CHAPTER 1. INTRODUCTION

Gesture-based communication is a remarkable kind of correspondence that regularly goes understudied. Deaf or speech-disabled people tend to communicate with others using gesture-based communication to convey their contemplations and emotions. Sign language seems to be a far more complicated interaction medium, which involves changing hands, faces and various body stances. Each nation has its own local sign language [1]. Different sign languages have their own standard of language structure, different gestures for different words, and different pronunciations. Hearing/Speech Impaired individuals are not self-reliant henceforth prepared gesture-based communication translators or consultants are required during clinical, legitimate arrangements and instructional meetings. Though it's good to have access to translators yet it very well can be seen as problematic for various reasons: first and foremost, Hearing/speech impaired individuals face quandaries in expressing themselves, so an interpreter needs to understand their gestures and represent them on their behalf and secondly, the interpreter who is more than the judge of phonetic content acts as a medium between the signers and the non-signers [2]. Translators are masters of depiction but they are not commonly and easily offered all the days, if offered conjointly, they charge tons. The vast majority of ordinary people don't possess any knowledge of sign language. Thus, numerous corresponding hindrances exist for hearing/speech impaired gesture-based communication users.

The number of deaf people has recently reached four hundred million [3], as per the data provided by World Health Organization. Following this cause, the latest experiments are being intensified to make it easier for disabled people to communicate [3]. While there are many sign languages, the most often used is the American Sign Language (ASL). American Sign Language, in contrast, has a wide variety of lingos, and is used in the U. S. and English conversing Canada. There are 22 hand-shapes relate to the 26 alphabets in order and one can sign the 10 digits on a single hand.

With the wish to accelerate the capacity of hearing/speech deprived people to receive assistance in their own society without the aid of a skilled consultant in sign language. These kinds of virtual machines can address the unmet need for professional decoding facilities and increase the quality of living considerably.

A comprehensive understanding of different systems, approaches and methodologies for the presentation of a sign language is the main objective of this research paper. The recent development situation in this field is further explored and helps to recognize prospects and obstacles for future research work.

A language barrier is formed between normal people and deaf and speech disabled people in the form of a sign language structure that differs from normal text. As a result, they rely on vision-based contact to communicate.

The signs can be readily interpreted by others if there is a standard GUI that translates sign language to text. As a result, testing has been conducted on a vision-based interface device that allows Deaf and Speech Disabled individuals to communicate without having to know one another's dialect.

The goal is to create a user-friendly human-computer interface (HCI) in which the device recognises human sign language. Across the globe there are numerous gesture oriented languages, among them are French Sign Language (FSL), Indian Sign Language, American Sign Language (ASL), British Sign Language (BSL) and Japanese Sign Language.

1.1 Keywords and Definitions

1.1.1 Linguistic Communication / Gesture-based Communication

Gesture based languages (also known as signed languages) are used for linguistic communication. In these languages, the optical-manual method is used to express the meaning. Manual articulations are used in conjunction with non-manual components to convey sign languages. Gesture based communication, with their own structure and vocabulary, are universal dialects in their very own way. While the symbolic languages are nearly identical, they are not standard and not commonly understandable. Both spoken and signed speech are considered forms of natural language by linguists, implying that they originated from an abstract, long-term ageing phase and developed over time without careful preparation. Body language, a form of nonverbal communication, should not be confused with sign language.

In deaf communities all over the world, sign languages have arisen as effective method of interaction and are at the core of local deaf societies. While sign languages are often used by listeners and individuals who need hearing aids, they are also used by listeners who are not capable of communicating, are having problems using vocalised communication because of a disability or disease, or have deaf families, such as kids of deaf elders.

The number of sign languages in the world today remains uncertain. Each country has its certain indigenous sign language and several countries have many sign languages. The 2020 edition of Ethnologue lists 144 sign languages, while the SIGN-HUB Atlas of Sign Language Structures lists over 200, with the caveat that there are likely to be more that have yet to be reported or found.

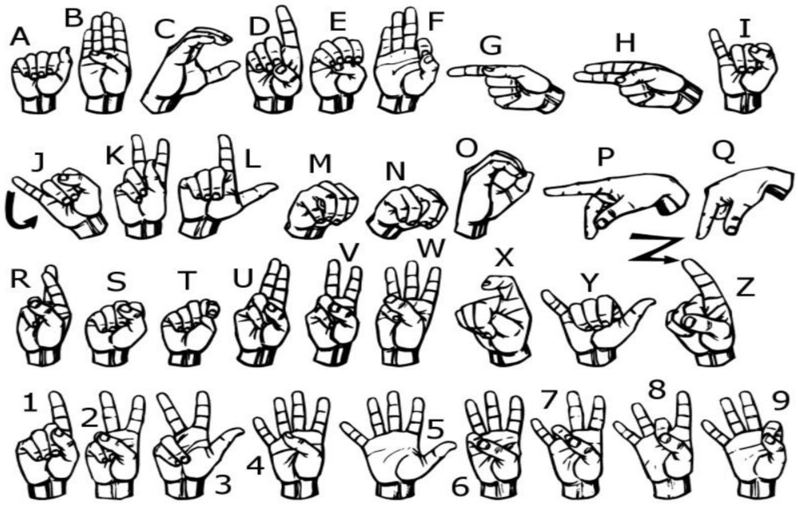


Figure 1.1. Hand gestures for American Sign Language

1.1.2 Deep Learning

Algorithms like artificial neural networks are inspired by the configuration and operation of the mind and deep learning is a subclass for machine learning. Machine Learning(ML), on the other hand, is a subset of Artificial Intelligence(AI), and Deep Learning is a subset of Machine Learning(ML). Artificial intelligence (AI) is a broad term that refers to methods that enable computers to imitate human behaviour. All of this is made possible by machine learning, which is a series of algorithms trained on data. Deep Learning, on the other hand, is a form of Machine Learning that is influenced by the human brain's structure. Deep learning algorithms analyse data using a predetermined conceptual form in order to draw similar results as humans. Deep learning does this by using a multi-layered system of algorithms known as neural networks.

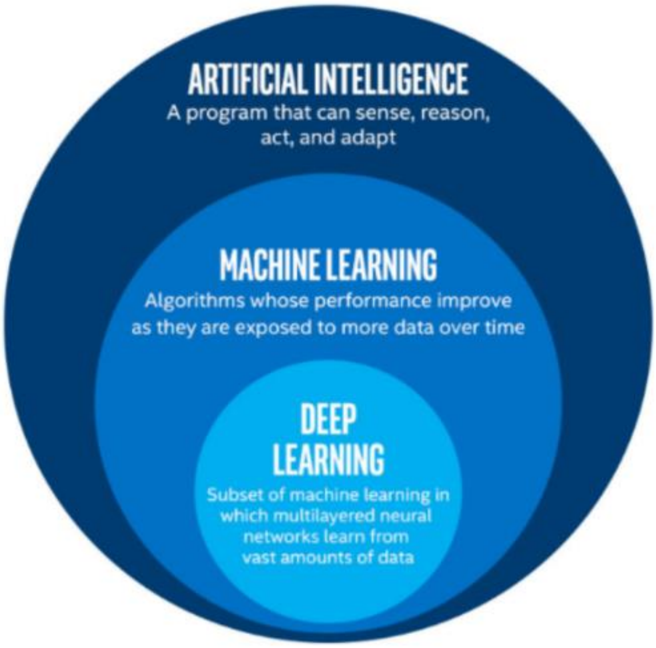


Figure 1.2 AL v/s ML v/s DL

1.1.3 Convolutional Neural Network

In contrast to normal Neural Networks, the neurons in CNN layers are organised in three dimensions: distance, height, and depth. The nerve cells in a stratum will only be attached to a small area of the stratum (window size) before the layer, instead of completely connected in a layer. Furthermore, since we can minimise the entire picture into a single vector of class scores by the end of the CNN architecture, the ultimate outturn stratum will have measurements (number of classes).

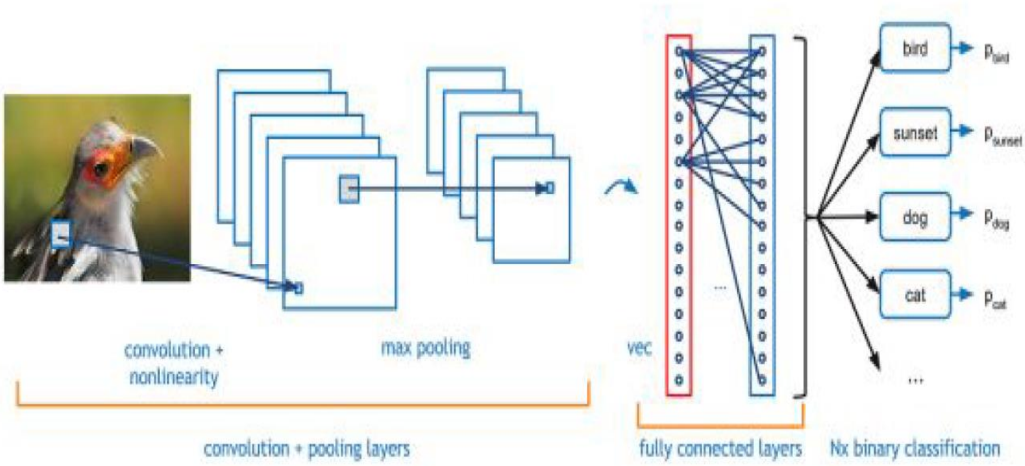


Figure 1.3 Convolution Neural Network

1.1.4 Image Processing

Image processing is a method of improving or retrieving information from a photograph by adding operations to it. It is a mode of gesture analysis where a picture serves as the and also the output is either that image or its characteristics/features. One of the most rapidly emerging technology today is image recognition. Image recognition is one of today's most rapidly evolving technologies. It is also a critical research field in engineering and computer science.

