Journal of Positive School Psychology http://journalppw.com
2022, Vol. 6, No. 7, 4953-4960

Hand Gesture Recognition And Voice Conversion For Deaf And Dumb

Mr. Dr N.V Ganapathi Raju[1] , A. Rajender Reddy[2] , P. Sai Pranith[3] , L. Sai Nikhil[4] , Sarwar Pasha[5] , P. Vishwanath Bhat[6]

[1]*Professor, MTech (Ph.d), Department of IT, nvgraju@griet.ac.in*
[2]*BTech, Department of IT, annamrajenderreddy@gmail.com*
[3]*BTech, Department of IT, patibandlasaipranith@gmail.com*
[4]*BTech, Department of IT, nikhillokam173@gmail.com*
[5]*BTech, Department of IT, sarwarpasha22@gmail.com*
[6]*BTech, Department of IT, pvishwanathbhat@gmail.com*

ABSTRACT: Communication is main method about inter-person communication. Number about deaf & dumb persons has dramatically increased in recent years due to birth defects, mishaps, & mouth diseases. Individuals who are deaf or dumb must use a visual medium to communicate among others because they are unable to do so among hearing people. Many different languages are used & translated all around world. People who have trouble hearing & speaking are referred to as "Special Persons." other person has trouble understanding what "The Dumb" & "The Deaf" persons are trying to say, respectively. In order for dumplings to communicate among regular people, sign language is essential. Speaking among others who aren't silent is really difficult for silent people. because public is not taught hand sign language. They find it quite challenging to communicate in an emergency. solution to this problem is to translate sign language into audible speech. There are numerous efficient techniques for spotting hand motions or gestures, such as voice-to-text conversion using CNN & SVM algorithms. Here proposed study utilised SVM technique, however Python SVM is not accurate in identifying hand gesture, therefore we are using deep learning Convolution Neural Network to train hand gesture photographs, & then this trained model can be used to predict those taught hand gesture from webcam.

Keywords: Sign language, Hand gesture, Feature extraction, Gesture recognition.

1. INTRODUCTION

Modern society is continuously discussing new technology that improves our way about life & makes our lives easier. way people live has been dramatically altered through technology. Technology has a gear that human race has assigned it, & they are not in mood to change that gear. There is a lot about research being done in several technology-related sectors, such as artificial intelligence, cellphones, & others. New insights from this research made everyone's lives easier. However, very little research has been done on Deaf & Dumb persons. This sector has gotten less attention than others, in my opinion. One about biggest difficulties for this exceptional person is communication gap between them & average person. Deaf & blind people frequently experience communication difficulties. They are uneasy due to widespread discrimination & strife in society. Due to their perceived inability to communicate, people who are deaf or dumb find it difficult to express their emotions. HGRVC (Hand Gesture Recognition & Voice Conversion) technology finds & tracks hand motions about dumb & deaf in order to maintain connection

among general public. Hand gesture detection is achievable among a web camera. After that, pre-processing is used to make photos standard size.

Sign language is most widely used form about communication for persons who cannot speak. It's a language where hand gestures are used to convey letters & words. Sign language users have been noted in China alone. It reaches 80 million people in total, & communication will always be challenging, especially for those who are unable to understand sign language. use about visual approach for sign identification has grown in prominence over past few decades. It is a system that collects information using finger motions & detects it among help about a camera. It is most widely used visual method. It has taken a lot about effort to develop vision-based sign recognition systems on a global basis. Systems for recognising gestures using vision can employ both direct & indirect methods. Previously, a method based on eyesight was used to identify hand movements. recognised image, however, is significantly impacted through this process on environment. Another disadvantage is that they have to show their hands to camera. Here, a flex sensor detects hand motion & converts it to voice.

