Multiple Hidden Neuron based model for accurate American Sign Language translation

Sowjanya M
Assistant Professor, Department of MCA,
VTU-RRC, Mysuru, Karnataka,India

N Thimmaraju S N
Professor. Department of MCA,
VTU-RRC, Belagavi, Karnatak, India

Abstract: The origin of language marked a dawn of new era in human history. Ever since then, there have been amazing developments in formalizing languages and aiding their growth. There have been so many languages that were developed as a hobby or for motion pictures but have been formalized by the fan following and extensively used in their communities. However, one language that binds them all is sign language and it has become a universal tool for communication across nations, continents and even species. Animals have been trained to use sign language successfully. One of the most commonly used sign languages is the American Sign Language (ASL) and there have been a large number of experiments in its translation to formal languages. However there are so many other sign languages developed intrinsic to various cultures and regions, which have not been translated into formal languages due to lack of resources or effort. Machine learning systems can change that precisely. In this paper one such attempt has been presented.

Keywords: Sign language, ANN(Artificial Neural Network) , translation, mapping, error correction

I. Introduction

The understanding of sign language has evolved for all animals greatly with the advent of technology. Now there are systems capable of understanding various animal or bird calls based on their physical features. Sign language developed by humans however, has so many variations based on culture, geography, requirements and native language that it cannot be modeled using a single learning system. Each language has its unique feature set and nuance that requires a change in the methodology of learning and hence the design of the model. Thus, it becomes increasingly difficult to design and develop a single machine learning model which is fast as well as accurate and robust to adapt to the changing language systems. Most of the research in this direction as studied in the following literature review has been happening in small pockets and there is no unified approach towards a model that can learn a lot if not all of the sign languages. In this research, an artificial neural network with multiple input, hidden and output neurons has been developed and tested on one of the most widely utilized sign languages, the American Sign Language (ASL). It has about 16000 symbols varying from 1 character to 17 characters in length and offers an amazing chance for a machine learning system to develop an exhaustive understanding of its grammar.

II. Literature review

Thai translation system using upright speed-up robust feature and c-means clustering" [1] 2012 and "Thai translation system using upright speed-up robust feature and dynamic time warping" [2], 2012 has shown that Thai sign language has both finger spelling and symbol gestures, that need two separate translators.

"Computer translation system for hearing impaired users" [3], 2012 illustrated that Russian had peculiarities language translator has been presented. The stage of the Russian text analysis, including morphological, syntactic, and semantic analysis, is characterized .

"Research on Chinese-Japanese Translation System" [4], 2010 illustrated that main way for express their thought and emotions, obtain information and participate in social activities. All sign languages of the world are different from one another. With the increasing of economic and cultural exchanges between China and Japan, more and more deaf from countries learn partner's . At the same time, the translation of Chinese-Japanese sign languages has become important means of communication.

“Plenary talk II: Recent developments in recognition systems” [5] has shown that Automated translation systems for sign languages are needed world where physically challenged individuals in communicating and contributing to the society and the workforce. For the hearing-impaired, such systems can the equivalent of speech-recognition systems speaking people to interact with machines more natural way.

“Sign language localization: Learning to eliminate language dialects” [6] presented that of into spoken languages is yet non-trivial work. In any sign language there are so many sheer variety of dialects exist which makes to divide sign language into so many parts. The system works in two phases. In the training phase the correspondence between users hand gestures against each symbol is learn through feed forward neural network with back propagation learning algorithm. Once the training is complete, user can utilize the system for translation or communication with . “Toward transcription model in XML for Processing gloss annotation system”[7] illustrated that Representation of (SL) in written form is required for the generation via avatar or tools of synthesis. Also, the new transcription system useful for any processing tool like or building n-gram models.

“Evaluating a System for Deaf People”[8] have shown that of System for Deaf People and its field evaluation real application domain. For language translation, the system integrates three technologies: an example-based strategy, a rule-based translation method and a statistical translator.

“Vision-based translation device”[9] claim that an automatic translation of into speech to act as a communicator between deaf people and hearing and/or speech impaired people. It used as a translator for understand , avoiding by the intervention of an intermediate person communication using their natural way of speaking. This allows instantaneous recognition from finger and hand movements to translation. This is recognize one handed sign representations of alphabets (A-Z) and numbers (0-9).The results are found to be highly consistent, reproducible, with fairly high precision and accuracy.