The three phases that make up image processing are as follows:

- Use data acquisition software to import an image. Analysing and modifying the image.
- Examining and altering the picture
- The output, which may be an altered image or a study dependent on image processing.

CHAPTER 2. LITERATURE SURVEY

Hand motion identification has been the subject of extensive study in recent years. The below are some of the common methods to gathering data on hand gestures:

2.1 Use of sensor-based devices

Electromechanical instruments are used to provide precise hand tuning and positioning. To retrieve information, various glove-based methods can be used. However, it is both costly and inconvenient to use.

2.2 Vision based approach

The computer camera is used as an input system in vision-based approaches to observe detail from hands or fingertips. The Vision Based approaches only involve a camera, allowing for normal human-computer interaction without the use of any additional equipment. By defining artificial vision systems that are applied in software and/or hardware, these systems aim to supplement biological vision. The key challenge in visual-oriented hand gesture recognition deals with vast variety of person's hand gesture impression because of large range of palm based gestures, various skin shade options, and differences in view-points, scales, and camera speed when recording the scene.

2.3 Data Pre-processing and Feature extraction for vision-based approach

- In [41] the approach for hand detection combines threshold-based color detection with background subtraction. Adaboost face detector was used to differentiate between faces and hands as both involve similar skin-color.
- The necessary image can be extracted which is to be trained by applying a filter called Gaussian blur. The filter can be easily applied using open computer vision also known as OpenCV and is described in[42].
- For extracting necessary image which is to be trained we can use instrumented gloves as mentioned in [43]. This helps reduce computation time for preprocessing and can give us more concise and accurate data compared to applying filters on data received from video extraction.

2.1 Gesture Classification

- In [41] HMMs (Hidden Markov Models) are used to classify data. This model is concerned with the complex dimensions of the gestures. Gestures are derived from a video sequence using following the blobs of skin colour that belong to the hand gesture. The user's face is concentrated in the facial space. Objective is to understand Deictic and symbolic gestures are the two types of gestures. A quick look–up indexing table is used to filter the image. During the filtering process, the pixels of skin colour are collected together to form blobs. Blobs are geometrical artefacts that are used to evaluate homogeneous regions using the position coordinates X and Y and colorimetry (Y, U, V) which belongs to the pixels of skin colour.
- According to [42] the Naive Bayes Classifier is used, which is an efficient and fast tool for recognising static hand gestures. It works by using geometric invariants obtained from image data after segmentation to categorise different motions. As a consequence, unlike many other methods of identification, this one does not rely on skin tone. For a static backdrop, the gestures are derived from each frame of the image. The first step is to segment, mark, and detach geometric invariants from the properties of interest. It works by using geometric invariants obtained from image data after segmentation to categorise different motions. As a consequence, unlike many other methods of identification, this one does not rely on skin tone.

- As per the paper "Human Hand Gesture Recognition Using a Convolution Neural Network," Hsien-I Lin, Ming-Hsiang Hsu, and Wei-Kai Chen, who are the alumini of National Taipei University of Technology Taipei, Taiwan, create a skin model to remove the gesture of hand from the image and then apply binary threshold to the entire image." They calibrate the threshold image along the principal axis after receiving it in order to focus the image around it. They used this picture to train and predict the outputs of a convolutional neural network model. They trained their model on seven hand gestures, and it produces an accuracy of about 95% for such gestures when used.

CHAPTER 3. METHODOLOGY

Decades after Human-Computer Interaction (HCI) became popular in the industry, Sign Language Recognition (SLR) based on gesture-based communication has been studied. The standard and widespread system for the identification of vision-based sign language is depicted in Figure 3.1, which involves user input, pre-processing, extraction of features, classification based on various methodologies of deep learning, and then output predictions. Sign gesture illustration covers the spot to represent the key data. One of the important parts of the whole frame-work is pre-processing because it involves different stages.

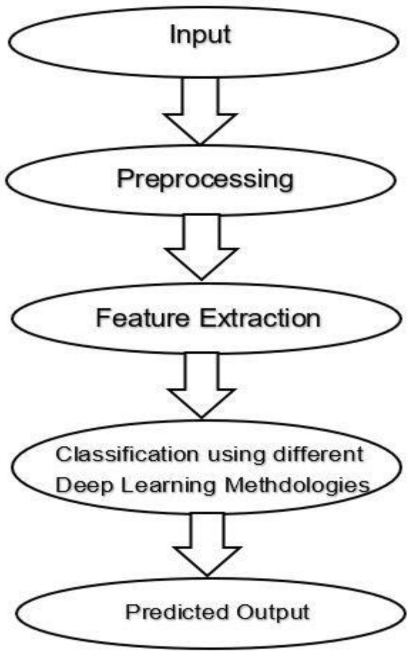


Figure 3.1. General flow-chart for Sign Language Recognition

3.1 Preprocessing

3.1.1 Creation of histogram

The very initial step is hand gesture recognition and for this, a threshold of the image is required. For the creation of the threshold of an image, a histogram is made. Graphical representation of the frequency of pixels in a grayscale image in which pixel value varies from 0 to 255 is known as histogram [5]. To get a good histogram of our hand with respect to surroundings, our hand should cover all those small squares to get a good histogram as shown in Figure 3.2. In determining the threshold value, the histogram is very significant.

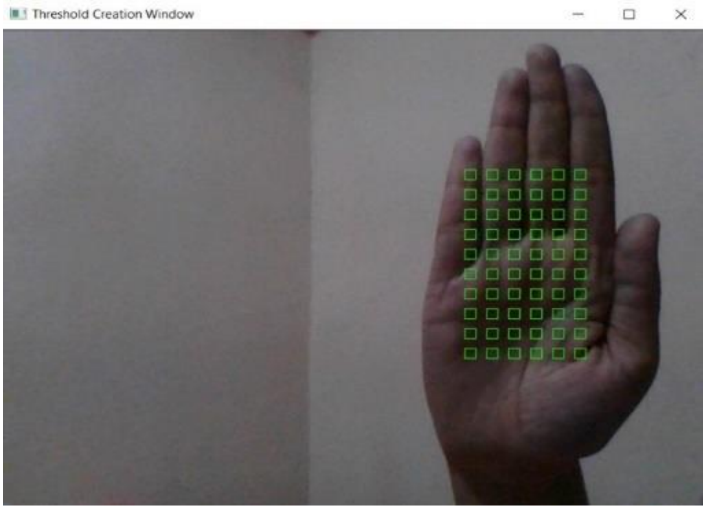


Figure 3.2 Hand covering all green spots to create a good histogram

3.1.2 Separation of background and creation of binary image

The hand gesture area must be removed from the context to make gesture identification easier. Hand and background have different pixel intensity hence, differences in their pixel intensity are used to separate them out from each other [6].

Pixels representing the hand region are assigned with a value so that those pixels could be identified. If pixel value exceeds a threshold value, it is assigned with a value that will represent that it is white, else it is assigned any other value which will represent that it is black. To remove our hand from the background, adaptive threshold method is used and images are scaled to 128 x 128 pixels.

3.1.3 Contour Detection using OpenCV

The contour is formed by joining all the continuous points along the boundary that have the same intensity or colour. Item identification and recognition contours demonstrated to be an effective tool for shape analysis. [7]. Binary images are preferred for better accuracy. Threshold or canny edge detection should be applied before finding contours [7]. Point to note that finding contours in OpenCV is like detecting a white entity on black background. Inbuilt OpenCV functions are used to identify the shape of the hand through the contour detection method. Different contour properties are required for the identification of gestures and to map with their meaning therefore outline is being drawn [6].

3.1.4 Image Labelling

After the creation of grayscale images, those grayscale gesture images are assigned a number that is in sequence and can be considered as a gesture identification number. Later these images are mapped with their corresponding meaning, this process is called image labelling. After this process, the labelled image is saved in the database.

3.1.5 Data Augmentation

While making gestures, restricting users to make gestures from one hand (either left hand or right hand) would not be convenient at all. If they are allowed to do so, it will affect the model accuracy. Therefore, data augmentation is one of the important aspects to get good accuracy. We can define data augmentation as the method by which slightly altered copies of pre-existing data are added to the already existing dataset in order to increase the amount of data which helps the model to learn from a different type of situation. This modification can be done by flipping images along a horizontal or vertical axis or by rotating images by few angles. It behaves as a regulator and helps to bring down the extent of overfitting while training deep learning or a machine learning model [8].

