REAL-TIME COMMUNICATION SYSTEM POWERED BY AI FOR SPECIALLY-ABLED: A LITERATURE SURVEY

Team ID : PNT2022TMID09385

I . INTRODUCTION:

In our society, we have people with disabilities. The technology is developing day by day but no significant developments are undertaken for the betterment of these people. Communications between deaf-mute and a normal person has always been a challenging task. It is very difficult for mute people to convey their message to normal people since normal people are not trained on hand sign language.

In emergency times conveying their message is very difficult. The human hand has remained a popular choice to convey information in situations where other forms like speech cannot be used. Voice Conversion System with Hand Gesture Recognition and translation will be very useful to have a proper conversation between a normal person and an impaired person in any language.

We aim to develop a system that converts the sign language into a human hearing voice in the desired language to convey a message to normal people. We are making use of a convolution neural network to create a model that is trained on different hand gestures. An app is built which uses this model. This app enables deaf and dumb people to convey their information using signs which get converted to human-understandable language and speech is given as output.

II. LITERATURE SURVEY:

1. **Based Real Time Communication for Physically and Speech Disabled People (Ong Chin Ann, Marlene Valeriu Lu – 2019)**

Communication is a social process of exchanging information from one entity to another in verbal and non-verbal form. It defines our existence and it is an important instrument that connects people together. It comes naturally as a raw skill embedded in most people at birth and we acquired the ways of communication through cognitive learning. Communication is the basis, which drives the process of development in all the fields (Manohar, 2008) and it is the very core of our civilisation. The ability to communicate allows us to express emotion, feelings, convey our thoughts and ideas as well as to relate our experiences. It plays an important role in the dis- semination of information and sharing of knowledge especially in the academic arena. Research has found that human started to learn how to communicate with each other since they are born not only through spoken and written languages but also body gesture, posture, facial expression and eye contacts (Busso, et al., 2004; Cohen, Grag & Huang, 2000).

Communication skill might come as a natural ability in majority of people. However, there are some people inflicted with some form of physical defects which affect their ability to communicate. One of the more severe disabilities is known as “cerebral palsy”, a congenital disorder at birth which causes abnormality in their Motor system. It affects their muscle movement and coordination, learning and Speech abilities. Their malfunctioned motor system causes an uncontrollable and involuntary movement. They are unable to control their oralfacial muscles, thus affects their ability to perform facial expression appropriately.

From the limitation of the existing tools reviewed (Novita, 2006; Macsolvers, 2009; Standup, 2006; Universiteit van Amsterdam, 2008; Crestwood, 2009; Sci-enceDaily, 2008), there is still a pressing need for more effective and efficient tools to alleviate this problem. One the possible methods are to implement a facial expres-sion recognition system to predict or determine the emotional state of a disabled person through his expression projected on his face. biometrics information system can be employed as a means to detect and classify the physiological aspect of a person in real time. Franco and Treves (2001) further support the notion that facial expression can be used for human computer interaction and usability enhancement.

Based on the problem statements deliberated above, we propose an improved real-time communication system using machine learning and computer vision. The aim is to create a communication channel between the specially abled and the society, so they can express there feelings, thoughts and understand other people’s feelings and thoughts through real time communication and facial expressions.

2. Systematic review of computer vision semantic analysis in medical (Antonio Victor Alencar Lundgren, Byron Leite Dantas Bezzerra – 2021)

Medical diagnosing techniques have fascinated us for a long time. It has been common for us to use them in our daily life and implement these technologies. Machine learning and especially computer vision contribute a lot in medical science, which make different difficult tasks easy for doctors and more tolerable for patients. They are widely useful in early detection of disease, and hence are a valuable tool to save human life. Cardio graphic techniques are a must for old age and infant safety.

These include:

- **Retinoscopy -** They although primitive in approach are a must once in a life time and retinoscopy have made yet successful to measure activities of rod and cone receptors in our eyes. Retina has three distinct areas for colors - erythrolabe, chlorolabe and cyanolabe…which are analogical to pixel fixation and identification algorithms on machine learning.

- **Tumor detection -** Cancer is spreading in the world affecting billions of lives both in terms of life and money… machine learning diagnosing systems apply their identification systems to further develop accurate detection in terms of size, location, quality of such tissues which are suspected to become malignant uncontrolled group of fast dividing cells.

- **CT scan - CT scan –** A very common term for cancer patients which uses electromagnetic radiations under manually operated controlled computer vision gratings which are so accurate that it can measure a pigment called c-125 in blood.

