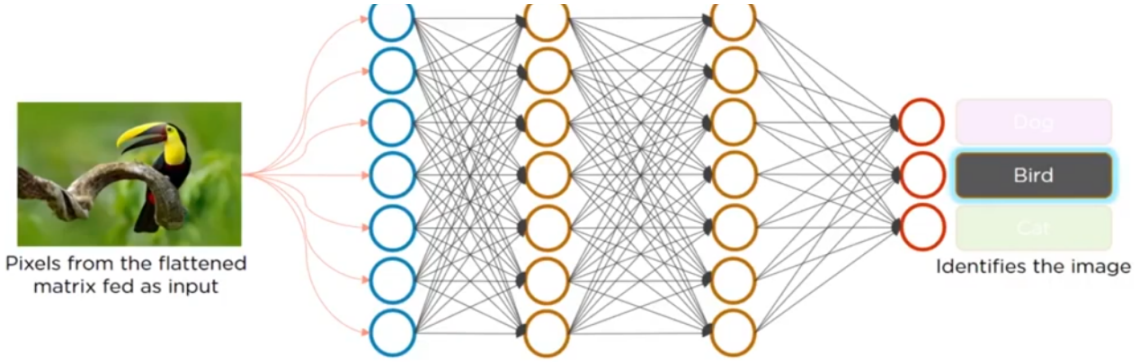




Identifies the edges, corners and other features of the bird

4.Flattening helps to change 2-D arrays that are pooled feature maps into a single long continuous linear vector .

5.Fully Connected Layer:- In convolution layer neurons are connected to a local region,while in a fully connected layer,well connect all the inputs to neuron.



5.3 Keras:-Keras is a deep learning API written in python that works as a wrapper to tensorflow.When we need a minimal line of code to build and test the neural network we used keras. It contains tools like layer,activation function,optimizer that make it easy to work with images.

5.4 Tensorflow:-It is a open source library for numeral computation.It is widely used in Machine Learning

5.5 OpenCV:- Open source computer Vision is an open source library of programming function used for real time computer vision. It is mainly used for image processing,video captures and analysis of feature like face recognition. It was written in c++.

6. Implementation & Methodology

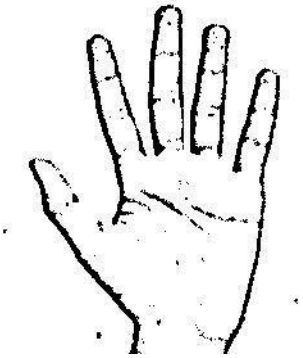
6.1 DATASET GENERATION

We tried to find already made datasets for the project but we could not get dataset in the form of raw images that fulfill our requirement. Therefore we created our own data set. Given below is steps for creating our own dataset.

In order to produce the dataset we used Open computer vision(Open CV).We capture 800 images of each and every symbol for training purposes and for testing purpose around 200 images also captured for each and every symbol.
By our webcam of the machine we capture each frame.We define region of interest in ach frame which is symbolized as blue bounded square as given figure below.

We extract our ROI which is RGB from this whole image and converted it into gray scale image as shown below.

At last we use gaussian blur filter to our image that helps in extracting various feature of our image.Given below figure is obtained by applying gaussian blur .



6.2 HAND GESTURE CLASSIFICATION

In Our Project we have used two layers of algorithms to predict the final result of the user's hand gesture.

Algorithm 1:

Step 1:- The image which we have taken by opencv, we have used gaussian blur and threshold there.
Step 2:- Now, the processed images is passed to our CNN model. If the frame of the detected letter is more than 50, then the letter will be printed on the dashboard.
Step 3:- If the blank symbol remains for (8-12) sec, then the space between the words will considered.

Algorithm 2:

- Many symbols show similar results like others.

  Like :- If we try to detect U it give result like D and T, similarly for R it shows D and U.

- We have made separate classifiers for that particular set.

Layer 1:

Model :

1st CNN Layer :

The hand's symbol picture is of 128 x 128 pixels. By using 32 filter weights(3 x 3 pixels each) we reduce its size to 126 x 126 pixels.

1st Pooling Layer :

For downsampled the images,we have used max pooling function and we have used 2x2 square matrix. Now, after this process we have got 63x63 pixels.

2nd CNN Layer :

Now, the images which we have got from the 1st CNN layer (63x63 pixels) will processed in the 2nd layer. By using 32 filter weights(3 x 3 pixels each) we reduce its size to 60 x 60 pixels.

2nd Pooling Layer :

Now again, we have used max pooling function for downsampled the images, we have used 2x2 square of matrix. Now, after this process we have got 30x30 pixels.

1st connected layer:

2nd CNN layer give the final image that is assed through 128 neuron of fully connected layer.And it is reshaped to 28800 arrays(30*30*32).To avoid overfitting we used 0.5 dropout layer.