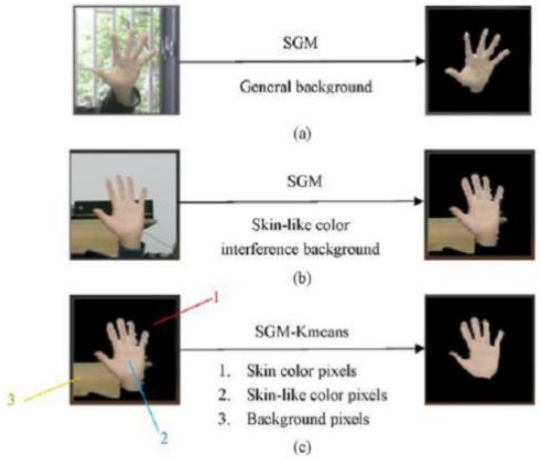

Fig.1: (a) SGM segmented image in general background. (b) SGM segmented image in skin-like background. (c) SGM-Kmeans segmented image in skin-like background.

This sector has gotten less attention than others, in my opinion. One about biggest difficulties for this exceptional person is communication gap between them & average person. Speaking among regular people might be challenging for Deaf & Dumb people. They are uneasy due to widespread discrimination & strife in society. Due to their perceived inability to communicate, people who are deaf or dumb find it difficult to express their emotions. HGRVC (Hand Gesture Recognition & Voice Conversion) technology finds & tracks hand motions about dumb & deaf in order to maintain a line about communication among other persons. Hand gesture detection is achievable among a web camera. photos are then resized via pre-processing to their original sizes. This project aims to develop a system that can convert hand gestures into text. project's objective is to use database matching to convert photos to text & place them into dataset. As part about detection process, hands are observed in motion. method generates text output, reducing communication gap between humans & deaf-mutes.

2. PROPOSED SYSTEM

Here about this research is constructing a machine learning model that can predict hand gesture from a camera & then turn recognised gesture into voice so that non-Deaf & non-Dumb people may understand what Deaf & Dumb people are saying. We are using a deep learning Convolution Neural Network to train hand gesture images, & we are using that trained model to predict those learnt hand motions from webcam. we used SVM technique in suggested investigation, although Python SVM is not reliable for distinguishing hand motion.

This project aims to develop a system that can convert hand gestures into text. project's objective

is to add photographs to database, which will match them & convert them to text. As part about detection process, hands are observed in motion. method generates text output, reducing communication gap between humans & deaf-mutes.

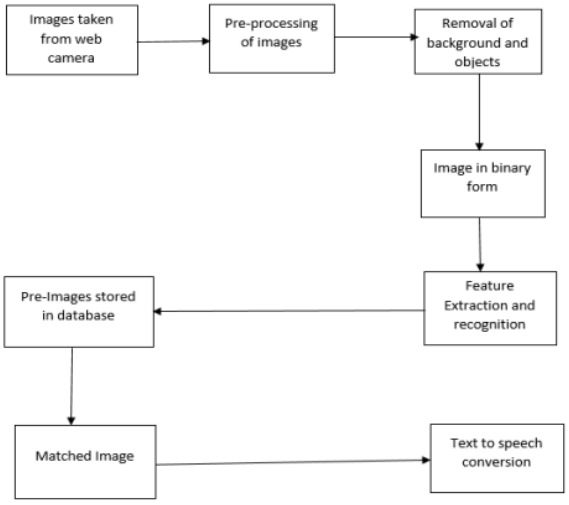

Fig.2: System architecture

Here, we will discuss process about implementing our system, which is depicted as a flowchart in Figure 2.

2.1 Training about System:

The user must enter desired number in order to store samples in database. More samples should be used to get greater accuracy than 5. user must select folder where photos will be saved. Click start video to start web camera & start database creation process. Click capture image to add specified number about photos to training folder for each sample. When number about images displayed matches number about successfully taken images, database construction is complete.

2.2 Image Pre-Processing:

The acquired images are pre-processed to enhance their intrinsic features. Pre-processing basically involves removing foreground & backdrop about an image to focus solely on hand gestures. preprocessed image is then shown as a series about binaryized (all-black & all-white) pixels.