3.2 Feature Extraction

Feature extraction is the method of extracting simplified and discriminative descriptors that are then used as training data for the classifier to improve accuracy and identification. One of the difficult issues that are being faced by the model is to pull out the discriminative feature from different kinds of videos or images exposed to different lighting conditions. The Red-Green-Blue (RGB) input images are converted to grayscale and gaussian blur is applied to minimize unwanted noise.

Figure 3.3 Image after applying Gaussian Blur Filter

3.3 Model Training

The selection of a Deep Learning method among the pool of various deep learning methodologies which can give promising results over varied data is another considerable issue. Among various Deep Learning methodologies, Convolutional Neural Network (CNN) gives promising results while making predictions in vision-based SLR. The architecture of CNN which is being followed is shown in Figure 3.4. The 128x128 pixels is the size of the input images. These are processed using 32 filter weights in the first convolutional layer (3x3 pixels each). This produces a 126x126 pixel representation for each of the filter-weights.

Further, they are down-sampled using max-pooling of 2x2 resulting in images of 63x63 pixels. The input picture was subjected to max pooling with a pool size of (2, 2) and to the Rectified Linear Unit activation function. This decreases the number of parameters, lowering both the computing expense and the risk of overfitting. The resulting 63 x 63 images from the first pooling layer are then served to the second convolutional layer as an input. It is then processed in the second convolutional layer using 32 filter weights (3x3 pixels each). The resulting 60 x 60 pixels images are again down-sampled using a max pool of 2x2 and further reduced to 30x30 resolution of images. These images are fed into a 128-neuron which is a completely connected layer, and the output of the second convolutional layer is then reshaped into a 30x30x32 =28800-value array. This layer receives a 28800-value array as input. The second densely connected layer receives the contribution of all these layers. To prevent overfitting, a dropout layer is used with a value of 0.5. The output from the first densely connected layer is fed into a 96-neuron entirely connected layer. The second densely connected layer's output is fed into the final layer, which would have the same number of neurons as the number of groups were classified (alphabets + blank symbol).

The prediction layer calculates the likelihood of the picture falling into one of the groups. As a result, the output is scaled between 0 and 1, and the sum of each class's values equals 1. We were able to accomplish this by using the Softmax function. The prediction layer's contribution will initially be a little off from the actual value. To improve it, we used labeled data to train the networks. Cross-entropy is the performance measurement that is used in classification.

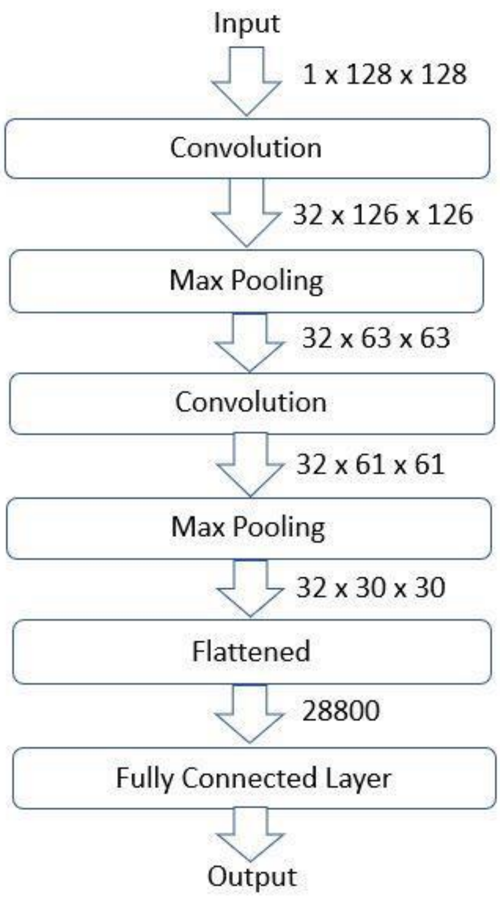


Figure 3.4. CNN Architecture

3.4 Software and Tools Used

3.4.1 Ubuntu 20.04

Ubuntu is a Debian-based Linux distribution which is free to use and has open-source applications. It is available in three different editions for IoT computers and robots: Core, Server, and Desktop. Both editions are compatible with single computer and with a virtual machine. It is a widely used cloud computing operating system that supports OpenStack. Since version 17.10, GNOME has become Ubuntu's default desktop. It is updated twice in a year, with long-term support (LTS) upgrades every 2 years. Version 20.04, which is subsidised until 2025 under public funding and until 2030 as a paid option, is the most recent long-term support release as of October 22, 2020. The most recent standard edition is 20.10, and is supported for 9 months. Canonical and a group of other developers work on Ubuntu under a meritocratic governance paradigm. Beginning on the release date and ending before the release's specified end-of-life (EOL) date, Canonical provides security patches and maintenance for each Ubuntu version. Canonical makes money by selling premium Ubuntu-related services.

Figure 3.5. Ubuntu

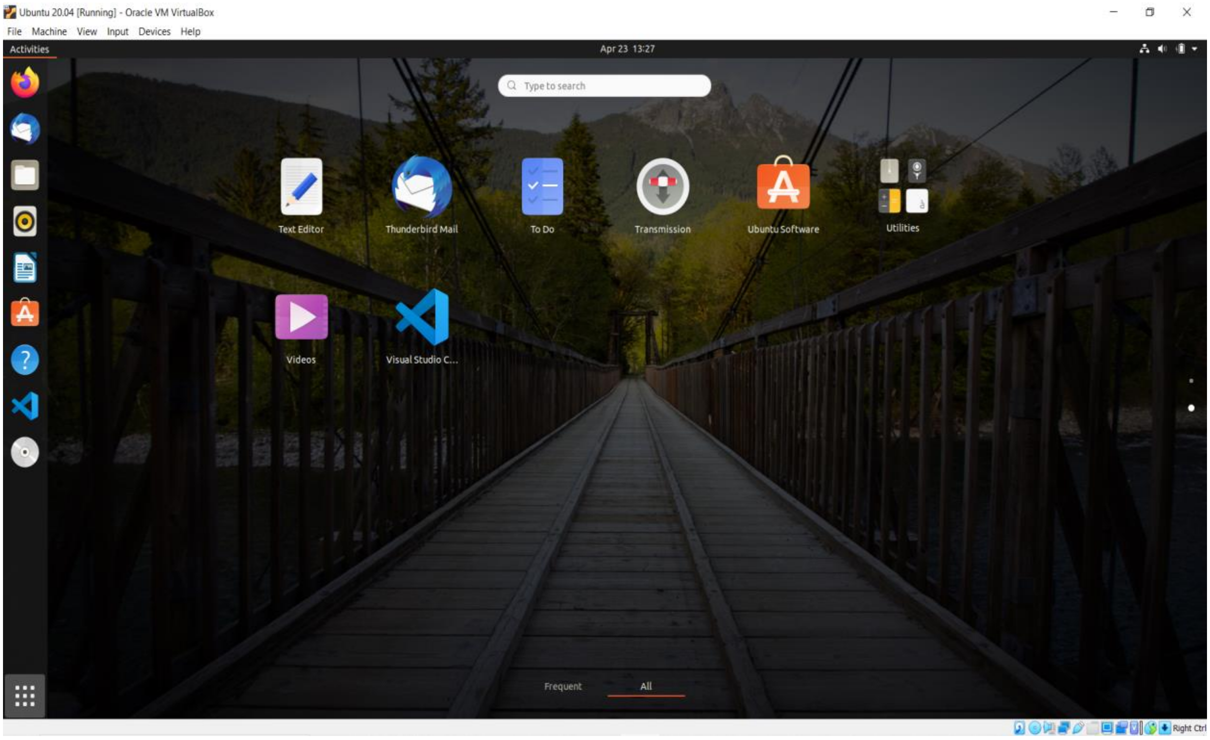


Figure 3.6. Ubuntu 20.04 Interface

3.4.2 TensorFlow

TensorFlow is an open -source software library for machine learning and is free to use. It has a wide range of applications, but it primarily focuses on the training of deep neural network and its inference. The TensorFlow was developed by the Google Brain Team for the internal use of Google. It was made available in public domain in 2015 under the Apache License 2.0. It has a scalable and broad ecosystem of tools, databases, and community resources that provides researchers the state-of-the-art advances in machine learning and help developers to build and deploy machine learning applications rapidly.

Figure 3.7 TensorFlow

3.4.3 Keras

Keras is a software library which is open-source has a wide application in python for artificial neural networks. TensorFlow has a user interface called Keras. Till version 2.3, it supported a variety of backends, including TensorFlow, Microsoft Cognitive Toolkit, Theano, and PlaidML. As of version 2.4, among all only TensorFlow is allowed. It's developed to be user-friendly, scalable, and extensible, with the intention of making fast deep neural network experiments. It was developed as part of the ONEIROS (Open-ended Neuro-Electronic Intelligent Robot Operating System) research project, and is maintained by a Google engineer François Chollet who is also its primary author. It is useful when we need to create and validate a neural network easily and with few lines of code. Layers, goals, activation mechanisms, optimizers, and methods for working with images and text data are all implemented.

