



# INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X

Impact factor: 4.295

(Volume 5, Issue 3)

Available online at: [www.ijariit.com](http://www.ijariit.com)

## MINI: Multidimensional Integrated Neural Interface

Juned Javed Kadri

[junedkadz@gmail.com](mailto:junedkadz@gmail.com)

Bharatiya Vidya Bhavan's Sardar Patel Institute of Technology,  
Mumbai, Maharashtra

### ABSTRACT

Data science today, have reached up to great heights and it has achieved many milestones with great innovation and ideas. On the basis of that have come with an idea where it's a help to mankind. In today's world with such a growing technology we have many people with different disabilities in them but with the help of technology, they are able to recover it to a great extent. Increasing the level of technologies alternately put an effect on the increase of its uses in the real world which helps the human being in various aspects. With such hope, to help mankind we have come up with an idea where we can help the people who cannot speak due to various reasons. So, we can come up with a technology where it will help them to express their feelings to the world without speaking.

**Keywords**— Data science, AI, Convolution Network, Neural Network, Dataset, Interface, Video Processing

### 1. INTRODUCTION

Technology in the real world has changed a lot from past where in past we use to think of some application, but we have now revolutionized the technology in today's world which was impossible in past but today with the new form of technology it is possible, to have those things in the real world. The human being is always in a race to enhance the technology which indirectly improved its way of life towards the modern world. Humans invent different technology to make their own life easy. Many times, people with various disabilities which keeps them a little backward from the race of the new world. In this, we think of having such a technology which can help the people. Some people have seeing disability, some have a hearing disability and many of them have speaking disability and various different disabilities.

Speaking nowadays is very essential, to communicate with anyone or sharing our feelings with them. So many a time we need the help of a second person to help us to express our feelings to the world. But with today's technology, it is possible that we can express our feeling with the use of technology. We don't need any other human intervention for our help technology will help us to do that. Many technologies have been evolved with their innovation to deliver the world with the best it needs and many a time it is noticed that the technology has fulfilled the

worlds need to a great extent. Expecting a technologically bright future ahead.

### 2. TECHNOLOGY OVERVIEW

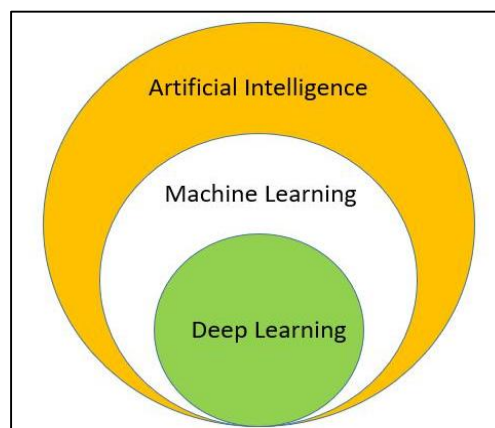


Fig. 1: Technology overview

The Artificial intelligence, machine learning, and deep learning have revolutionized the world of technology which we have thought of in the past about some application to be moving from the imaginary world to the real world. Machine learning, Deep learning, and Artificial intelligence really challenging the world with its various features and it's also changing the vision of the world to see the different thing in its own significance ways. It is helping us to make the things which were imaginary but know we can have it into the real world like for example image recognition, prediction on various expect self-driving vehicles, etc. It has helped to advance more in the automation field along with more accuracy. Making machines becoming our slaves to do our work efficiently and in a more appropriate way [6].

### 3. DATASET

The dataset is a piece of information which contains the data of all the gesture which is required for the system. The system lens which captures the images, that images the system will check in the dataset and according to the match the system will provide the relevant output. In our system the dataset contains two sets of information one information will be predefined as it contains the information of the all the gestures of the human being and on

the other hand it will contain the information which the lens will capture when it will first identify an individual on which the process is going to do the operation [7].

#### 4. CONVOLUTION NEURAL NETWORK

The convolution neural network or we can say CNN which will be playing the main role for the system to perform its operational task. The CNN with the help of dataset will do the comparison and verification and will provide the output [6].