2.3 Feature Extraction & Recognition:

Feature extraction is part about dimensionality reduction process, which reduces size & complexity about a starting collection about raw data into manageable chunks. As a result, processing will be less complicated for you.

MODULES:

1. Upload Hand Gesture Dataset
2. Preprocess dataset
3. Model Generation
4. Train CNN Gesture Images
5. Sign Language Recognition from Webcam
6. Extract image from webcam
7. Convert image to binary or grey format & back ground removal
8. Extract features from image
9. Recognition & play audio

3. RELATED WORK

3.1 Two Way Communicator between Deaf & Dumb People & Normal People.

One about nature's most priceless gifts to mankind is ability to express oneself through responding to world around him. average individual watches, listens, & then speaks in response to situation. Some poor people, however, are not permitted to receive this wonderful gift. This deepens gap between rich & people. This software makes it easier for them to communicate among one another. Two components make up majority about system. Indian Sign Language (ISL) gestures are translated into speech that can be comprehended through humans in first module. As a result,

utilising as its input, second module will convert animated movements in Indian Sign English into natural language. processing from video to voice method will include image mapping among linguistic knowledge bases, area about interest (ROI) identification, & frame generation from videos. After that, pertinent audio will be produced via Google Text-to-Speech(TTS) API. Natural language is mapped among analogous Indian Sign Language movements utilising Google Speech-to-Text(STT) API to convert speech to text. relevant animated motions are then further mapped to text from database.

3.2 Orientation sensing for gesture-based interaction among smart artefacts:

For implementation about embedded technology augmented artefacts, also known as "smart artefacts," that exhibit embodied modes about interaction based on their position, orientation, & pertinent dynamics, orientation sensing is recognised as a key instrument. When considering artefacts subject to manual (or "by-hand"") manipulation through user, we distinguish hand worn, hand carried, & (hand) graspable real world objects as displaying distinct artefact orientation dynamics, necessitating an analysis along these three categories. We provide a general framework for orientation sensor-based gesture identification and, in an abstract sense, call orientation dynamics "gestures.. framework description describes a group about gestures that are independent about classification methods & sensor technologies & are not particular to any one application. It enables interoperability between various sensors & supports a configurable number about sensors. Among framework's most important components is a library about gestures, which is broken down into three groups: hand gestures, artefact gestures held continuously, & artefact gestures periodically taken from hand. Using an inertial orientation sensing-based system for gesture detection & recognition, a framework for creating gesture-based interactions is developed. development about physical, software-based, & hardware-based media player remote controls serves as an example about how this paradigm is used.

3.3 Automated speech recognition approach to continuous cue symbols generation:

The study covered in this paper intends to develop a system that aids people who are deaf-dumb through converting spoken language into sign language. American Sign Language is translated into spoken signals through this technique. A portable computer's monitor plays a prerecorded American sign language (ASL) demonstration about sign as words that go among it are read from lexicon about American sign language. If there is no sign for word in sign language, it is finger spelled. In actual world, deaf people use this technique for things like proper names that lack obvious signs. Hidden Markov Model (HMM) is used to convert user's spoken signal into cue symbols for those among vocal impairments. proposed job is an extension about previous study on "Boltay Haath," a speech signal used to recognise a vocally disabled person's finger movement. When used in conjunction among Boltay Haath system, proposed AISR system might bridge communication gap between people without speech issues & typical person.

3.4 Finger Detection for Sign Language Recognition:

Computer recognition about sign language is a key scientific problem for enhancing communication among hearing-impaired people. number about fingers spread out in an American Sign Language gesture that represents an alphabet is counted quickly & accurately in this research. concept about finger detection is based on boundary tracing & finger tip detection. hand does not need to be correctly aligned to camera, nor does system require use about any specific

markings or input gloves on hand. Index Terms: Boundary tracing, sign language recognition, image processing, finger detection, & computer accessibility for disabled.