Figure3.8 Keras

3.4.4 OpenCV

OpenCV is a large open-source repository for computer vision and has a wide range of applications in machine learning as well as in image recognition, and now it has become a integral part of various real-time operations, that is essential in day-to-day applications. It can be used to recognise objects, faces, and even human handwriting in photographs and videos. Python will process the OpenCV array structure for review when it is combined with different

libraries such as Numpuy. To define image patterns and their various features, openCV has large set of functions to perform various mathematical operations on these features.
It is free for both commercial and academic use and was released under a BSD licence. The first version of openCV was 1.0. It has interfaces for C, Python, C++ and Java interfaces and supports Windows, Mac OS, iOS, Android, and Linux. When OpenCV was first created, the main aim was to make real-time apps as effective as possible. All is written in optimised C/C++ code to take benefit of multi-core processors.

Figure 3.9 OpenCV

3.4.5 Oracle Virtual-box

Oracle VM VirtualBox (formerly Sun VirtualBox, Sun xVM VirtualBox, and Innotek VirtualBox) is Oracle Corporation's open-source hosted hypervisor for x86 virtualization and is free to use. It was created by Innotek and bought by Sun Microsystems in 2008. Sun Microsystems was later acquired by Oracle Corporation in the year 2010.
Windows, Mac OS X, Linux, Solaris, and OpenSolaris are all supported by VirtualBox. FreeBSD and Genode ports are also available. It aids in the development and management of guest virtual machines running BSD, Windows, OS/2, Linux, OSx86, Solaris, and Haiku, as well as the macOS guest virtualization on Apple hardware. We have used oracle virtual box to launch Ubuntu, a Linux distribution in our windows machine.

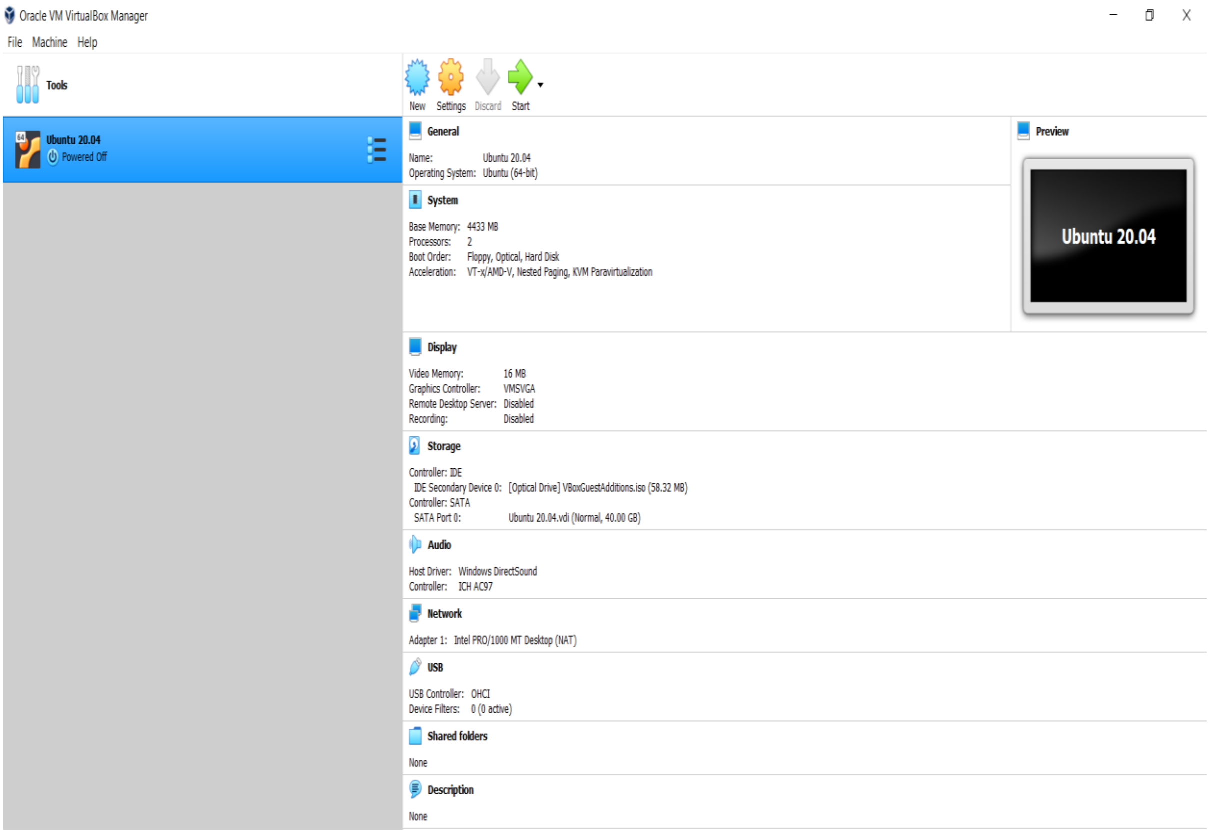


Figure 3.10 Oracle VirtualBox

3.4.6 Visual Studio Code

Visual Studio Code is a open source and free code editor available for Windows, macOS, and Linux from Microsoft. Among the features are intelligent code completion, debugging, snippets, syntax highlighting, embedded Git and code refactoring. Developers can change the preferences, keyboard shortcuts and theme, as well as install plugins to add additional functionality. Although Microsoft's releases are freeware. In the Stack Overflow 2019 Developer Survey, Visual Studio Code was rated the most common programming environment tool, with 50.7 percent of the 87,317 respondents claiming they use it. On April 29, 2015, Microsoft announced the introduction of Visual Studio Code at the 2015 Build conference. A demo prototype was released shortly after. It was released under the MIT License in the year 2015 and source code was made available on GitHub. Extension support was also announced. We have used Microsoft visual studio code to develop our whole project.

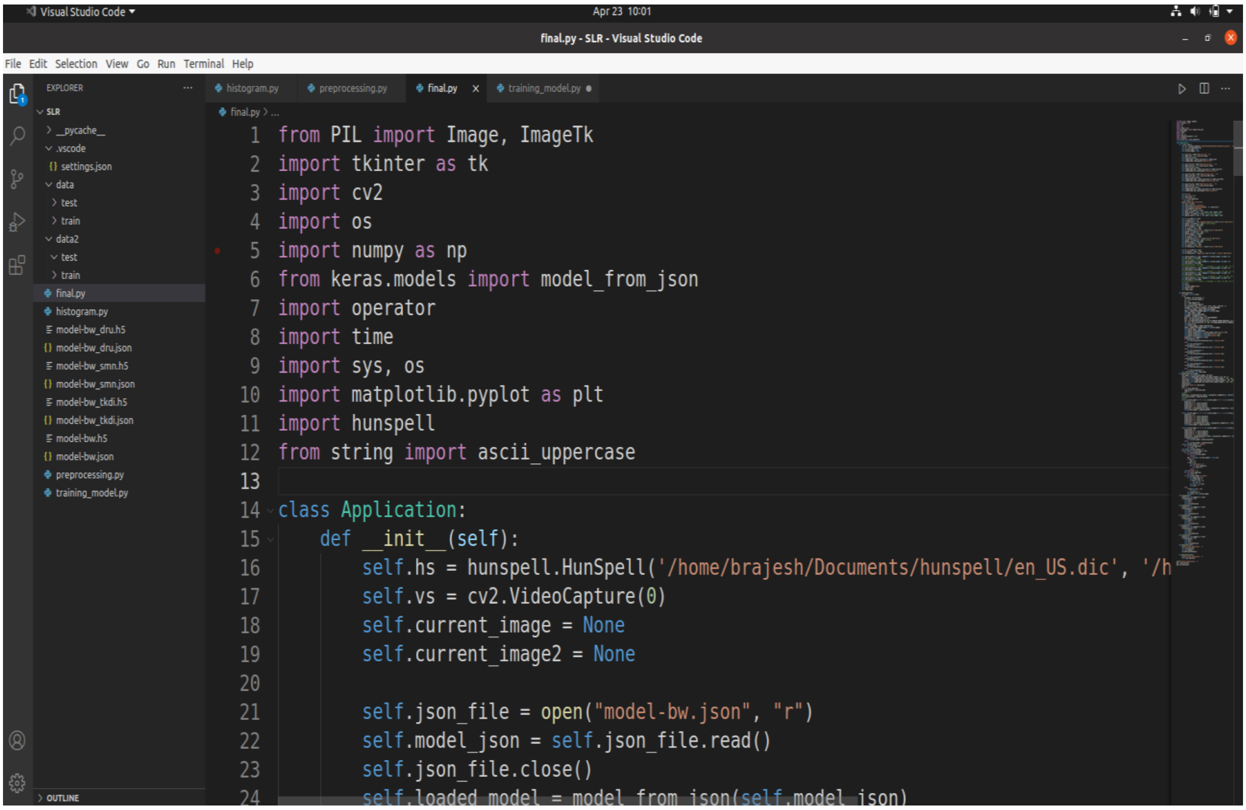


Figure 3.11 Microsoft Visual Studio Code

CHAPTER 4. RESULTS AND DISCUSSION

4.1 Project Outcome

The primary goal was to improve the model's performance with readily available re-sources while also improving its accuracy. Two layers of the CNN model algorithm were used for this. The histogram was used to create a threshold in the first layer, and then a gaussian blur filter was used to remove features from the captured hand motion in the second layer. The data was then augmented to generate more inputs, allowing the model to identify movements from multiple angles and sides. The processed image input was then transferred to the Layer 1 CNN model, which yielded the same results for a particular set of symbols. The symbols for which the model did not make sufficient pre-dictions are described below

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

Then such sets were categorized using 4 different CNN-classifier designed specifically for each of them in layer 2. The principle behind layer 2 is that if the layer 1 model decides that the input belongs to one of those sets, the input would then proceed through the CNN Classifier in layer 2, which was designed explicitly to distinguish ambiguous symbols.

