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

A Comparative Approach an Alphabet Detection in Sign Language Using Machine Learning Algorithm"

Pranay Bante[1],Sarika Kamble[2] ,Nitikesh P.Raut[3], Shrinath L.Kashid,[4]S.S.Bhosale[5], C. G. Patil[6]

"Dept. of E&TC Department of E&TC,SKNCOE,SPPU,Pune

[1]pranaykbante123@gmail.com
[2]sarika.kamble_skncoe@sinhgad.edu
[3]nitikeshpr@gmail.com
[4]shrinathkashid@gmail.com
[5]sonali.bhosale_skncoe@sinhgad.edu
[6]chandrashekhar.patil_skncoe@sinhgad.edu

ABSTRACT

The only way speech and hearing-impaired people can communicate is by using the sign language. The main problem with this kind of communication is that the non-impaired people, who cannot understand the sign language, would not be able to communicate with these people or vice versa. The project is intentionally designed to allow deaf and dumb communities to convey messages and connect with the society. It aims to bridge the gap between speech and hearing impaired people and the non-impaired people. Speech impairment is a disability that affects an individual's ability to communicate effectively. Many existing studies have proposed methods for sign language recognition The project uses a machine learning system to identify, especially English alphabetic sign language, and convert it into text. The focus of this work is to create a vision based application that offers American Sign Language translation to text thus aiding communication between signers and non-signers. The proposed model takes image sequences and extracts temporal and spatial features from them.

Keywords- Machine learning Alphabet Detection, Sign recognition, Text-to-speech algorithms**.**

1.1 INTRODUCTION

Automatic conversion of sign language to text or speech is indeed helpful for interaction between deaf or mute people with people who even do not have knowledge of sign language. This is the demand of current times to develop an automatic system to convert ISL signs to normal text and vice versa. This will be beneficial for both communities to express their fillings to one another in accessing publicly available facilities like ticketing, banking services, traffic signals, etc. A new feature extraction and selection technique using structural features and some of the best available classifiers are proposed to recognize ISL signs for better communication for computer-human interface. This paper narrates a system for automatic recognition of ISL immobile numeric signs, in which a standard digital camera was only used to acquire the signs, no wearable devices are required to capture electrical signals. The system is intended to convert isolated digit signs into text, that is, each entered sign image should contain precisely one numeric sign.

which contains 840 images, 10 images for each numeral sign (0-9). To effectively achieve this, a sign language (ASL – American Sign Language) image to text as well as speech conversion was aimed at in this research. The techniques of image segmentation and feature detection played a crucial role in implementing this system. We formulate the interaction between image segmentation and object recognition in the framework of FAST and SURF algorithms. The pre-determination of the ROI of each image using SURF and FAST, has

demonstrated the ability of the proposed algorithm to limit image modelling to relevant region within the image. In recent years, there have been ongoing efforts to develop automated methods for the completion of numerous linguistic tasks using advanced algorithms that can 'learn' based on past experience [33]. Sign language recognition (SLR) is an area where automation can provide tangible benefits and improve the quality of life for a significant number of people who rely on sign language to communicate on a daily basis [34]. The successful introduction of such capabilities would allow for the creation of a wide array of specialized services, but it is paramount that automated SLR tools are sufficiently accurate to avoid creating confusing or dysfunctional responses. In this section, we provide a brief background regarding some important approaches that have been utilized for automated SLR. The machine learning concept encompasses a number of stochastic procedures that can be used to predict the value of a certain parameter based on similar examples that the algorithm was previously exposed to. A simple example, illustrated by Algorithm 1, shows how a general formalization of the learning process takes place. There are many different methodologies that belong to this group; some of the best-known methods include naïve Bayes, random forest, K-nearest neighbour, logistic regression, and the support vector machine. All of these methods undergo a training phase, which can be either supervised (using label input data) or unsupervised (without label data), and use input features to establish connections among variables and acquire predictive power. However, owing to their simplicity, such methods have limitations when there is a need to capture nuanced semantic hints, as is the case with most linguistic tasks. On the other hand, they can often provide the foundation for the development of more powerful analytic tools and serve as a measuring stick to evaluate progress.