2nd connected layer:

Then the output of the 1st connected layer will passed through 96 neurons of connected layer.

Final layer :

Now, the result which we have got from the 2nd connected layer will classify on the basis of there number of neurons (alphabets + blank symbol).

Pooling Layer :

This function is used to reduce computational cost and for reducing overfitting. Here we have used max pooling function of (2x2) matrix.

Activation function :

Here, RELU function we have used. It removes gradient which are erased and by reducing the computational cost it increases the process of training. RELU is used to calculate maximum between two parameters, like max(a,0).

Layer 2:

Many symbols show similar results like others. To handle this we have made two layers of algorithm.

So that we can predict or comes closer what we want to predict. In our project model we have found some symbols which are not giving proper result and they are predicting other symbols also:

- For S : M and N
- For D : R and U
- For I : T, D, K and I

So, due to this we have made three classifiers to handle it.

These are:-
(S,M,N),(D,R,U),(T,K,D,I)

Sentence Formation:

- When the count of predicted symbol reaches or exceeds a specific value i.e 50. And by threshold i.e 20, there is no other word is close to it, then we simply concatenate the string.
- Else, we just simply clear the record of current dictionary, to avoid the chances of predicting of wrong words.
- When the count of a plain background exceeds a specific value, then it will detect no space, if the current buffer is empty.
- Else, It will predicts the last words and with printing the space the word will append in the sentence.

Autocorrect :

For suggestion of the word, we have used a python library Hunspell suggest. We display three set of words which are are matching to the current word. From the suggested word user can easily select a word and add to the sentence. By this, we have reduces the chances of mistakes.

Training and Testing

Images(RGB), we convert it into grayscale, and for eliminating superfluous commotion we apply gaussian haze. We have resize our pictures to 128 x 128 by applying versatile threshold. The expectation layer appraises how logical the picture will fall under one of the classes. So with the assistance of softmax work the result is standardized somewhere in the range of 0 and 1 and to such an extent that the amount of each qualities in each class totals to 1. Initially the result of the expectation layer will be entirely unexpected from the genuine worth. In this way, for improving the model we have prepared the organization utilizing named information. The cross entropy is a consistent capacity which is positive at values which isn't same as marked worth and is zero precisely when it is equivalent to the named esteem. Subsequently we upgraded the cross-entropy by limiting it as near nothing. , It is utilized as execution estimation in the grouping. There is inbuilt capacity in tensorflow to ascertain the cross entropy.We have improved entropy work by utilizing angle plummet streamlining agent called Adam Optimizer.

7. Results

By using only 1 layer of algorithm we have achieved 94 % of accuracy. But after combining the layer 1 and layer 2 algorithm we have achieved 98.3%, which is better than the most of the current research paper on ASL. In research paper [5], they have applied SVM algorithm on ASL kaggle dataset which we have used for comparison , and they have achieved an average accuracy of 97%. In [6] they have applied different machine learning algorithms like SVM, CNN and RNN on Indian Sign language dataset generated by them. This dataset contains 42 classes (26 for alphabets, 9 for numbers 1-9 and 7 for english words) . They have compared each results obtained after applying each algorithm and finds the efficient among these three algorithms. Among them SVM has the highest training and testing accuracy (97%) followed by CNN (88.9%) and the by RNN (82.3%). In research paper [17] , they have applied LSTM algorithm on Chinese Sign Language Dataset, and achieved average accuracy of 86%. In [16] , they have applied ML and NLP algorithm and achieved average accuracy of 77% .In [4] research paper, 11 ASL gestures dataset were used . Principal component analysis (PCA) and multi-class SVM algorithm is applied to get better accuracy .

```
Epoch 1/5
1286/1286 [==============================] - 133s 103ms/step - loss: 1.1486 - accuracy: 0.5323 - val_loss: 0.3419 - val_accurac
y: 0.8954
Epoch 2/5
1286/1286 [==============================] - 200s 156ms/step - loss: 0.6787 - accuracy: 0.8318 - val_loss: 0.1186 - val_accurac
y: 0.9666
Epoch 3/5
1286/1286 [==============================] - 225s 175ms/step - loss: 0.3179 - accuracy: 0.9453 - val_loss: 0.0290 - val_accurac
y: 0.9774
Epoch 4/5
1286/1286 [==============================] - 217s 169ms/step - loss: 0.2110 - accuracy: 0.9548 - val_loss: 0.0216 - val_accurac
y: 0.9803
Epoch 5/5
1286/1286 [==============================] - 210s 163ms/step - loss: 0.1307 - accuracy: 0.9735 - val_loss: 0.0188 - val_accurac
y: 0.9835

17]: <tensorflow.python.keras.callbacks.History at 0x174d70b6190>
```

8. Future Scope

We will plan to accomplish higher exactness(accuracy) even in complex foundation by evaluating different foundation deduction algorithms.And additionally consider working on the preprocessing to anticipate signals in low light conditions with higher precision.