3. **A survey on Facial Emotion Recognition Techniques** (Felipe Zago Canal, Tobias Rossi Muller, Gustavo Gino Scotton – 2022)

Facial expressions recognition is an ability to recognize people by their facial characteristic and differentiate it with one another. Human is born with the ability to recognize other people easily by identifying their facial features such as shape, appearance, skin texture and skin complexion. Other than that, humans also have the ability to express, interpret and differentiate facial expressions. The regular recur-ring ones are happiness, anger, disgust, fear, surprise and sad (Ekman & Friesen, 1978). The six facial emotions stated above are important and play a major role in expressing emotion as well as recognising facial expression (Busso, et al, 2004).

In real life, inter personal human interaction are performed not only using speech or spoken language, but also nonverbal cues for example hand gesture, body gesture, facial expression and tone of the voice. All these cues are sometimes being used for expressing feeling and give feedback (Busso, et al, 2004; Cohen, et Al., 2000). We can see how human interact with each other using non-verbal cues every day. For example a child cries in front of his mother because he is not happy or dissatisfied with something. Other people might interpret it differently thinking that the child might be in pain.

Facial expression interaction is relevant mainly for community social life, teacher and student interaction, credibility in difference contexts, medicine and so on. Besides, facial expression recognition is useful for designing new interactive devices which offers the possibility of new ways for human computer interaction - HCI (Franco & Treves, 2001). Cohen, et al. (2000) conducted survey on their users and noticed that they have been through traditionally HCI consists of the keyboard, mouse, joystick, trackballs, data gloves and touch screen monitors.

Facial Expression Recognition System (FER) has been a topic for research since Ekman and Friesen (1978) who pioneered this research and worked from the psychology perspective. In the past 20 years, many researchers have tried to adopt their idea and make improvement, innovation and modification on facial expression recognition by introducing different techniques, mainly concentrated on the improvement in term of accuracy, efficiency, mobility, and speed (Kotsia & Pitas, 2007). With all the enhancements on techniques for facial detection and recognition, the development of the facial expression recognition has also improved (Zhan & Zhou, 2007). The most active researches in computer vision and pattern recognition is face recognition in forensic identification, access control, user interface design (Wang, Plataniotis & Venetsanopoulos, 2005), emotion analysis, interactive video, indexing and retrieval of image and video database, image understanding and synthetic face animation (Zhan & Zhou, 2007).

Human can interpret and generate major facial expressions but a computer is not built with any facial recognition ability unless through the use of some software. It is even more complicated for the computer to interpret irregular facial expression, especially from those suffering from cerebral palsy. Due to their disorder, they do not have the ability to reflect their emotions like a normal typical person. Thus, a more natural and naive method has to be employed for the system to work by a manual labelling of the image captured with the emotion of the user.

4. Machine Learning based techniques in data analysis (Lavanya Vemulapalli, Dr.P.Chandra Sekhar – 2018)

A lot more applications available for us in play store, app store, amazon, etc., which are dependent machine learning. There are significant number of organizations and startups which turn towards optimum machine learning, and have proved that investing in machine learning is the best in today's world.

Google Street View - It is a pervasive city imagery datasets application.

It is an application from which we can virtually explore streets of cities. It uses a dense geosampling tool to shows the streets of cities. Streets are captured through a fleet of vehicles equipped with a specialized camera.

After collection of photos, they are digitally processed and combined together and looks like a single image. From files reported for privacy, Google pixelated faces of pedestrian and license plate which is captured. Web mapping technologies have been embraced by discipline such as geography, archeology and ecology, but also by several social scientific disciplines. Researchers working in the discipline of geography, archeology, and ecology quickly incorporated webbased mapping technologies into their research designs. There are various applications of google street view in research field, although the number still remains limited. It is also used for better estimation of fish catching, estimation of forestry biomass in India, estimation of area of different regions or lakes, etc.

Google also helps in the criminological studies that have implemented in the google maps and streets view in their research design. Public and some law enforcement agencies and offenders are familiar with the power of online mapping technology through their day-to-day life and work. The social sciences have also embraced web-mapping technologies. But to google maps still remain limited in social science.We can see google maps and its street view can be used in various fields. It can be used in mapping or developing or maintaining cities' streets. We can use Google's street view to make an infrastructure of building or apartment, park, bridges, water

reservoir, etc. Google maps and Google street view can be used in some research field of detecting the population or urbanization in some areas or throughout the globe.