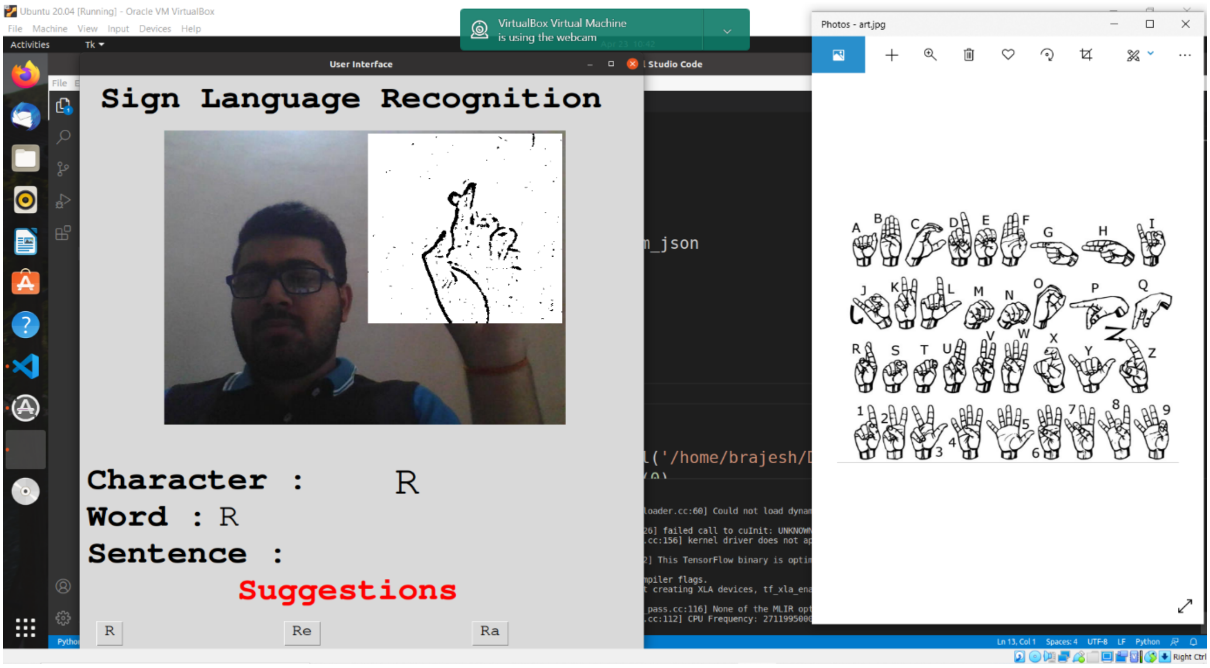


Figure 4.1 Final output of SLR

We obtained 85.4 percent accuracy in our model using just layer 1 of our algorithm, and 95.7 percent accuracy using a mix of layer 1 and layer 2 of our algorithm, which is higher than most existing study papers on American sign language. Figure 4.2 depicts the uncertainty matrix for our model. The model was updated using the Adam optimizer in response to the loss function's output. Adam combines the advantages of two stochastic gradient descent extensions, namely adaptive gradient algorithm (ADA-GRAD) and root mean square descent algorithm (RMSD).

						P	r	e	d	i	c	t	e	d		V	a	l	u	e	s					
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
	A	147	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	2	0	0
	B	0	139	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	0	0	0
	C	0	0	152	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	D	0	0	0	153	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	E	0	0	0	0	152	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	F	0	0	0	0	0	135	0	0	0	0	0	4	0	0	0	0	0	0	0	0	3	10	0	0	0
C	G	0	0	0	0	0	0	150	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
o	H	1	0	0	0	0	0	7	143	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	1
r	I	0	0	0	0	0	0	0	0	150	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
r	J	0	0	0	0	0	0	0	0	0	153	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
e	K	0	0	0	0	0	0	0	0	0	0	153	0	0	0	0	0	0	0	0	0	0	0	0	0	0
c	L	0	0	0	0	0	0	0	0	0	0	0	153	0	0	0	0	0	0	0	0	0	0	0	0	0
t	M	0	0	0	0	0	0	0	0	0	0	2	0	152	0	0	0	0	0	0	0	0	0	0	0	0
	N	0	0	0	0	0	0	0	0	0	0	0	0	0	152	0	0	0	0	0	0	0	0	0	0	0
V	O	0	0	0	0	0	0	0	0	0	0	0	0	0	0	154	0	0	0	0	0	0	0	0	0	0
a	P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	153	0	0	0	0	0	0	0	0	0
l	Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	147	1	0	0	0	0	0	0	0
u	R	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	150	0	0	0	0	0	0	0
e	S	0	0	0	0	1	0	0	0	0	0	0	0	0	0	10	0	0	0	133	0	0	0	0	8	0
s	T	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	151	0	0	0	0	0
	U	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	150	0	0	0	0
	V	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	151	1	0	0
	W	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	149	0	0
	X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	148	0
	Y	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	151
	Z	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 4.2 Confusion Matrix of the proposed model

4.2 Object recognition techniques

In the case of single object recognition, for its clear recognition, 2-Dimensional features like Histogram of Oriented Gradient [10], Scale Invariant Feature Transform (SIFT) [9], Kernel descriptors [11] are used. All these methods are good only for single object detection. Therefore, these methods might not perform well when gestures involve multiple parts of the body as a part of the gesture. SIFT [9] can compute at the edges which will be invariant to scaling, rotation, the addition of noise. One of the advantages of depth images is that it is not get affected by illumination changes. Therefore, this ad-vantage gives an edge to reduce the texture and colour variability which occurs due to background, skin, and hair. Space-Time Occupancy Pattern (STOP) [12] is proposed to resolve the problems like occlusions and noise in-depth images. Space-Time Occupancy Pattern is focused on characterized space-time patterns of human gestures (which are 4-Dimensional in nature), to utilize the temporal and spatial contextual knowledge and permitting intra-class variations [3].

4.3 Sensor Based Technologies

The category in which physical intervention of the user is required includes magnetic, visual, or acoustic sensor-based devices which need to be connected with hands or other body parts to locate their position. Few examples are: information gloves [13], accelerometer [15], digital camera sensor [14]. All these sensor-based technologies vary with different parameters like latency, resolution, accuracy, the vary of motion, price, and user comfort. In this category, individuals are required to wear a bulky gadget and to hold a load of wires or cables that link the device to a computer system. The ease and naturalness of user engagement are impeded by this. Also, when handling such instruments, a lot of precautions and calibration are required. Therefore, this class adds a layer of complexity to the user experience.

On the other hand, there is a category in which physical intervention is not required. This category gets much more appreciation than the previous one when Microsoft Kinect was first introduced. Google Tango Project [18], Microsoft Kinect [16], Leap motion controller [17] provides depth images or maps. As already discussed before, the advantage of depth images is that it does not get affected by illumination changes. Hence, the effect of illumination variability can be decreased to a larger extent. In the past recent years, it has been seen the increment in the studies carried out by researchers focused on human gesture-based recognition by utilising depth information provided by different sensing devices. Google Tango Project [18] is a limited-run experimental device which has a Microsoft Kinect-like vision along with innovative System-on-Chip (SoC) integrated inside. It has also a dedicated and very subtle 3D scanner. All these recent advances would make methods based on depth maps even more convenient and useful for signers [3].

4.4 Different sign languages across the globe

There is no specific standard for sign languages. Diverse gesture-based communications are utilized in numerous nations or locales. Table 1 shows the most commonly used sign languages and their respective geographical region in which they are being spoken.