“Signal processing for low cost optical data glove” [10] illustrated clearly that the aim is to produce a low cost data glove for sign language translation by using an optical detector technology and a 3D positioning system. The main innovation here is the optical system used for the glove which allows a great cost reduction. The glove also makes translation accessible. In this we describe the optical signal analysis as well as the 3D positioning. These two elements will the demonstration of translation methodology.

“Gesture recognition using kinect for translation” [11]. This paper presents method for identification of an isolated gesture using Microsoft Kinect. This paper presents the way of extracting some highly robust features from the depth image provided by Kinect and to use them in creating and accurate gesture recognition system, for of translation. The proposed algorithm helps in translating language gesture performed by a user, which in-turn used as an input applications.

“An Avatar-based Interface for the Italian ” [12] presented a virtual interpreter of the Italian (Italian Sign Language, LIS). developed as part of the ongoing ATLAS project, on the automatic translation from Italian to Italian Sign Language. The translation system communicates with the user through a virtual signer: the system takes as input representation of language sentence and produces the corresponding animation of the avatar. The architecture of the virtual signer consists of a resource planner, an executor of the planned sign animations, and an animation system.

“Spectral domain cross correlation function and generalized Learning Vector Quantization for recognizing and classifying Indonesian Sign Language”[13] illustrated that of a Kinect camera, Discrete Cosine Transform (DCT), Cross Correlation Function and classifying algorithm Generalized Learning Vector Quantization (GLVQ) can create system alphabet A to Z and to 10 in Indonesian Sign Language.

”Sign Language synthesis using hand motion acquisition” [14] illustrated that This paper addresses the critical issues we found during experiments for recording sign language with motion capture. This technology to capture hands motion with considerable precision and allows to store hand motion trajectories space. The goal is to provide an Italian to Italian Sign Language translation through a virtual character that is a computer graphic realization of a Sign Language interpreter.

“Development of computer sign translation technology for deaf people”[15] has shown an in depth characteristic of the text analysis stage, including prior semantic and semantic interpretation within the framework of developing a program for text computer sign language translator into colloquial sign language is given. Problems arising at this stage are also emphasized. Peculiarities of the sign language grammatical system necessary to take into account when developing a computer sign language translator are singled out.

"Finger-gesture Recognition Glove using Velostat (ICCAS 2011)" [16] illustrated that there are many researches into signing translation by finger-gesture recognition. Recognition technique of finger alphabet can be a substitute for the deaf that are unable to use sound recognition devices instead of computer input devices. They have developed a finger-gesture recognition glove that recognizes sign language with a use of inexpensive and conductive material. This material is called `Velostat', made of a film surfaced with carbon particles, used for anti-static package of electronic parts. Its resistance varies consistent with the strain of the film.

"Comparison of methods for hand gesture recognition supported Dynamic Time Warping algorithm" [17] illustrated that Gesture recognition may find applications in rehabilitation systems, signing translation or smart environments. The various methods tested were DTW - Dynamic Time Warping, DDTW - Derivative Dynamic Time Warping, PDTW - Piecewise Dynamic Time Warping, based on Dynamic Time Warping algorithm, which is used for hand gesture recognition using small wearable three-axial inertial sensor.

"Dataglove for consumer applications"[18] illustrated that Sensory gloves are promising technologies for human/machine interface and signing translation. For these applications, all degrees of freedom of the human hand have to be monitored with a good precision. For now, there's no sensory glove on the market meeting these requirements at low cost for widespread consumer applications. This paper describes a prototype of low cost sensory glove supported coupling loss fiber optics flexion sensors.

"Mobile motion gesture design for deaf people" [19] have shown that, so as to efficiently communicate with non-hearing-impaired (NHI) persons especially locations in real-time, deaf people need a more intelligent and straightforward to use tool beyond their signing . It can quickly translate sign language into text, with one mobile phone, through organized vocabularies in context of different particular locations and integrated touch screen with gesture recognition technology.

"Toward HMM based machine translation for ASL"[20] HMM-based models are widely used in many fields such as pattern recognition, speech recognition or Part-of-speech tagging. However, A HMM are often considered as a simplest dynamic Bayesian network. This network allows us to style a probabilistic graphical model which will be utilized in MT field especially for signing MT . They present a Bayesian Learning based method to train the alignment between a simple GLOSS form and a more complicated GLOSS form using sign language specificities such as space locative and classifier predicates.

