JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Speech Impaired Support Application

[1]Shreeyukta, [2]Sudiksha Khatri, [3]Prof. Ajay K. Srivastava

[1]Student, [2]Student, [3]Head of Department
Information Technology Department, Bachelor of Technology
Shri Ramswaroop Memorial College of Engineering and Management, Lucknow, India

Abstract: Sign Language is the method of communication of deaf and dumb people all over the world. However, it has always been a difficulty in communication between a verbal impaired person and a normal person. Sign Language Recognition is a breakthrough for helping deaf-mute people to communicate with others. The commercialization of an economical and accurate recognition system is today's concern of researchers all over the world. Thus, sign language recognition systems based on Image processing and neural networks are preferred over gadget system as they are more accurate and easier to make.

Index Terms **– Speech Impaired Support Application, Machine Learning, CS Project.**

INTRODUCTION

Sign Language is the most natural and expressive way for the hearing impaired people. People, who are not deaf, never try to learn the sign language for interacting with the deaf people. This leads to isolation of the deaf people. But if the computer can be programmed in such a way that it can translate sign language to text format, the difference between the normal people and the deaf community can be minimized. Indian sign language (ISL) uses both hands to represent each alphabet and gesture. ISL alphabets are derived from British Sign Language (BSL) and French Sign Language (FSL). Most of the researchers in this area concentrate on the recognition of American Sign Language (ASL) since most of the signs in ASL are single handed and thus, complexity is less. Another attractive feature is that ASL already has a standard database that is available for use. When compared with ASL, Indian Sign Language relies on both hands and thus, an ISL recognition system is more complex. A few research works carried out by the researchers in the recognition of ISL. Currently, more researchers have started doing research in ISL. Here this proposed system is able to recognize the various alphabets of Indian Sign Language; this will reduce the noise and give accurate result.

The important research problem in computer recognition is the sign language for enabling communication with hearing impaired people. This system introduces efficient and fast techniques for identification of the hand gesture representing an alphabet of the Sign Language. Currently, more interest is created to do research in the field of sign language recognition system. Deaf and Dumb people rely on sign language interpreters for communications. A real time Sign Language Recognition system was designed and implemented to recognize 26 gestures from the Indian Sign Language by hand gesture recognition system for text generation. The signs are captured by using web cam. This signs are processed for feature extraction using some colour model. The extracted features are compared by using pattern matching algorithm. In order to calculate the sign recognition, the features are compared with testing database. Finally, recognized gesture is converted into text. This system provides an opportunity for a deaf-dumb people to communicate with non-signing people without the need of an interpreter.

Our system aims to get the deaf and dumb people more involved to communicate and the idea of a camera-based sign language recognition system that would be in use for converting sign language gestures to text and audio. Our objective is to design a solution that is intuitive and simple which simplifies the communication for the majority of people with deaf and dumb.

Problem Definition

Sign language is an incredible advancement that has grown over the years. Unfortunately, there are some drawbacks that have come along with this language. Not everyone knows how to interpret a sign language when having a conversation with a deaf and dumb person. There is always a need to communicate using sign language. One finds it hard to communicate without an interpreter.