Table 1 Different sign languages and their respective countries/regions

Sign Languages	Respective Countries/Regions	Reference
British Sign Language (BSL)	United Kingdom	[19]
British, Australian, and New Zealand Sign Language (BANZSL)	United Kingdom, Australia, New Zealand	[19]
French Sign Language (LSF)	France and French-speaking areas of Switzerland	[20]
American Sign Language (ASL)	United States, Canada, West Africa, and Southeast Asia	[21]
Irish Sign Language (ISL)	Ireland	[19]
Chinese Sign Language (CSL or ZGS)	China	[19]
Brazilian Sign Language (Libras)	Brazil	[19]
Indian Sign Language	India	[22]
Thai Sign Language	Thailand	[23]
German Sign Language (DGS)	Germany	[24]
Arabic Sign Language	Arab Middle East	[25]

4.5 Comparison of various deep learning methodologies

Deep neural network (DNN) based methodologies can take in reasonable contextual information from crude pictures or video fragments straightway, which is sturdy and adaptable. Table 2 shows the comparison between different Deep Learning methodologies based on their accuracy.

Table 2 Deep Learning Methodologies and their accuracy

Various Deep Learning Methodologies	Accuracy (%)	Reference
Motion Sensor Information + SVM k-NN	79.83	[26]
Edge Orientation Histogram	88.26	[27]
DBN (3 RBM layers + 1 translation layer)	79	[28]
CNN	82	[29]
DNN + CRF	92	[30]
HMM + BLSTM-NN	96.2	[31]
PCA + HMM	89.10	[32]
Depth Comparison + Random Forest	90	[33]
Full Dimensional Feature + k-NN	99.6	[34]
SURF + SVM	97.13	[35]
AdaBoost + HAAR	98.7	[36]
Ring Projection and Wavelet Transform + GRNN	90.44	[37]

4.6 Challenges Faced

Throughout the project, we faced numerous challenges. The first problem we ran into was a lack of details. Initially, a dataset containing raw images of all ASL characters was to be compiled, however, no fitting dataset could be identified in either of the public dataset repositories. As a result, a new dataset was developed. The next issue was to choose the filter for feature extraction. Various filters, such as gaussian blur, binary threshold, canny edge detection, and others, were tested, with the gaussian blur filter proving to be the most efficient. Aside from that, there were problems with the model's accuracy due to lighting conditions; the model did not have promising accuracy when lighting conditions were poor. However, the advances and other techniques in this area are discussed in further sections which have high cost and the resources used in these methodologies are not easily available.

4.7 Advantages and Disadvantages of Sign Language Interpreter

4.7.1 Advantages

- Friends, customers, and strangers who are Deaf-Blind, Deaf, or hearing-Mute can speak directly with us.
- Using Pro-Tactile, we can explain environmental details even more effectively and in greater detail. We can also get much more detailed explanations, especially for photos and videos. We've been more mindful of aspects of the visual world that we would not have seen otherwise.
- We should fully appreciate the impact of body language and how it shapes language. Based on the nuances of how they sign, I get a very vivid sense of the person's personality and mood. Hearing all of the detail in someone's speech has less of an impact. I'm able to access more subconscious data. I even learn how to more accurately imitate my own body language.
- The mental mapping skills changes dramatically after learning ASL. When we learn ASL and then PTASL, orientation and agility skills improved dramatically.
- We can speak effectively in both noisy environments (concerts, loud restaurants) and quiet environments (libraries, hearing person in the room sleeping).

4.7.2 Disadvantages

- To talk, we need to be in near vicinity. We can't talk at all when we're around a room or through a window or door. We can't talk easily around a table or clock, either.
- Instead of being able to do both at the same time, we often have to split up things, such as talk, then take notes, or speak, then eat.
- At any given time, we can only talk with up to two people (one person signing under each hand). In community contexts, We’ll need interpreters to obey ASL.

- We can fully appreciate the impact of body language and how it shapes language. Based on the specifics of how they sign, we get a very vivid sense of the person's personality and mood. Hearing all of the detail in someone's speech has less of an impact. We are able to access more subconscious data. We even learn how to more accurately imitate our own body language.

CHAPTER 5. CONCLUSION AND FUTURE SCOPE

On the whole, this paper provides the different advances in sign language recognition. Experts in different disciplines have primarily performed past studies on sign language comprehension for numerous sign languages, reducing real-world utility. More than 25 studies were reviewed to get the knowledge of the past researches and then the different works based on sign language recognition using various Deep Learning methodologies and other technological approaches are discussed. This paper also highlights the various terms and processes involved in this area.

From the various studies, it can be depicted that there is no universal sign language. Sign language changes with concerning countries and regions. Therefore, when a deaf or speech-impaired person travels to another country or region, he/she might face the problem because the device in which SLR is integrated might not have the sign language of the region he/she belongs to. Also, Deep Neural Network (DNN) shows promising performance in various practical aspects. But It can learn good features from multi-modal histogram because these histograms can enhance the accuracy of recognition by providing more knowledge or features about sign gestures. So, further studies should be concerned about integrating all the sign languages used across the globe into one device with a convenient and customizable user interface to make it more robust and barrier breaker. In conjunction, in future research work, studying high-level functionality across deeper networks remains questionable to see the potential in terms of usability and accuracy development.

Conclusively, everyone is of value in society and has the right to communicate with others regardless of their physical impairments contemplating this, let us try to embrace hearing and speech impaired people in our everyday life and fostering them ahead in their life by incorporating innovation with their exigencies and other needs.

CHAPTER 6. REFERENCES

1. M. M. Islam, S. Siddiqua, and J. Afnan, “Real time Hand Gesture Recognition using different algorithms based on American Sign Language,” *2017 IEEE Int. Conf. Imaging, Vis. Pattern Recognition, icIVPR 2017*, 2017, doi: 10.1109/ICIVPR.2017.7890854.
2. A. Young, R. Oram, and J. Napier, “Hearing people perceiving deaf people through sign language interpreters at work: on the loss of self through interpreted communication,” *J. Appl. Commun. Res.*, vol. 47, no. 1, pp. 90–110, 2019, doi: 10.1080/00909882.2019.1574018.
3. L. Zheng, B. Liang, and A. Jiang, “Recent Advances of Deep Learning for Sign Language Recognition,” *DICTA 2017 - 2017 Int. Conf. Digit. Image Comput. Tech. Appl.*, vol. 2017-Decem, pp. 1–7, 2017, doi: 10.1109/DICTA.2017.8227483.
4. “The 26 letters and 10 digits of American Sign Language (ASL). | Download Scientific Diagram.” https://www.researchgate.net/figure/The-26-letters-and-10-digits-of-American-Sign-Language-ASL_fig1_328396430 (accessed Feb. 22, 2021).
5. https://www.geeksforgeeks.org/opencv-python-program-analyze-image-using-histogram/ (accessed Feb. 22, 2021).
6. K. Manikandan, A. Patidar, P. Walia, and A. B. Roy, “Hand Gesture Detection and Conversion to Speech and Text,” *arXiv*, 2018.
7. https://docs.opencv.org/3.4/d4/d73/tutorial_py_contours_begin.html (accessed Feb. 22, 2021).
8. https://nanonets.com/blog/data-augmentation-how-to-use-deep-learning-when-you-have-limited-data-part-2/ (accessed Feb. 22, 2021).
9. D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004, doi: 10.1023/B:VISI.0000029664.99615.94.
10. R. Rai, S. Shukla, and B. Singh, “Reactive Power Based MRAS for Speed Estimation of Solar Fed Induction Motor with Improved Feedback Linearization for Water Pumping,” *IEEE Trans. Ind. Informatics*, vol. 16, no. 7, pp. 4714–4725, 2020, doi: 10.1109/TII.2019.2950094.
11. L. Bo, K. Lai, X. Ren, and D. Fox, “Object recognition with hierarchical kernel descriptors,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 1729–1736, 2011, doi: 10.1109/CVPR.2011.5995719.
12. A. W. Vieira, E. R. Nascimento, G. L. Oliveira, Z. Liu, and M. F. M. Campos, “STOP: Space-Time Occupancy Patterns for 3D action recognition from depth map sequences,” *Lect. Notes*

Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 7441 LNCS, pp. 252–259, 2012, doi: 10.1007/978-3-642-33275-3_31.