Machine learning techniques are used to aid in sign language recognition and have achieved some degree of success. Some of the earliest studies in this field were based on data input from wearable sensors, which provide a very direct translation of a user's movements. The data can be filtered using techniques such as SVM to provide a reasonably accurate recognition of the intended sign. Some of the aforementioned machine learning methods are used primarily to analyse static content (i.e., individual signs isolated in time and space), while in some cases, there have been attempts to interpret continuous segments of sign language speech, necessitating the use of dynamic models such as dynamic time warping or relevance vector machines. In general, basic stochastic models are better suited for simple SLR tasks, which is why they were extensively used in the early stages of research. These statistical models typically require less computing power than more complex architectures, although this depends on the number of analysed features as well as the size of the dataset. As more complex ASLR applications naturally require the inclusion of additional variables and sometimes additional modalities, the simplicity of basic models remains attractive.

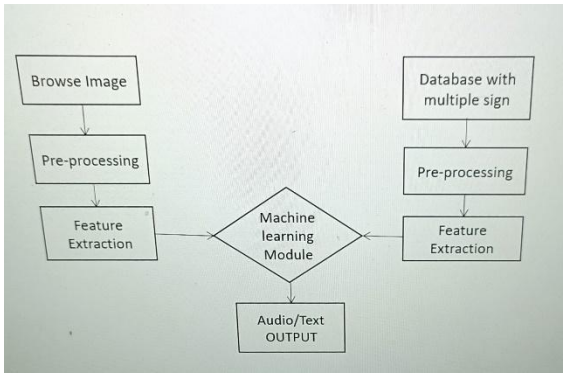
Communication is very crucial to human beings, as it enables us to express ourselves. We communicate through speech, gestures, body language, reading, writing or through visual aids, speech being one of the most commonly used among them. However, unfortunately, for the speaking and hearing impaired minority, there is a communication gap. Visual aids, or an interpreter, are used for communicating with them. However, these methods are rather cumbersome and expensive, and can't be used in an emergency. Sign Language chiefly uses manual communication to convey meaning. This involves simultaneously combining hand shapes, orientations and movement of the hands, arms or body to express the speaker's thoughts. Sign Language consists of fingerspelling, which spells out words character by character, and word level association which involves hand gestures that convey the word meaning. Fingerspelling is a vital tool in sign language, as it enables the communication of names, addresses and other words that do not carry a meaning in word level association. In spite of this, fingerspelling is not widely used as it is challenging to understand and difficult to use. Moreover, there is no universal sign language and very few people know it, which makes it an inadequate alternative for communication.

A system for sign language recognition that classifies finger spelling can solve this problem. Various machine learning algorithms are used and their accuracies are recorded and compared in this report.

The main aim of this project is to recognize the Sign Language using image processing and Machine Learning Algorithm.

This chapters gives introduction about Sign Language and used different techniques, from that techniques we have used one of technique for sign language We have successfully developed sign language detection project. This is an interesting machine learning python project to gain expertise. This can be further extended for detecting the English alphabets.

DESIGN AND DRAWING



Block Diagram of Sign Language

BROWSE IMAGE

In this system, we will know the HTML Image, how to add the image in HTML, along with knowing its implementation & usage through the examples. In earlier times, the web pages only contains textual contents, which made them appear quite boring and uninteresting. Fortunately, it wasn't long enough that the ability to embed images on web pages was added for users. In this article, we will know how to add images to the web page that will make the website attractive & various methods to insert the images.

By providing a full path or address (URL) to access an internet file. By providing the file path relative to the location of the current web page file.

We will first discuss inserting the image to the webpage & simultaneously, we will understand both the above approaches.

IMAGE PRE-PROCESSING

Image preprocessing contains cropping, filtering, brightness &contrast adjustment & many more. To do such process Image enhancement, Image cropping & Image Segmentation methods are used. Captured Images are in the form of RGB. So the first step is to convert RGB images to binary images then cropping of image is to be done so that unwanted part of i images can be removed. And now enhancement can be done in certain selected area. In Image segmentation, Edge detection method is used which can detect the boundary of cropped images which is further used for feature extraction method.

Image pre-processing are the steps taken to format images before they are used by model training and inference. This includes, but is not limited to, resizing, orienting, and colour corrections.