9. Conclusion

Our main project building goal was to develop a system capable of translating static sign language into its corresponding word equivalent which includes numbers, letters and basic static signs to familiarize users with the fundamentals of the sign language. In this report, a functional real time vision American Sign Language recognition for D&M people has been developed. Using the ASL dataset, we obtained an accuracy of 91.19%.
98.3% accuracy was obtained on our own data set. After implementing two layers of algorithms, prediction has been improved where we check and predict symbols that are more similar to each other.

10. References

[1] D. Singh, "Text to sign language," Thapar University, 2016. (3.1)

[2] J. Huang, W. Zhou, H. Li, and W. Li, "Sign language recognition using 3D convolutional neural networks," in Proc. IEEE Int. Conf. Multimedia Expo (ICME), Jun. 2015, pp. 1–6. (3.4)

[3] Chandra, M.M., Rajkumar, S. and Kumar, L.S., 2019, October. Sign Languages to Speech Conversion Prototype using the SVM Classifier. In *TENCON 2Ol9-2Ol9 IEEE Region lO Conference (TENCON)* (pp. 1803-1807). IEEE. (3.5)

[4] Zhou, Z., Chen, K., Li, X., Zhang, S., Wu, Y., Zhou, Y., Meng, K., Sun, C., He, Q., Fan, W. and Fan, E., 2020. Sign-to-speech translation using machine-learning-assisted stretchable sensor arrays. *Nature Electronics*, *3*(9), pp.571-578. (3.6)

[5] Kohsheen Tiku, Jayshree Maloo 2020. Real-time Conversion of Sign Language to Text and Speech Proceedings of the Second International Conference on Inventive Research in Computing Applications (CIRCA-2020) IEEE Xplore Part Number: CFP20N67-ART; ISBN: 978-1-7281-5374-2 (3.2)

[6] Krishnaa Hari K B , Divyashree S , 2021. An Efficient Approach for Interpretation of Indian Sign Language using Machine Learning 2021 3rd International Conference on Signal Processing and Communication (ICPSC) | 13 – 14 May 2021 | Coimbatore (3.3)

[7] Abraham, E., Nayak, A. and Iqbal, A., 2019, October. Real-Time Translation of Indian Sign Language using LSTM. In *2Ol9 Global Conference for Advancement in*

Technology (GCAT) **(pp. 1-5). IEEE. (table)**

[8] Fatima, S., Agarwal, A. and Gupta, P., 2016, March. Different approaches to convert speech into sign language. In *2Ol6 3rd International Conference on Computing for Sustainable Global Development (INDIACom)* (pp. 180-183). IEEE.

[9] Dhanjal, A.S. and Singh, W., 2019, February. Comparative Analysis of Sign Language Notation Systems for Indian Sign Language. In *2Ol9 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP)* (pp. 1-6). IEEE

[10] Verma, A. and Kaur, S., 2015. Indian sign language animation generation system for Gurumukhi script. *International Journal of Computer Science and Technology*, *6*(3), pp.117-121.

[11] Ghanem, S., Conly, C. and Athitsos, V., 2017, June. A survey on sign language recognition using smartphones. In *Proceedings of the lOth International Conference on PErvasive Technologies Related to Assistive Environments* (pp. 171-176).

[12] Vij, P. and Kumar, P., 2016, August. Mapping Hindi Text To Indian sign language with Extension Using Wordnet. In *Proceedings of the International Conference on Advances in Information Communication Technology & Computing* (pp. 1-5).

[13] Grif, M. and Manueva, Y., 2016, June. Semantic analyses of text to translate to Russian sign language. In *2Ol6 llth International Forum on Strategic Technology (IFOST)* (pp. 286-289). IEEE.

[14] Dutta, K.K. and GS, A.K., 2015, December. Double handed Indian Sign Language to speech and text. In *2Ol5 Third International Conference on Image Information Processing (ICIIP)* (pp. 374-377). IEEE.

[15] Bairagi, V.K., 2014. Gloves Based Hand gesture recognition using Indian Sign Language. *International Journal of Latest Trends in Engineering and Technology,* Vol.(8)Issue(4-1)(pp.131-137)

[16] Bhavinkumar Devendrabhai Patel, Harshit Balvantrai Patel, Manthan Ashok Khanvilkar, Nidhi Rajendrakumar Patel, and Thangarajah Akilan, *Member, IEEE* 2020 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE) (3.1 ref)

[17] BIAO XU , , SHILIANG HUANG , AND ZHONGFU YE 2021 , National Engineering Laboratory for Speech and Language Information Processing, Institute of Statistical Signal Processing, University of Science and Technology of China, Hefei 230027, China