13. M. Mohandes and S. O. Aliyu, "(12) United States Patent," vol. 2, no. 12, 2017.
14. H. Hongo, M. Ohya, M. Yasumoto, Y. Niwa, and K. Yamamoto, "Focus of attention for face and hand gesture recognition using multiple cameras," *Proc. - 4th IEEE Int. Conf. Autom. Face Gesture Recognition, FG 2000*, pp. 156–161, 2000, doi: 10.1109/AFGR.2000.840627.
15. X. Zhang, X. Chen, Y. Li, V. Lantz, K. Wang, and J. Yang, "A framework for hand gesture recognition based on accelerometer and EMG sensors," *IEEE Trans. Syst. Man, Cybern. Part ASystems Humans*, vol. 41, no. 6, pp. 1064–1076, 2011, doi: 10.1109/TSMCA.2011.2116004.
16. K. Lai, J. Konrad, and P. Ishwar, "A gesture-driven computer interface using Kinect," *Proc. IEEE Southwest Symp. Image Anal. Interpret.*, pp. 185–188, 2012, doi: 10.1109/SSIAI.2012.6202484.
17. C. H. Chuan, E. Regina, and C. Guardino, "American sign language recognition using leap motion sensor," *Proc. - 2014 13th Int. Conf. Mach. Learn. Appl. ICMLA 2014*, pp. 541–544, 2014, doi: 10.1109/ICMLA.2014.110.
18. https://en.wikipedia.org/wiki/Tango_(platform) (accessed Feb. 22, 2021).
19. "A Guide to the Different Types of Sign Language Around the World." https://k-international.com/blog/different-types-of-sign-language-around-the-world/ (accessed Feb. 22, 2021).
20. https://www.tradonline.fr/en/blog/french-sign-language-a-language-in-its-own-right/ (accessed Feb. 22, 2021).
21. https://en.wikipedia.org/wiki/American_Sign_Language (accessed Feb. 22, 2021).
22. https://indiansignlanguage.org/ (accessed Feb. 22, 2021).
23. https://en.wikipedia.org/wiki/Thai_Sign_Language (accessed Feb. 22, 2021).
24. https://en.wikipedia.org/wiki/German_Sign_Language (accessed Feb. 22, 2021).
25. "Arab sign-language family - Wikipedia." https://en.wikipedia.org/wiki/Arab_sign-language_family (accessed Feb. 22, 2021).
26. A. Vaitkevičius, M. Taroza, T. Blažauskas, R. Damaševičius, R. Maskeliunas, and M. Woźniak, "Recognition of American Sign Language gestures in a Virtual Reality using Leap Motion," *Appl. Sci.*, vol. 9, no. 3, pp. 1–16, 2019, doi: 10.3390/app9030445.

27. J. R. Pansare and M. Ingle, "Vision-based approach for American Sign Language recognition using Edge Orientation Histogram," *2016 Int. Conf. Image, Vis. Comput. ICIVC 2016*, pp. 86–90, 2016, doi: 10.1109/ICIVC.2016.7571278.
28. L. Rioux-Maldague and P. Giguere, "Sign language fingerspelling classification from depth and color images using a deep belief network," *Proc. - Conf. Comput. Robot Vision, CRV 2014*, no. May 2014, pp. 92–97, 2014, doi: 10.1109/CRV.2014.20.
29. S. Ameen and S. Vadera, "A convolutional neural network to classify American Sign Language fingerspelling from depth and colour images," *Expert Syst.*, vol. 34, no. 3, 2017, doi: 10.1111/exsy.12197.
30. T. Kim *et al.*, "Lexicon-free fingerspelling recognition from video: Data, models, and signer adaptation," *Comput. Speech Lang.*, vol. 46, pp. 209–232, 2017, doi: 10.1016/j.csl.2017.05.009.
31. P. Kumar, H. Gauba, P. Pratim Roy, and D. Prosad Dogra, "A multimodal framework for sensor based sign language recognition," *Neurocomputing*, vol. 259, no. 2017, pp. 21–38, 2017, doi: 10.1016/j.neucom.2016.08.132.
32. M. M. Zaki and S. I. Shaheen, "Sign language recognition using a combination of new vision based features," *Pattern Recognit. Lett.*, vol. 32, no. 4, pp. 572–577, 2011, doi: 10.1016/j.patrec.2010.11.013.
33. C. Dong, M. C. Leu, and Z. Yin, "American Sign Language alphabet recognition using Microsoft Kinect," *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, vol. 2015-Octob, pp. 44–52, 2015, doi: 10.1109/CVPRW.2015.7301347.
34. D. Aryanie and Y. Heryadi, "American sign language-based finger-spelling recognition using k-Nearest Neighbors classifier," *2015 3rd Int. Conf. Inf. Commun. Technol. ICoICT 2015*, pp. 533–536, 2015, doi: 10.1109/ICoICT.2015.7231481.
35. C. M. Jin, Z. Omar, and M. H. Jaward, "A mobile application of American sign language translation via image processing algorithms," *Proc. - 2016 IEEE Reg. 10 Symp. TENSYMP 2016*, pp. 104–109, 2016, doi: 10.1109/TENCONSpring.2016.7519386.
36. V. N. T. Truong, C. K. Yang, and Q. V. Tran, "A translator for American sign language to text and speech," *2016 IEEE 5th Glob. Conf. Consum. Electron. GCCE 2016*, 2016, doi: 10.1109/GCCE.2016.7800427.
37. Y. Fiagbe, "World Journal of Engineering," no. April, 2020.

38. M. Taskiran, M. Killioglu, and N. Kahraman, "A Real-Time System for Recognition of American Sign Language by using Deep Learning," *2018 41st Int. Conf. Telecommun. Signal Process. TSP 2018*, pp. 1–5, 2018, doi: 10.1109/TSP.2018.8441304.
39. A. Joshi, H. Sierra, and E. Arzuaga, "American sign language translation using edge detection and cross correlation," *2017 IEEE Colomb. Conf. Commun. Comput. COLCOM 2017 - Proc.*, 2017, doi: 10.1109/ColComCon.2017.8088212.
40. O. Koller, S. Zargaran, H. Ney, and R. Bowden, "Deep sign: Hybrid CNN-HMM for continuous sign language recognition," *Br. Mach. Vis. Conf. 2016, BMVC 2016*, vol. 2016-Septe, no. August, pp. 136.1-136.12, 2016, doi: 10.5244/C.30.136.
41. R. Shafieian, Literature Review Literature Review Literature Review," vol. 27, no. 8, pp. 22–39, 2020.
42. "Smoothing Images — OpenCV 2.4.13.7 documentation."https://docs.opencv.org/2.4/doc/tutorials/imgproc/gausian_median_blur_bilateral_filter/gausian_median_blur_bilateral_filter.html (accessed Apr. 10, 2021).
43. "Sign language to audio translator by Parmjeet Singh." https://prezi.com/p/walg3xnnbtjy/sign-language-to-audio-translator/ (accessed Apr. 10, 2021).

APPENDIX 1: Time-Table of the Project Implementation

Table 3 Complete work plan with timeline

S.No.	TIMELINE	WORK DONE
1.	2^{nd} week of January 2021	Project Selection
2.	3^{rd} week of January 2021	Project discussion with our mentor and discussing its application and impact on society.
3.	4^{th} week of January	Selection of appropriate research and review papers which was relevant to our project and their study.
4.	1^{st} week of February	Studying selected research papers and review papers
5.	2^{nd} week of February	Started research paper writing and side-by-side started project implementation.
6.	2^{nd} week of March	Searching for different Scopus/SCI/WoS/UGC listed journals and conferences related to our project.
7.	2^{nd} week of March	Project Analysis and study of different topics and additional tools required during implementation of project.
8.	3^{rd} week of March	Found bugs in the project and fixed it up and tested it for output.
9.	4^{th} week of March	Plagiarism check of prepared manuscript by our mentor and approaching to different journals and conferences for publication.
10.	1^{st} week of April	Started preparing capstone project report.

11.	2nd week of April	Submission of final capstone project report and Power-Point presentation and got the acceptance from the conference for paper publication.
12.	3rd week of April	Plagiarism check of capstone report prepared and sent all the required documents to the conference organising institute i.e, Kurukshetra University, Haryana.
13.	4th week of April	Presentation of paper at ETBS Confernece-2021 organised by Kurukshetra University, Haryana at 28th - 29th April 2021. Also, details about research publication is updated on UMS portal.

APPENDIX 2: Acceptance from International Conference on Emergent Converging Technologies & Biomedical Systems 2021, Organised by UIET, Kurukshetra University, Haryana

Springer

International Conference on
Emergent Converging Technologies and Biomedical Systems, (ETBS-2021)
Organized By:
UNIVERSITY INSTITUTE OF ENGINEERING &TECHNOLOGY
KURUKSHETRA UNIVERSITY, KURUSHETRA, HARYANA, INDIA
Website: https://www.uietkuk.ac.in/etbs2021
28th -29th April, 2021
Sponsored by: Springer & MHRD, TEQIP-III

Session 6

Session Chairs: Dr. Roshani Raut, Associate Professor and Associate Dean International Relations, Pimpri Chinchwad College of Engineering, PCCOE, Pune, India,
Dr.Pradnya Borkar, Associate Professor and R&D Cell Incharge, Jhulelal Institute of Technology, Nagpur, India
(Session Co-ordinator: Dr. Deepak Sood, Mrs. Sonia Saini)
(Date: 29/04/2021 Time: 09:30 AM to 10:30AM)

Paper ID	Presenter Name	Title
242	Deepa Jose, Pavithra , Sasipriya S, Satya Venkata Santosh, Jhonatan Fabricio Meza Cartegana	Effective Plant Leaf Disease Detection System for Farmers
250	Muskan Jain, Manpreet Singh Bajwa, Hemant Kumar	AGRICULTURE ASSISTANT FOR CROP PREDICTION AND FARMING SELECTION USING MACHINE LEARNING MODEL WITH REAL TIME DATA USING IMAGING THROUGH UAV DRONE
252	Dr. Sapna Gambhir, Dr. Parul Tomar, Preeti Sharma	Handling Security Issues In Fog Computing Environment Using Blowfish Encryption Algorithm
306	Kamaldeep Sharma, Brajesh Kumar, Deepti Sehgal and Ayan Kaushik	A COMPREHENSIVE ANALYSIS ON TECHNOLOGICAL APPROACHES IN SIGN LANGUAGE RECOGNITION
308	Santosh Kumar Satapathy, Hari Kishan Kondaveeti, Ravisankar Malladi	Accurate Machine Learning-Based Automated Sleep Staging using Clinical Subjects with Suspected Sleep Disorders
321	Shivani and Nidhi Gaur	Design of Visual Image Classifier for Web Application

Important Note:** Due to the evolving public health concerns and travel advisories issued for avoidance of non-essential travel in view of COVID-19, the **Conference ETBS-2021 is being conducted online. The link for joining conference and Instructions for online presentation will be sent through email of the registered participants separately

Gmail   etbs

Compose
Inbox 7,092
Starred
Snoozed
Important
Sent
Drafts 22
Categories
Personal
Unwanted 47
Meet
New meeting
Join a meeting
Hangouts
Brajesh
No recent chats

1 of 4

UIET (Conference) .
to kamaldeep.16958, me, deese2398, ayan171299
Apr 13, 2021, 12:13 PM (9 days ago)

Dear Author,
Congratulations!