Image augmentation are manipulations applied to images to create different versions of similar content in order to expose the model to a wider array of training examples. For example, randomly altering rotation, brightness, or scale of an input image requires that a model consider what an image subject looks like in a variety of situations.

Image augmentation manipulations are forms of image pre-processing, but there is a critical difference: while image pre-processing steps are applied to training and test sets, image augmentation is only applied to the training data. Thus, a transformation that could be an augmentation in some situations may best be a pre-processing step in others.

FEATURE EXTRACTION

As describing our own features may not result in higher efficiency, we started with features as it computes the key points in the image which is more apt than describing features manually. So, after the skin segmented images were obtained using the model, we used the following approaches for extracting feature vectors.

Manual feature extraction requires identifying and describing the features that are relevant for a given problem and implementing a way to extract those features. In many situations, having a good understanding of the background or domain can help make informed decisions as to which features could be useful. Over decades of research, engineers and scientists have developed feature extraction methods for images, signals, and text. An example of a simple feature is the mean of a window in a signal.

Automated feature extraction uses specialized algorithms or deep networks to extract features automatically from signals or images without the need for human intervention. This technique can be very useful when you want to move quickly from raw data to developing machine learning algorithms. Wavelet scattering is an example of automated feature extraction.

With the ascent of deep learning, feature extraction has been largely replaced by the first layers of deep networks – but mostly for image data. For signal and time-series applications, feature extraction remains the first challenge that requires significant expertise before one can build effective predictive models.

Dataset with Multiple sign

In this system we have stored 26 alphabets from 'A to Z'. Each alphabet store 30 Images. We Added two more Images such as Delete and Space. In this Dataset Every Alphabet having Different sign. We are going to use trained dataset which will preprocessed. and send to SVM classifier/random forest classifier preprocessing means the process of transforming raw data into an understandable format Support Vector Machine" (SVM) is a supervised machine learning algorithm that can be used for both classification or regression challenges.

However, it is mostly used in classification problems Using webcam we can capture image. Captured image is uploaded then it is preprocessed an send to trained model Trained model will recognize which character is there and then it will convert text to speech using this speaker will recognize which character is there.

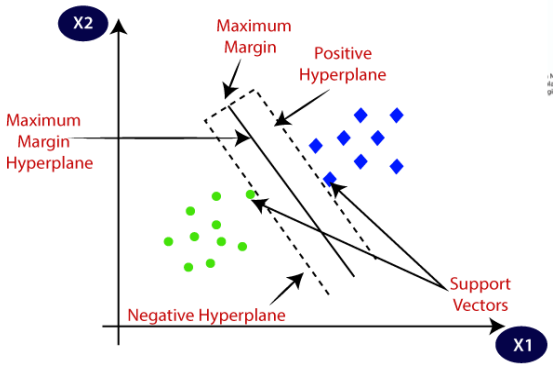
Comparison of features matching with database

Alphabets						
A	**284**	13	37	25	16	26
B	8	**332**	16	13	13	13
C	4	43	19	35	18	16
D	31	21	12	12	39	24
E	90	30	30	43	32	32
F	7	16	**392**	22	47	39
G	13	9	34	31	49	43
H	12	23	69	32	45	39
I	24	5	23	26	23	23
K	16	6	19	31	32	22
L	16	27	32	**387**	39	15
M	10	14	42	26	41	40
N	4	18	41	12	35	33
O	21	10	20	31	53	36
P	17	7	16	22	43	41
Q	16	15	34	38	33	50
R	12	10	32	10	29	22
S	7	9	24	17	**400**	36
T	20	17	34	33	38	39
U	4	16	19	23	53	12
V	2	15	16	57	20	26
W	20	7	42	34	50	68
X	24	12	40	38	28	92
Y	14	12	33	16	35	**597**
Z	11	5	12	27	25	34

ALGORITHM:

Support Vector Machine:

According to Wikipedia, SVM is a supervised machine learning model with associated learning analysis data. In this given a set of training example, we divide data into two classes on the basis of its labeling. If data is labeled it is put in category of supervised else in the category of unsupervised. Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. They are used in text and hypertext classification. SVM is used in hand written characters Recognization. They are used in image classification.



Graph of Support Vector Machine

It is simple to learn and use, but does that solve our purpose? Of course not! Because you can do so much more than just Regression!