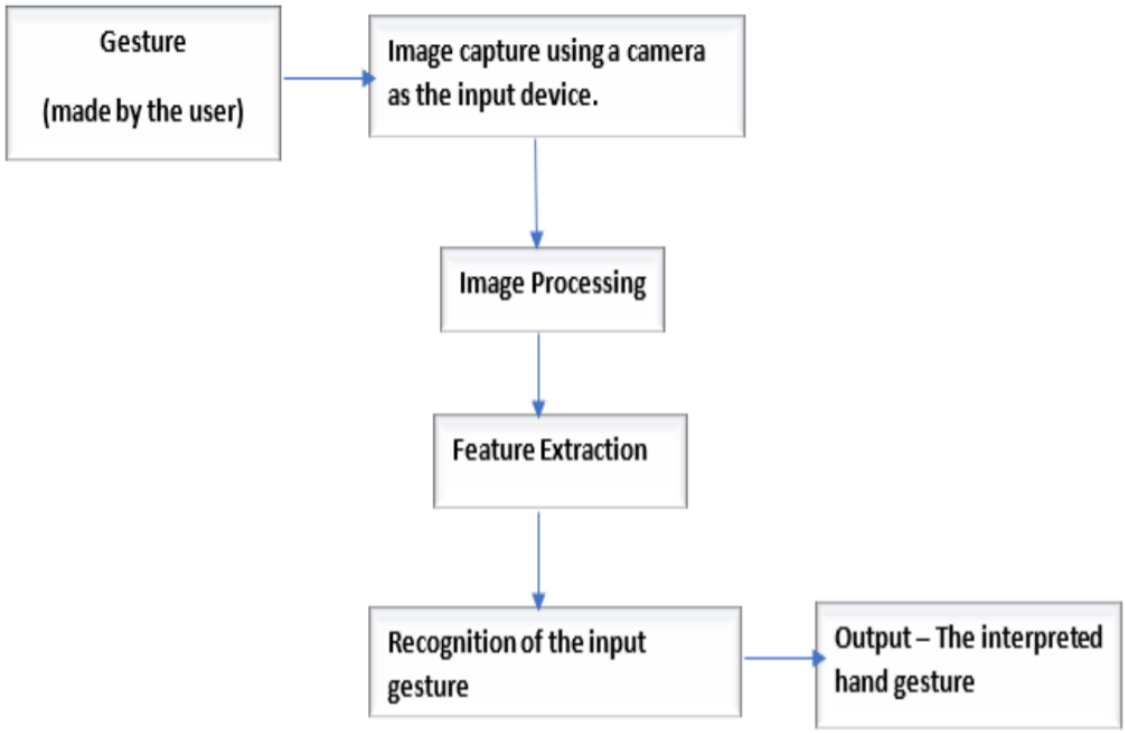
To solve this, we need a product that is versatile and robust. We need to convert the sign language so that it is understood by common people and will help them to communicate without any barriers. The main purpose of this project is to eliminate the barrier between the deaf and dumb and the rest.

Proposed Objective

The main objective is to translate sign language to text/speech. The framework provides a helping-hand for speech-impaired to communicate with the rest of the world using sign language. This leads to the elimination of the middle person who generally acts as a medium of translation. This would contain a user-friendly environment for the user by providing speech/text output for a sign gesture input.

Proposed Methodology

The most basic explanation of workflow of the system goes as follows - A hand gesture is performed in front of the webcam. This sign gesture is converted to text and the text output is converted to audio and is served as an input to the assistant. The assistant processes the question and responds in audio format. This audio format is converted to text output. The text output will be then displayed on the display screen result will be spoken.



METHODOLOGY USED

Tensor Flow

TensorFlow in our system helps us in training the model using the provided dataset. TensorFlow object recognition algorithms help us classify and identify different hand gestures when combined with use of OpenCV. By analyzing thousands of photos, TensorFlow can help classifying and identifying real-time hand gestures. It makes possible to develop a model which can help identify 3D images and classify it on basis of 2D images from its feed dataset. TensorFlow is capable of processing more information and spot more patterns.

Deep Learning

Deep Learning is basically a subset of Machine Learning model which consists of algorithms that make use of multilayer neural networks. Deep Learning makes use of Neural Network most of the times to implement its functioning. A Neural Network is a collection of layers that transforms the input in some way to produce output.

Initially, the computer program might be provided with training data -- a set of images for which a human has labelled each image *sign* or *not sign* with metatags. The program uses the information it receives from the training data to create

a feature set for *sign* and build a predictive model. In this case, the model the computer first creates might predict that anything in an image that has four legs and a tail should be labelled *sign*. Of course, the program is not aware of the labels *four legs* or *tail*. It will simply look for patterns of pixels in the digital data. With each iteration, the predictive model becomes more complex and more accurate.

OpenCV

OpenCV is an open source library for Computer Vision. Now since all the training and classification is ready to be executed when it needed an eye for the designed system to capture real-time images of Hand Gestures which can then be sent for classification and identification. OpenCV adds intelligence to Deep Learning models for visualization image processing.

OpenCV is the huge open-source library for the computer vision, machine learning, and image processing and now it plays a major role in real-time operation which is very important in today's systems. By using it, one can process images and videos to identify objects, faces, or even handwriting of a human. When it integrated with various libraries, such as NumPy, python is capable of processing the OpenCV array structure for analysis. To Identify image pattern and its various features we use vector space and perform mathematical operations on these features.

ALGORITHM AND APPROACH

ALGORITHM USED

Algorithms used in this project are :-

Rule Based Classifier:

Rule-based classifiers are just another type of classifier which makes the class decision depending by using various "if-else" rules. These rules are easily interpretable and thus these classifiers are generally used to generate descriptive models. The condition used with "if" is called the **antecedent** and the predicted class of each rule is called the **consequent**.