3.5 Recognition about Arm Gestures Using Multiple Orientation Sensors: Gesture Classification:

We present a method for identifying gestures using Euler angles gleaned from a variety about orientation sensors. This algorithm is used in a system for controlling unmanned aerial vehicles (UAVs) while manned aircraft are present on an aircraft deck. We investigate various techniques for arm gesture recognition before analysing a real-time arm gesture identification system using InterSense's IS-300 Pro Precision Motion Tracker. In our study, we characterise low-level gestures, analyse alternative gesture detection methodologies before selecting an active sensor, model gestures using Euler angles, & create model-based algorithms for categorising gestures. We tested proposed real-time arm gesture detection system in a lab environment among a robot that represents a UAV.

4. IMPLEMENTATION

• Here proposed study utilised SVM technique, however Python SVM is not accurate in identifying hand gesture, therefore we are using deep learning Convolution Neural Network to train hand gesture photographs, & then this trained model can be used to predict those taught hand gesture from webcam.

CNN ALGORITHM:

Convolutional neural networks, often known as CNNs or ConvNets, are particularly skilled at processing input among a grid-like design, such as images. A digital image is a binary representation about visual data.

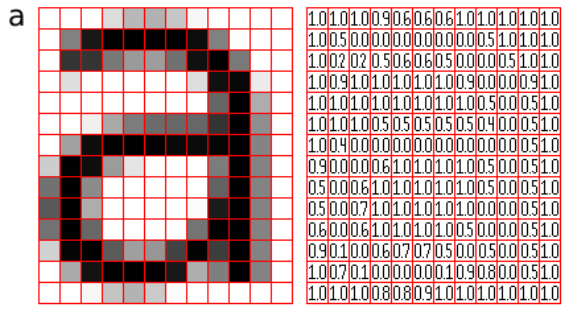

Fig.3: Representation about image as a grid about pixels

The human brain starts processing a huge amount about data as soon as we perceive a picture. Each neuron has a unique receptive field, & because they are connected to one another, they collectively encompass whole visual field. Similar to how each neuron in biological vision system responds to stimuli only in confined area about visual field known as receptive field, each neuron in a CNN processes data only in its receptive field. layers first pick up on lines, curves, & other simpler patterns before moving on to more complex patterns like faces & objects. Using a CNN, one can give computers sight.

Convolutional Neural Network Architecture:

The conventional three layers about a CNN are fully connected, pooling, & convolutional layers.

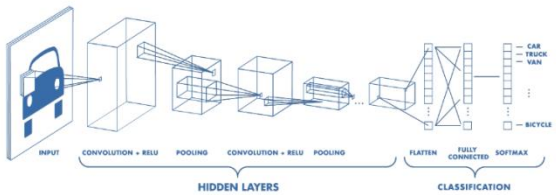

Fig.4: CNN architecture

Convolution Layer: convolution layer is a key component about CNN. It carries majority about network's computational load. kernel—a group

about learnable parameters—and limited region about receptive field are two matrices that are combined in this layer to form a dot product.

Pooling Layer: At specific points, pooling layer takes place about network's output through getting a summary statistic from nearby outputs. This helps to reduce spatial size about representation, which minimises amount about computation & weights required. pooling procedure is applied to each slice about representation independently.

Similar to a regular FCNN, fully connected layer has complete connections between every neuron in it & every neuron in layer above & below it. As a result, it can be estimated using standard matrix multiplication & bias effect. among help about FC layer, representation between input & output is mapped.