Think of machine learning algorithms as an armory packed with axes, swords, blades, bows, daggers, etc. You have various tools, but you ought to learn to use them at the right time. As an analogy, think of 'Regression' as a sword capable of slicing and dicing data efficiently, but incapable of dealing with highly complex data. On the contrary, 'Support Vector Machines' is like a sharp knife – it works on smaller datasets, but on complex ones, it can be much stronger and powerful in building machine learning models.

By now, I hope you've now mastered Random Forest, Naive Bayes Algorithm, and Ensemble Modelling. If not, I'd suggest you take out a few minutes and read about them as well. In this article, I shall guide you through the basics to advanced knowledge of a crucial machine learning algorithm, support vector machines.

If you're a beginner looking to start your data science journey, you've come to the right place! Check out the below comprehensive courses, curated by industry experts, that we have created just for you:

Random Forest –

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of ensemble learning**,** which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model.

As the name suggests**,** "Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset."

In Random forest n number of random records are taken from the data set having k number of records Individual decision trees are constructed for each sample Each decision tree will generate an output Final output is considered based on Majority Voting or Averaging for Classification and regression respectively.

Adding one further step of randomization yields extremely randomized trees, or Extra Trees. While similar to ordinary random forests in that they are an ensemble of individual trees, there are two main differences: first, each tree is trained using the whole learning sample (rather than a bootstrap sample), and second, the top-down splitting in the tree learner is randomized. Instead of computing the locally optimal cut-point for each feature under consideration (based on, e.g., information gain or the Gini impurity), a random cut-point is selected. This value is selected from a uniform distribution within the feature's empirical range (in the tree's training set). Then, of all the randomly generated splits, the split that yields the highest score is chosen to split the node. Similar to ordinary random forests, the number of randomly selected features to be considered at each node can be specified.

As part of their construction, random forest predictors naturally lead to a dissimilarity measure among the observations. One can also define a random forest dissimilarity measure between unlabeled data: the idea is to construct a random forest predictor that distinguishes the "observed" data from suitably generated synthetic data.[9][25] The observed data are the original unlabeled data and the synthetic data are drawn from a reference distribution. A random forest dissimilarity can be attractive because it handles mixed variable types very well, is invariant to monotonic transformations of the input variables, and is robust to outlying observations. The random forest dissimilarity easily deals with a large number of semi-continuous variables due to its intrinsic variable selection; for example, the "Addcl 1" random forest dissimilarity weighs the contribution of each variable according to how dependent it is on other variables. The random forest dissimilarity has been used in a variety of applications, e.g. to find clusters of patients based on tissue marker data.

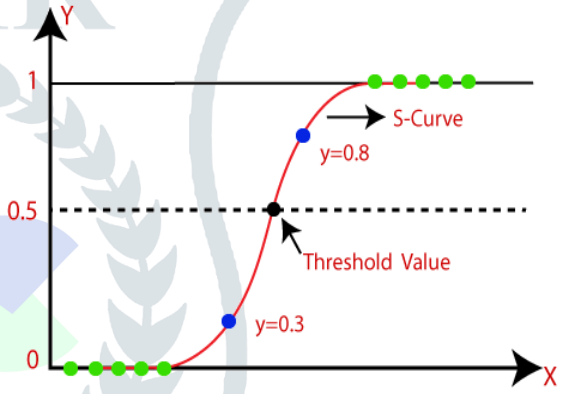
Logistic Regression:

Logistic regression is very popular in machine learning and statistics. It can work on **both binary and multiclass classification very well**. I wrote tutorials on both binary and multiclass classification with logistic regression before. This article will be focused on image classification with logistic regression. Logistic regression is one of the most popular Machine Learning algorithms, which comes under the Supervised Learning technique. It is used for predicting the categorical dependent variable using a given set of independent variables.

Logistic regression predicts the output of a categorical dependent variable. Therefore the outcome must be a categorical or discrete value. It can be either Yes or No, 0 or 1, true or False, etc. but instead of giving the exact value as 0 and 1, **it gives the probabilistic values which Lie between 0 and 1**.

Logistic Regression is much similar to the Linear Regression except that how they are used. Linear Regression is used for solving Regression problems, whereas Logistic regression is used for solving the classification problems.