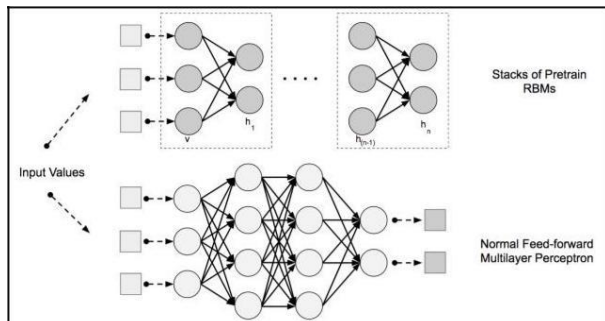


Fig. 2: CNN

The CNN works on the basis of getting the input and the process will be done on the hidden layers and after the process, the system will provide the relevant output.

#### 5. VIDEO PROCESSING

As we are aware that a video comprises of multiple additions of an image which is called frames. The image frames run so fast that it looks like a moving picture. In this system, the CNN will work, do the operation on these frames and then provide the output. When the system captures the frames, we can see that many of the frames are similar to each other so if the system identifies such frames it will discard those frames that will allow the system to process fast and less space will be occupied in memory[8].

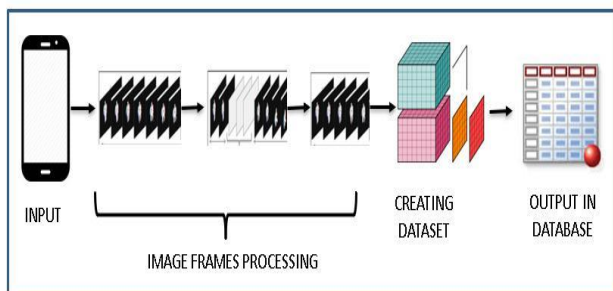


Fig. 3: Video processing

#### 6. MINI (MULTIDIMENSIONAL INTEGRATED NEURAL INTERFACE)

It is a technology which will be used to help the deaf and dumb people to express their feeling to the world, we would be requiring the various aspects for the development of such an application like an interface which will take the input, process in the backend and deliver the output to the world. The most important thing in the whole process is the accurate dataset which will be having all the gestures data from where the system will get the data. These gestures will be created and will be saved in the database which will be called as a gesture dataset. As it has multiple dimensions in it for its comparison so it compares always with the dataset to provide the output.

#### 7. THE PROCESS

##### 7.1 Capture the images/ frames [4]

With the help of the camera lens, we can get the input from the external world, like we can capture the video from the external

world then the system job will start it will break the video into frames and we can then again break the frames into pixels. The pixels will identify the hand shape like it will first identify the user, generate the outline of the body parts of the users and keep that outlines into its database so when the user provides the gestures with hand or any other part of the body the system will be able to identify the correct gesture[5]. Later the system will pre-process the outlined image and this is how the system will differentiate the body parts and provide the output [2].

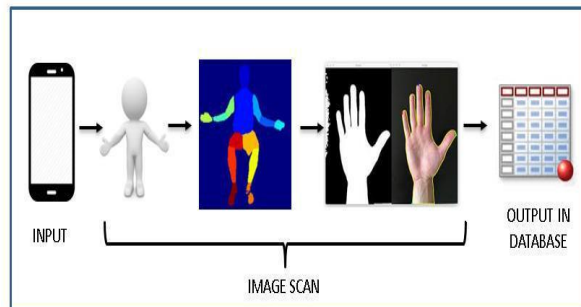


Fig. 4: Capture the images

##### 7.2 Creating the dataset

The dataset is the kind of database which will keep all the data imported from the real world which will be used during the process. Dataset is the collection of the data where we will be having all the dataset of gestures stored in the system. All the gesture images will be stored into the dataset. So, whenever the system needs to compare it will be available.