5. EXPERIMENTAL RESULTS

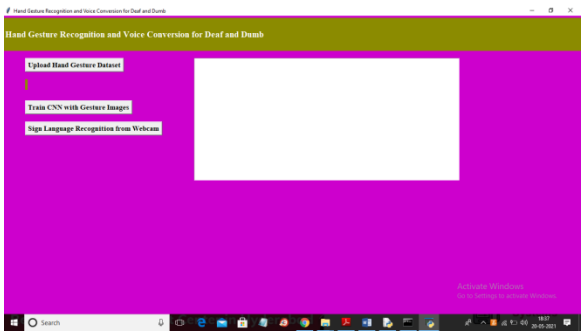

Fig.5: Home screen

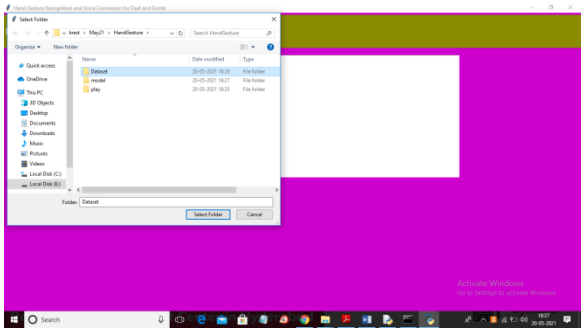

Fig.6: Upload hand gesture dataset

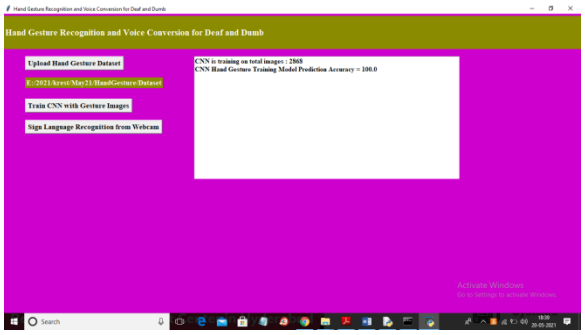

Fig.7: Train CNN gesture images

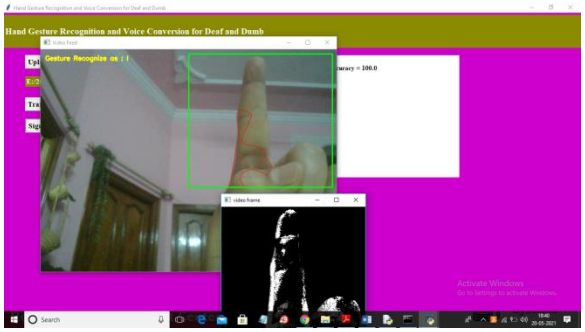

Fig.8: Sign language recognition from webcam

Here, all you have to do is act out motion as it appears on screen above. If your hands are adjusted, forecast might be off; however, if your gesture is fixed, it will be on money. When project is run, modules listed below are executed for each forecast.

• Webcam image extraction, binary or grayscale conversion, & background removal

• Play music & recognise & extract visual features

6. CONCLUSION

For dumb & deaf persons, image processing has been used to translate voices & recognise hand movements. method takes an image as input & outputs text & audio. implementation about this system offers accuracy about up to 90% & performs well in most test situations. objective about this project is to build a machine learning model that can forecast hand motions from

webcam footage & then convert recognised hand gestures into voice, enabling hearing & hearing-impaired people to converse among regular people. finished product is shown as static text after captured image has been retrieved from image dataset. A hand motion is transformed into a picture through feature extraction & categorization.

REFERENCES

[1] Shinde, Shweta S., Rajesh M. Autee, & Vitthal K. Bhosale. "Real time two way communication approach for hearing impaired & dumb person based on image processing." Computational Intelligence & Computing Research (ICCIC), 2016 IEEE International Conference on. IEEE, 2016.

[2] Shangeetha, R. K., V. Valliammai, & S. Padmavathi. "Computer vision based approach for Indian Sign Language character recognition." Machine Vision & Image Processing (MVIP), 2012 International Conference on. IEEE, 2012.

[3] Sood, Anchal, & Anju Mishra. "AAWAAZ: A communication system for deaf & dumb." Reliability, Infocom Technologies & Optimization (Trends & Future Directions)(ICRITO), 2016 5th International Conference on. IEEE, 2016.

[4] Ahire, Prashant G., et al. "Two Way Communicator between Deaf & Dumb People & Normal People." Computing Communication Control & Automation (ICCUBEA), 2015 International Conference on. IEEE, 2015.