Logistic Regression is a significant machine learning algorithm because it has the ability to provide probabilities and classify new data using continuous and discrete datasets.



Graph of Logistic Regression

Steps in Logistic Regression: To implement the Logistic Regression using Python, we will use the same steps as we have done in previous topics of Regression. Below are the steps:

1) Data Pre-processing step

2) Fitting Logistic Regression to the Training set

3) Predicting the test result

4) Test accuracy of the result (Creation of Confusion matrix)

5) Visualizing the test set result.

1. Data Pre-processing step: In this step, we will pre-process/prepare the data so that we can use it in our code efficiently. It will be the same as we have done in Data pre-processing topic.

2.Logistic Regression to the Training set: We have well prepared our dataset, and now we will train the dataset using

the training set. For providing training or fitting the model to the training set, we will import the Logistic Regression class of the sklearn library. After importing the class, we will create a classifier object and use it to fit the model to the logistic regression.

3. Predicting the Test Result: Our model is well trained on the training set, so we will now predict the result by using test set data.

4. Test Accuracy of the result: Now we will create the confusion matrix here to check the accuracy of the classification. To create it, we need to import the confusion matrix function of the sklearn library. After importing the function, we will call it using a new variable cm.

5. Visualizing the training set result: Finally, we will visualize the training set result. To visualize the result, we will use Listed Colormap class of matplotlib library.

IMPLEMENTATION

SOFTWARE SPECIFICATION

Image processing library: Open CV2

Open Source Computer Vision (Open CV) is an image processing and computer vision library mainly developed for artificial vision. It has a BSD license (free for commercial or research use). Open CV was originally written in C, but currently, it's a whole C++ interface, and there's additionally an entire Python interface to the library. Open-source computer Vision Library, also called Open CV, is associated with a freeware software package aimed toward computer vision. It is used in this project because of its versatility and the fact that it has a C++ interface. Open CV runs on most major Operating Systems (OS), making it worthwhile to use another computer to program or test.

SOFTWARE REQUIREMENT

The application will give the desired results only if the specified software requirements are satisfied. The backend coding is done in Python and VSS/HTML. The client or the web browser sends a POST request containing the data. The server processes the request and sends response back to client.

Language: Python 3.8

Python is a high-level programming language extensively used for programming. Python, an interpreted language, supports several programming scripts and a syntax that allows you to use programs in most languages such as Python and Vss/html. The language provides constructions designed to permit clear programs at each scale. Python is easy and simple to know, the python code is way easier than alternative languages.

Installation

I. Download Anaconda Application from Google Chrome

II. After download we have Install anaconda.

III. After installation we open this application.

IV. Then we have Launch a Spyder Application.

V. We open the google chrome after we have open a new window then we Select a python 3.8.0.

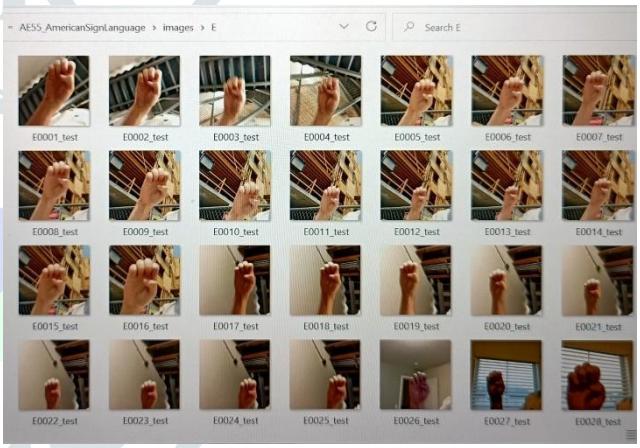
VI. Click the download button and you will see Python 3.8.