##### 7.3 Identifying the images (gesture)

The lens after collecting all the data it will perform the outlier test like the system will make the outline of the body parts and make the relevant decision from where it will get the information and later it does a comparison check and after getting a match from the dataset it will provide an output. [1]

##### 7.4 Classification step

Classifying the correct image is very important to provide a correct output so for that first we need to have a proper and efficient dataset. Later when the system gets the image it will compare and classify it with the dataset and provide the relevant output. This step will help in providing a system to provide an easy output without any confusion to the system. We can use the k-nearest classification concept. To classify the images which can be the nearest match to the dataset towards the gesture capture by the system [7]

##### 7.5 System performing analysis [6]

This the most important and crucial step where all the processing will be done. The system will use the CNN which is the Convolution Neural Network method for processing the images classify it to compare it with the dataset and after doing various operations will provide an output.

##### 7.6 Providing the output into text format

As the system will get the correct output after the comparison the images will be having information about what images express which kind of word. Like for every image will be specifying a specific word so that as per the continuous flow of the images the system will compare the image from dataset get the relevant image and then it will give the proper wording for that images which it will get from the dataset.

##### 7.7 Converting the text into a voice [4]

As we get the output from the system that will be in alphanumeric format, so to hear those words we need to convert

those words into speech for that we can use text to speech algorithm.

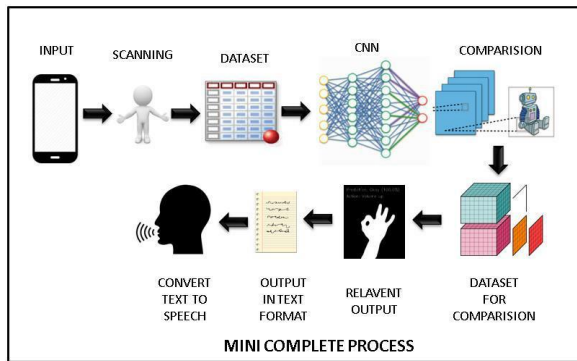


Fig. 5: Converting the text into a voice

## 8. CONCLUSION

After the whole process, we conclude that this process is a very fast process which requires a powerful processor to process it on a very faster rate. MINI (Multidimensional Integrated Neural Interface) is a multi-tasking process which scans and also process at the same time making the action to convert to text than to speech.

This system can give maximum efficiency with more accuracy. Many a time many different algorithms need to use and needs to function at the same time to achieve recognition speed and provide the relevant output very fast. This process will definitely help dumb people to express their feelings to the world and also can explore themselves in the field where speech is required. The ultimate gain of the proposed system is to convert the action into text and then ultimately into speech accurately with less time used in the processing and providing relevant desire output.

## 9. ACKNOWLEDGEMENT

It is my honor and privilege to thank my mentor and guide, Prop. Taruna Sharma for sparking my interest in Machine learning, AI and deep learning technologies. I would also like to grant my privileged, and thank my HOD, Dr Pooja Raundale, for providing us with an opportunity, building our strength and motivating us to do research on such an upcoming and advance technology.

## 10. REFERENCES

- [1] Richard E. Woods, Rafael C. Gonzales, "Digital image processing", Prentice Hall.
- [2] D. M. Gavrilu, "The visual analysis of human movement: A survey on visual analysis."
- [3] P. Jenifer Martina, P. Nagarajan, P. Karthikeyan, "Hand Gesture Recognition Based Real-time Command System" IJCSMC, Vol. 2, Issue. 4, April 2013
- [4] S.A.Patil, P.M.Jadhav, Poonam shetake "Review of text to speech conversion method", International Journal of Electrical Engineering and Industrial Electronics, ISSN: 2347-6982.
- [5] Caroline Rebecca Pantofaru, James MacLean, "Fast Hand Gesture Recognition for Real-Time Teleconferencing Applications" RATFG-RTS '01 Proceedings of the IEEE ICCV Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems (RATFG-RTS'01)
- [6] Josh Patterson and Adam Gibson, "Deep Learning a practitioner's approach"
- [7] Peter Harrington, "Machine learning in Action" 2012 by Manning Publications Co.
- [8] A, Murat Tekalp, "Digital video processing", University of Rochester