Properties of rule-based classifiers:

Coverage:

- The percentage of records which satisfy the antecedent conditions of a particular rule.
- The rules generated by the rule-based classifiers are generally not mutually exclusive, i.e. many rules can cover the same record.
- The rules generated by the rule-based classifiers may not be exhaustive, i.e. there may be some records which are not covered by any of the rules.
- The decision boundaries created by them is linear, but these can be much more complex than the decision tree because the many rules are triggered for the same record.

Rule-Based Classifier classify records by using a collection of "if…then…" rules.

(Condition)→Class Label(Condition)→Class Label

(BloodType=Warm)∧(LayEggs=Yes)→Birds(TaxableIncome<50K)∨(Refund=Yes)→Evade=No(BloodType=Warm)∧(LayEggs=Yes)→Birds(TaxableIncome<50K)∨(Refund=Yes)→Evade=No

The Left Hand Side is rule antecedent or condition

The Right Hand Side is rule consequent

Coverage of a rule - Fraction of records that satisfy the antecedent of a rule

Accuracy of a rule - Fraction of records that satisfy both the antecedent and consequent of a rule

Background Subtraction Method:

Background Subtraction is a technique for separating out foreground elements from the background and is done by generating a foreground mask. This technique is used for detecting dynamically moving objects from static cameras. Background subtraction technique is important for object tracking. There are several techniques for background subtraction.

The running average of a function is used to separate foreground from background. In this concept, the video sequence is analysed over a particular set of frames. During this sequence of frames, the running average over the current frame and the previous frames is computed. This gives us the background model and any new object introduced in the during the sequencing of the video becomes the part of the foreground. Then, the current frame holds the newly introduced object with the background. Then the computation of the absolute difference between the background model (which is a function of time) and the current frame (which is newly introduced object) is done.

1. Data Set: In this step, a set of images in the sign language is fed to a database. The input obtained is then compared with the given images in the dataset to identify the gesture made. The reason for the number of images in the dataset is

to get the output with a good amount of accuracy and also to avoid ambiguity, which has high chances of occurring in sign language as one gesture might be similar to another one.

2. Image detection: This is the step that comes right after camera capture. Image detection refers to detecting the image that is obtained and, in this case, it is found out if the obtained image is that of a hand or not. A binary classifier is to be trained beforehand to check the same. A binary classifier has the task of classifying sets into two groups, depending on the criteria the sets meet. It checks for one or more qualities that a particular set should possess. It is according to that factor that a binary classifier decided to which group the set should be sent to.

3. Feature Extraction and Image Recognition: Feature extraction refers to extracting the details from the image captured. In a sign language interpreter, the image captured is a gesture made by a hand. Therefore, the features extracted from such images include the size of the palm, the number of fingers open, etc. These features are then used to recognize the gesture using certain algorithms

4. Output: All of these modules contribute to the successful working of sign language interpreting system. The flow of execution takes place in the following manner: The camera gets the input gesture image from the user, the detection process takes place to check if it is a hand or not using certain algorithms, image recognition is the next step where the image acquired from the user is compared with the images in the dataset, to interpret the shown gesture. The next step is the output where the recognized symbol is converted to text and audio form as the output.

RESULT ANALYSIS AND DISCUSSION

The requirement of machine-based sign language translator is very important in the present scenario. Even though we have found initial success in this regard, lot of work needs to be done.

- The main area where this can be used is in public places like ticket issuing counters, hospitals etc.
- This can be even used to teach the sign language to normal people.
- Further this can be used to take words and display the gesture for the same.
- Recognizing fingers will widen the training set for the machine.

How we split data in Machine Learning ?

Training Data:

Training data (or a training dataset) is the initial data used to train machine learning models. Training datasets are fed to machine learning algorithms to teach them how to make predictions or perform a desired task.

Validation Data:

The part of data that is used to do a frequent evaluation of the model, fit on the training dataset along with improving involved hyperparameters (initially set parameters before the model begins learning). This data plays its part when the model is actually training.

Testing Data:

Once our model is completely trained, testing data provides an unbiased evaluation. When we feed in the inputs of Testing data, our model will predict some values(without seeing actual output). After prediction, we evaluate our model by comparing it with the actual output present in the testing data. This is how we evaluate and see how much our model has learned from the experiences feed in as training data, set at the time of training.