VII. Click Python 3.8. ...

VIII. Next, right click the mouse button you will see open button click to open.

IX. Enable to add Python 3.8 to path and click install now.

X. Wait a few minutes and display setup was successful.

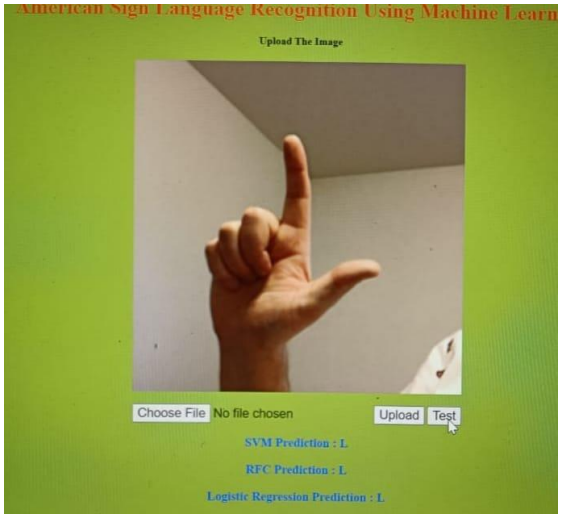


Dataset Hand Gesture

RESULTS AND DESCUSSION

The preprocessed data is passed through several classifiers (Random Forests, Support Vector Machines, Logistic Regression, Multilayer Perceptron) to draw effective results. The accuracy of the new models has been found significantly higher than the existing model.

Various types of techniques can be used to implement the classification and recognition of images using machine learning. Apart from recognizing static images, work has been done in the field of depth-camera sensing and video processing. A diversity of processes embedded in the system was developed using various other programming languages to implement the procedural techniques for the final system's maximum efficacy. The problem can be solved and systematically organized into three similar approaches which are, firstly using static image recognition techniques and preprocessing procedures.



Final Output

Output:- Trained model will recognize which character is there and then it will convert text to speech using this speaker will recognize which character is there.

CONCLUSION

In this work, we have gone through an automatic sign language gesture recognition system in real- time, using different tools. Although our proposed work expected to recognized the sign language and convert it into the text, there's still a lot of scope for possible future work.

Thus system was developed to translate the Sign Language into both text and speech This system helps and aid the hearing impaired and mute people to live independently. It develops confidence and will power to share their emotions, thoughts, ideas and difficulties with the normal people in the society. This eliminates the gaps among the people and achieve better society This system is also useful to all the people in the society gets as through this speech translation is possible, system output can be obtained in different languages.

REFERENCES

[1] ArtiThorat, VarshaSatpute, AratiNehe, TejashriAtre, Yogesh R Ngargoje, "Indian Sign Language Recognition Systern for Deaf People," IJARCCE, vol.3, issue 3,March,2014.

[2] Divya S, Kiruthika, S Nivin Anton A L and Padmavahi S, 'Segmentation, Tracking And Feature Extraction For Indian SignLanguage Recognition," IJCSA, vol.4, no.2, April,2014, pp. 57- 72.

[3] KonduruSatheesh Kumar Raju Sunny ArokiaSwamy B Anil Kumar G S KusumikaKrori Dutta, "Indian Sign Language to Speech," IJAER, vol.8 issue VI, Decernber, 2014, pp. 6-12.

[4] Rekha, K; Latha, B. "Mobile Translation System from Speech Language to Hand Motion Language", Intelligent Computing Applications (ICICA), 2014 International Conference on, On page(s): 411-415.

[5] Praveenkumar S HavalagiShruthiUrfNivedita, "The Amazing Data Gloves that give voice to the voiceless," IJAET, vol. 6, issue I,March,2013,pp. 471-480.

[6] MJerin Jose V.Priyadharshni et.al, "Indian Sign Language (ISL) Translation Systern For Sign Language Learning,"IJIRD, vol. 2, May,2013, pp. 358-365.

[7] YeUapuMadhuri, Anitha G, Anburajan MI, "Vision-Based Sign language Translation Device," International Conference on Infonnation Communication and Embedded Systerns (ICICES), IEEE, 2013, pp. 565-568.

[8] Shi, 1., and C. Tomasi, "Good Features to Track," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 1994, pp. 593600

[9] Rajam P S, Balakrishnan G,"Real Time Indian Sign Language Recognition Systern to aid deaf-dumb people", Communication Technology (ICCn, 2011 IEEE 13th International Conference, pp. 737- 742.

[10] M.W. Kadous, "Machine recognition of Australian signs using power gloves: Toward large-lexicon recognition of sign language," Proc. Workshop Integration Gesture Language Speech, pp. 165-174, 1996.

[11] http://www.nist.gov/speech/tests/mt/

[12] http://www.statmt.org/moses