[5] Ms R. Vinitha & Ms A. Theerthana. "Design & Development about Hand GestureRecognition System For Speech Impaired People."

[6] Kumari, Sonal, & Suman K. Mitra. "Human action recognition using DFT." Computer Vision, Pattern Recognition, Image Processing & Graphics (NCVPRIPG), 2011 Third National Conference on. IEEE, 2011.

[7] S. F. Ahmed, S. Muhammad, B. Ali, S. Saqib, & M. Qureshi, "Electronic Speaking Glove for Speechless Patients A Tongue to," no. November, pp. 56-60, 2010.

[8] A. Y. Satpute, A. D. Bhoi, & T. Engineering, "ELECTRONIC SPEAKING SYSTEM FOR DUMB," vo!. 6, no. 3, pp. 1132-1139, 2013.

[9] M. Wald, "Captioning for Deaf & Hard about Hearing People through Editing Automatic Speech Recognition in Real Time", Proceedings about 10th International Conference on Computers Helping People among Special Needs ICCHP 2006, LNCS 4061, pp. 683- 690.

[10] R. R. Itkarkar & A. V. Nandi, "Hand gesture to speech conversion using Matlab," in 2013 Fourth International Conference on Computing, Communications & Networking Technologies (ICCCNT), 2013, pp. 1-4.

[11] Jingdong Zhao, Li Jiang, Shicai Shi, Hegao Cai, Hong Liu, G. Hirzinger, "A Five-fingered Underactuated Prosthetic Hand System",Proceedings about 2006

[12] Harish, N. & Poonguzhali, S., 2015, May. Design & development about hand gesture recognition system for speech impaired people. In 2015 International Conference on Industrial Instrumentation & Control (ICIC) (pp. 1129-1133). IEEE.

[13] Vishal, D., Aishwarya, H.M., Nishkala, K., Royan, B.T. & Ramesh, T.K., 2017, December. Sign Language to Speech Conversion. In 2017 IEEE International Conference on Computational Intelligence & Computing Research (ICCIC)(pp. 1-4). IEEE.

[14] Vijayalakshmi, P. & Aarthi, M., 2016, April. Sign language to speech conversion. In 2016

International Conference on Recent Trends in Information Technology (ICRTIT) (pp. 1-6). IEEE.

[15] Heera, S.Y., Murthy, M.K., Sravanti, V.S. & Salvi, S., 2017, February. Talking hands—An Indian sign language to speech translating gloves. In 2017 International conference on innovative mechanisms for industry applications (ICIMIA) (pp. 746-751). IEEE.

[16] Bhaskaran, K.A., Nair, A.G., Ram, K.D., Ananthanarayanan, K. & Vardhan, H.N., 2016, December. Smart gloves for hand gesture recognition: Sign language to speech conversion system. In 2016 International Conference on Robotics & Automation for Humanitarian Applications (RAHA) (pp. 1- 6). IEEE.

[17] Jean, C., & Peter, B., "Recognition about Arm Gestures Using Multiple Orientation Sensors: Gesture Classification", IEEE Intelligent transportation systems conference on electronics, Vol. 13, No. 1, pp. 334-520,2004.

[18]. Otiniano, R., & Amara, C. "Finger spelling recognition from rgb-d information using kernel descriptor", IEEE Transactions on neural systems & rehabilitation engineering, Vol.28, No. 8, pp. 124- 184, 2006.

[19]. Zhengmao, Z., Prashan, R., Monaragala, N., Malin, P. "Dynamic hand gesture recognition system using moment invariants", IEEE Transactions on neural networking & computing, Vol. 21, No. 1, pp. 1034- 1320, 2010.

[20]. Recognition about sign language gestures using neural networks through Peter Vamplew. Department about Computer Science, University about Tasmania [2]. Fels & G Hinton (1993), Glove-Talk: A Neural Network Interface Between a Data-Glove & a Speech Synthesiser, IEEE Transactions on Neural Networks, 4, 1, pp. 2-8 [3]. Johnston (1989), Auslan: Sign Lang