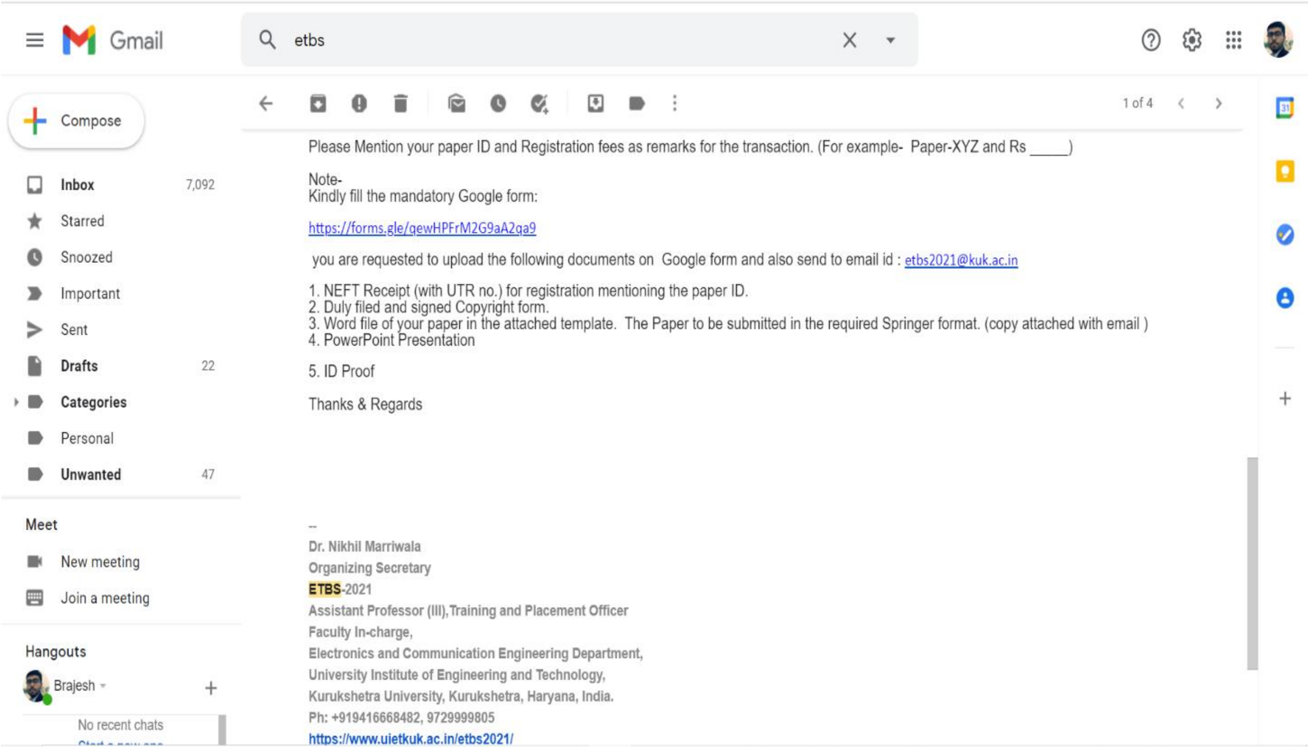
Your paper entitled "**A COMPREHENSIVE ANALYSIS ON TECHNOLOGICAL APPROACHES IN SIGN LANGUAGE RECOGNITION**" with paper ID UIETKUK/ECE/ETBS2021/306 has been accepted for publication in the International Conference on Emergent Converging Technologies and Biomedical Systems (ETBS-2021) happening during 28th – 29th April 2021 in University Institute of Engineering & Technology, Kurukshetra University, Kurukshetra, Haryana, India via (virtual/online mode). All the accepted papers will be published by Springer in the proceedings in the book series "Lecture Notes in Electrical Engineering" link: https://www.springer.com/series/7818.This series is indexed by EI Compendex and Scopus databases.
You are requested to reduce plagiarism below 10% and check plagiarism using Turnitin software.
Please find attached herewith the Springer word Template, CopyRight form and follow the below-mentioned guidelines (strictly), for the final submission of your manuscript.

You are required to fill the Google form link https://forms.gle/qewHPFrM2G9aA2qa9 and Join WhatsApp group with link: https://chat.whatsapp.com/Cl6Bly9xz6q4deT5jvaTul.

Due to the evolving public health concerns and travel advisories issued for the avoidance of non-essential travel in view of COVID-19, Team ETBS-2021 has decided to conduct the Conference online. The decision has been taken while keeping in mind stakeholders' health, safety and academic commitments. The conference schedule will be uploaded on the conference website soon (https://www.uietkuk.ac.in/etbs2021/). For further details and upcoming information about the conference, you are requested to constantly visit the website.

All the payments shall be made through RTGS or NEFT. Account Details are as follows-

ACCOUNT NO: 4673000100018617
ACCOUNT NAME: Scholarship Account UIET
Kurukshetra University, Kurukshetra
BANK: Punjab National Bank
BRANCH: PNB GURUKUL
IFSC Code: PUNB0467300



Please Mention your paper ID and Registration fees as remarks for the transaction. (For example- Paper-XYZ and Rs _____)

Note-
Kindly fill the mandatory Google form:

https://forms.gle/qewHPFrM2G9aA2qa9

you are requested to upload the following documents on Google form and also send to email id : etbs2021@kuk.ac.in

1. NEFT Receipt (with UTR no.) for registration mentioning the paper ID.
2. Duly filed and signed Copyright form.
3. Word file of your paper in the attached template. The Paper to be submitted in the required Springer format. (copy attached with email)
4. PowerPoint Presentation

5. ID Proof

Thanks & Regards

--
Dr. Nikhil Marriwala
Organizing Secretary
ETBS-2021
Assistant Professor (III),Training and Placement Officer
Faculty In-charge,
Electronics and Communication Engineering Department,
University Institute of Engineering and Technology,
Kurukshetra University, Kurukshetra, Haryana, India.
Ph: +919416668482, 9729999805
https://www.uietkuk.ac.in/etbs2021/

APPENDIX 3: First Page of Research Paper

A COMPREHENSIVE ANALYSIS ON TECHNOLOGICAL APPROACHES IN SIGN LANGUAGE RECOGNITION

Kamaldeep Sharma[1[0000-0002-4869-8757]], Brajesh Kumar[2[0000-0001-5101-7783]],

Deepti Sehgal[3[0000-0001-8931-4273]] and Ayan Kaushik[4[0000-0003-1012-7972]]

[1] Lovely Professional University, Phagwara, Punjab, India
kamaldeep.16958@lpu.co.in
[2] Lovely Professional University, Phagwara, Punjab, India
brajeshkumarjha98@gmail.com
[3] Lovely Professional University, Phagwara, Punjab, India
deese2398@gmail.com
[4] Lovely Professional University, Phagwara, Punjab, India
ayan171299@gmail.com

Abstract. Communication is one of the imperative sides of the 21st century. We as human-being can't even imagine this world without communicating with each other. For honest communication to happen two persons should agree on a common language that each of them understands. Keeping this context in mind, there'll be a necessity for a translator in between a hearing/speech impaired person and a normal person. But the translators or consultants who understand sign language do not seem to be simply and commonly offered all the days, if offered conjointly, they charge tons. To help the social interaction of hearing-impaired - speech debilitated individuals with the society economical interactive communication tools are made. As the importance of sign language interpretation is increasing day by day so several triple-crown applications for linguistic communication recognition comprise new forms of real-time interpretation with progressed artificial intelligence and image processing approaches. In this paper, we gave an entire outline of various methodologies and processes rooted in deep learning and discussed technological approaches for linguistic communication interpretation, also achieving 95.7% accuracy with 2 layers of CNN model classifier for the sign gestures of 26 English alphabets with readily available resources. This paper also provides current analysis and advances in this area and tends to distinguish openings and difficulties for future exploration.

Keywords: Linguistic Communication, Deep Learning, Convolutional Neural Network, Image Processing, Sign Language Recognition