Uber - Uber is one of the examples of using machine learning. It uses an algorithm which provides estimated time and real-time location on map, which is very useful and helpful for both drivers and riders. The company is also dealing with fraudulent behavior like face detection and invalid stolen credit cards.

Google Keyboard - Almost all android handset uses google keyboard. Gboard uses the neural spatial model to determine the pixels touched on the screen and making relevant words and emoji in handwriting mode. It predicts the next word by matching the currently typed word with its dictionary set, which helps user to type fast and accurate.

Snapchat - Snapchat uses machine learning to identify or face detection technology for applying filters on it. One may wonder about how Snapchat filter works? It first detects a face. Then locate facial features, and then create a mesh of 3D mask (pyramidal shape) over face. Snapchat not only apply filters but also a list of things they are doing like, language detection for very short texts, names entity recognition and disambiguation using multimodal NER (sound, text, etc.), normalizing text misspellings (phonetic, orthographic, semantic representations), emotion analysis (from emoji to actual pictures), speech, music recognition (keyword spotting), personalized neural conversational models. We can use this technology for detection of culprit's face if he or /she made some facial changes.

Virtual Voice Assistant - The world moves in the path of automation. People want their lives easy and comfortable like this hand free service provided by voice assistants. There are lots of virtual assistants available like google assistant, Apple's Siri, Cortana by Microsoft, Alexa by Amazon, Samsung's S voice, etc. As there more advances in machine learning voice assistants become more emotionally attached to human beings. Voice assistants remind us on time so that we do not skip some important stuff. Voice assistants along with computer vision can do many things that we even can't expect. It can do almost 70% of our daily work, from morning tea to evening supper.

Evernote - Evernote uses machine learning which automatically identifies the document file from device storage and applies filter on it, such that it appears clear and tidy.

5. Survey on Machine Learning Algorithm's

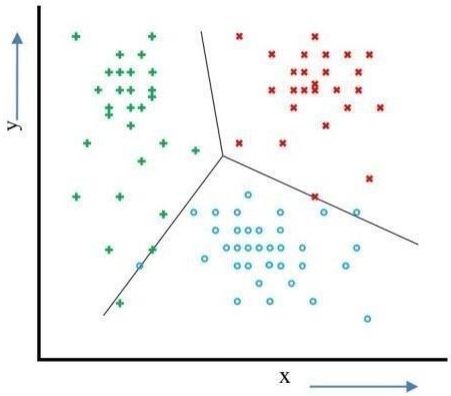
(Rekha Nagar, Dr. Yudhvir Singh – 2022)

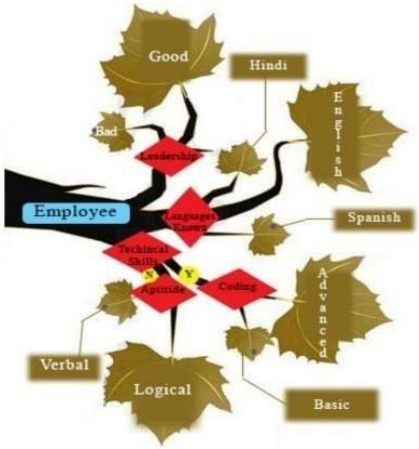
The subfield of artificial intelligence, machine learning has gained muchpopularity in last few couple of years. Many tech giants use machine learning algorithms, like Netflix's algorithms to make movie prediction from your previous watched movies. In this section, we would like to present some of the famous algorithms which use frequently.

They are:

- **Naïve-Bayes' algorithm -** This is the algorithm mostly used in machines and hardware. It simply applies Bayes' theorem along with strong independence assumptions. Let's take an example, to mark an email as spam, used for face detection software, etc.
- **K-means clustering algorithm -** This is a type of unsupervised learning which has various uses including business and management. This algorithm also lets us know profit at each stage of the product. It is also referred as Lloyd's algorithm. This algorithm is also used in grouping of features into different labels.