APPLICATIONS AND LIMITATIONS

Applications:

- The proposed system can be used in classrooms and public places where anyone and everyone can interact with the deaf and the dumb.
- The proposed system converts sign language to written form.
- The proposed system converts sign language to audio.
- It is helpful for the deaf and the dumb to easily communicate with the rest of the world as not everyone is familiar with sign languages.

Limitations:

- There are various forms of Sign Languages globally, so dataset used needs to be changed according to the standard Sign Languages used in that nation or region.

- It still needs to make it better by means of it's Overall Accuracy and Productivity in terms of generating end results.

- Need of a plain background, white being the most favourable

- Presence of good amount of light while presenting the hand gestures

CONCLUSION

Sign languages are kinds of visual languages that employ movements of hands, body, and facial expression as a means of communication. Sign languages are important for specially-abled people to have a means of communication. Through it, they can communicate and express and share their feelings with others. The drawback is that not everyone possesses the knowledge of sign languages which limits communication. This limitation can be overcome by the use of automated Sign Language Recognition systems which will be able to easily translate the sign language gestures into commonly spoken language. In this paper, it has been done by TensorFlow object detection API. The system has been trained on the Indian Sign Language alphabet dataset. The system detects sign language in real-time. For data acquisition, images have been captured by a webcam using Python and OpenCV which makes the cost cheaper. The developed system is showing an average confidence rate of 85.45%. Though the system has achieved a high average confidence rate, the dataset it has been trained on is small in size and limited.

FUTURE SCOPE

The current designed system completely works on basis of underlying Dataset which is used to train the system, thus making its use limited to certain group of people which communicates using the similar Sign Language.
However it is found out that there are various forms of Sign Languages globally, so dataset used needs to be changed according to the standard Sign Languages used in that nation or region.
Since the current system's Graphical User Interface is simple but not visually soothing, further it is aimed at designing and building a better Visual Interface that makes it even more vibrant and eye-catching eventually making it easier as well interesting to use, for our targeted audience.
As the current system can still not be called a complete Error Free product, it still needs to make it better by means of it's overall Accuracy and Productivity in terms of generating end results.
Our system has few limitations such as need of a plain background, white being the most favourable for better results, also presence of good amount of light while presenting the hand gestures. So it is required to overcome these difficulties in order to make system perform better.

REFERENCES

[1] Oi Mean Foong, Tan Jung Low, and Satrio Wibowo, Hand Gesture Recognition: Sign to Voice System S2V. Proceedings Of World Academy Of Science, Engineering And Technology Volume 32 AUGUST 2008 ISSN 2070-3740
[2] Maebatake, M., Suzuki, I.; Nishida, M.; Horiuchi, Y.; Kuroiwa, S.; Sign Language Recognition Based on Position and Movement Using MultiStream HMM, Second International Symposium on, ISBN: 978-0- 7695-3433- 6.
[3] T.S. Huang, and V.I. Pavlovic, "Hand gesture modeling, analysis and synthesis", Proceedings of international workshop on automatic face and gesture recognition, 1995, pp.73-79.
[4] J. Rekha, J. Bhattacharya, and S. Majumder, Hand Gesture Recognition for Sign Language: A New Hybrid Approach, 15th IPCV 11, July 2011, Nevada, USA.
[5] A.W. Fitzgibbon, and J. Lockton, "Hand Gesture Recognition Using Computer Vision", BSc.Graduation Project,
[6] Tushar Chouhan, Ankit Panse, Smart Glove with Gesture Recognition Ability for The Hearing and Speech Impaired, Global Humanitarian Technolgy Conference, September 2014.
[7] T.H. Speeder, "Transformation human hand motion for Telemanipulation," Presence, Vol.1, no. 1, pp.63-79,1992.
[8] L.Bretznar & T.Linderberg, "Relative orientation from extended sequence of sparse pointand line correspondances using the affine trifocal sensor," in Proc. 5 Eur. Conf. Computer Vision, Berlin, Germany, June 1998, Vol.1406, Lecture Notes in Computer Science, pp. 141-157, Springer Verlag.
[9] D. Xu, "A neural network approach for hand gesture recognition in virtual reality driving training system of SPG," presented at the 18 Inc.Conf. Pattern recognition, 2006.
[10] D. K. Sarji, "HandTalk: assistive technology for the deaf,"Computer, vol. 41, pp. 84-86, 2008
[11] Emil M. Petriu, Qing Chen, Nicolas D. Georganas, Real-time vision-based hand gesture recognition using haar-like features, 2017.