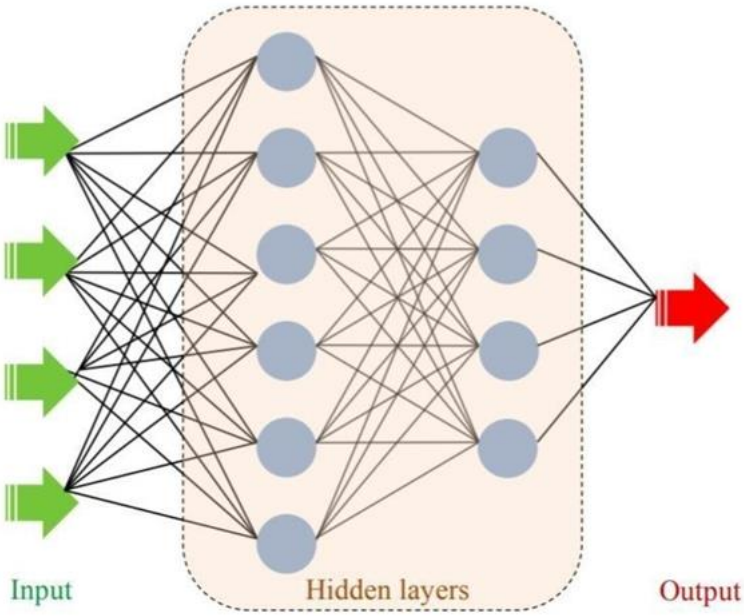
  Decision Trees - These are trees in which decisions are made by the computer at each stage based upon recurrence relations.





- **Neural Network -** Our neurons in body play a major role in determining the steps to process a single task. Similarly, artificial neurons are those which help the nervous

System of transistors in any sequential or combinational circuit to take up a decision and execute it conditionally. This again depends on activity of the neurons. An artificial neuron is an actual piece of hardware machinery which helps the system to take up a decision based on the receptors, as such several optoelectronic devices have already been developed. This algorithm helps us to build any machine functioning exactly as human reflex arcs.



Algorithms used in machines have several important implementations. We also have regression of value as well as regression trees, which help us to do different useful kind of job. The algorithms are also useful in health care industries, for example, random forest distribution algorithm, this algorithm is mostly derived from statistical studies… they are useful in calculating people densities and mass or chunk density. The most important is the artificial neural networks algorithm. This algorithm is related to artificial intelligence and neural networking. Though for mass application we must have machine learning. Through computer vision these algorithm judge systems on basis of their reactance to external stimuli.

6. Survey on application of Artificial Intelligence in Cyber Security (Shidawa Baba Atiku, Achi Unimke Aaron, Fatima Shittu – 2020)

Cyber security refers to protecting your personal computer from malicious software. Machine learning has a lot many algorithms and systems which protect users from threats. Such as the Paypal app which was developed in December 1998, uses machine learning algorithms to protect its users from different threats and online spoofing. It uses three types of machine learning algorithms that are linear, neural network, and deep learning algorithm.

They are:

- **Waterhole -** It is like a pit surrounded by greenery. Hackers access other people's information by using sites which are more accessible to the public more than anything else…. for example, networks in a coffee shop is accessed by so many users such that these users load their pc 's with whatsoever data is provided to them. Like this there ae so many sites to put on viruses and worms. Machine learning has algorithms that detect path of these malware blocking them with a firewall thereafter.

- **Webshell -** These are piece of code which is loaded into a working device which provokes the user to misjudge and then taking advantage, entry is gained into the full database.

- **Ransomware -** Similar to webshell, but here the user is vulnerably threatened externally by a group of software brokers who have corrupted the users' personal files. Such scenarios can be totally avoided by using machine level language which was early detection.

III. CONCLUSION:

From the above literature survey, we can conclude that all those papers follow a more or less similar methodology. We also like to follow that methodology with some improvements to overcome some of the limitations mentioned above. The input image is processed to isolate the hand. Then it is passed to a trained convolution neural network to identify the gesture with greater accuracy.

IV. REFERENCES:

1. Based Real Time Communication for Physically and Speech Disabled People (Ong Chin Ann, Marlene Valeriu Lu – 2019)
2. systematic review of computer vision semantic analysis in medical (Antonio Victor Alencar Lundgren, Byron Leite Dantas Bezzerra – 2021)
3. A survey on Facial Emotion Recognition Techniques (Felipe Zago Canal, Tobias Rossi Muller, Gustavo Gino Scotton – 2022)
4. Machine Learning based techniques in data analysis (Lavanya Vemulapalli, Dr.P.Chandra Sekhar – 2018)
5. Survey on Machine Learning Algorithm's (Rekha Nagar, Dr. Yudhvir Singh – 2022)
6. Survey on application of Artificial Intelligence in Cyber Security (Shidawa Baba Atiku, Achi Unimke Aaron, Fatima Shittu – 2020)