"A survey of image-based Arabic signing recognition"[21] present that signing is that the language of deaf and hearing impaired people which they like to use on their lifestyle . Few interpreters are available to facilitate communication between deaf and vocal people. However, this is often neither practical nor possible for all situations. Advances in information technology encouraged the event of systems which will facilitate the automated translation between signing and speech , and thus removing barriers facing the mixing of deaf people in the society. A system has been that translate sign languages into spoken words and the reverse. By image based automatic recognition of the sign language.

"Image-Based and Sensor-Based Approaches to Arabic signing Recognition"[22] illustrated that signing continues to be the well-liked method of communication among the hearing-impaired. Advances in information technology have prompted the event of systems which will facilitate automatic translation between signing and speech . More recently, systems translating between Arabic sign and speech became popular. This paper reviews systems and methods for the automated recognition of Arabic signing . Additionally, this paper highlights the most challenges characterizing Arabic signing also as potential future research directions.

"Combining depth image and skeleton data from Kinect for recognizing words within the sign system for Indonesian language (SIBI [Sistem Isyarat Bahasa Indonesia])"[23] illustrate that a model for recognizing the SIBI words by using Microsoft Kinect as the input sensor can be developed. This model may be a a part of automatic translation from SIBI to text. The features for every word are extracted from skeleton and color-depth data produced by Kinect. Skeleton data features indicate the angle between human joints and Cartesian axes. The Generalized Learning Vector Quantization (GLVQ) and Random Forest (RF) training algorithm from WEKA data mining tools are used as the classifier of the model and image classified by Random Forest.

III. PROPOSED METHOD

The system proposes an artificial neural network for the automatic translation of sign language symbols into formal language (English). The data for the sign language symbols is obtained from the micro soft American sign language data set (MSASL). There are 20000 symbols in the ASL and new symbols are being added into its grammar. For the purpose of this research, the set as on date has been considered. The data set consists of the signed text, original text, its various meanings implied in conversation, target text as per the meaning, the classification labels (1 and 0) for correctness. A sample of the data is shown in Table 1.

Table 1. The data modified to suit the research requirement

Sign	Text	Target	Length	IT_1	IT_2	IT_3	IT_4	Class label
i	i	i	1	0.111006	0.283734	0.792799	0.928898	1
25	25	25	2	0.126291	0.190798	0.814862	0.863288	1
IN	in	in	2	0.006766	0.174534	0.727799	1	1
Inside	inside	in	2	0.20987	0.240643	0.703553	1	0
My, mine	my	my	2	0.209505	0.258152	0.7004	1	0
Hi	hi	hi	2	0.021881	0.034529	0.785179	1	1
IN	in	in	2	0.138068	0.23682	0.681766	0.88993	1
WE	we	we	2	0.135274	0.252501	0.713513	0.888308	1
UP	up	up	2	0.139129	0.323976	0.73971	0.883079	1
DO	do	do	2	0.127188	0.221336	0.812446	0.892901	1
Me	me	me	2	0.108728	0.266097	0.931328	1	1
SUPPOSE	if	if	2	0.006245	0.189094	0.861913	0.988685	0
HOT	hot	hot	3	0.098007	0.30826	0.753625	1	1
YOU	you	you	3	0.102086	0.323807	0.757818	1	1
MAD	mad	mad	3	0.103844	0.331641	0.761977	1	1

Since there need to be as many words as there are symbols (a one to many mapping system) it is proposed to utilize a multiple input, hidden and output neuron based ANN for the fast and accurate translation. Ideally where one output is triggered by a single hidden neuron, the multiple mapping would lead to errors. However in this approach, the weights and biases are computed to compensate for the errors introduced by the mapping. The ANN weights and biases are assigned using the entire data set as training set. The modeling continues till all samples are placed in their classes effectively. Instead of having more hidden layers, this man to many mapping with a single layer of hidden neuron, whose number itself is treated as a variable achieves higher accuracy as shown by the results. This is a supervised learning phase where human intervention is needed to add multiple target values for symbols that take different meanings according to usage, environment and context. Some examples are shown in figure 1 below. The ANN model is shown in Figure 2. As it can be seen, there are multiple weights and biases connecting each input to multiple hidden neurons and multiple output neurons which results in viable results with improved accuracy.

Figure 1. Hand gestures with multiple meanings (Courtesy: TripSavvy)

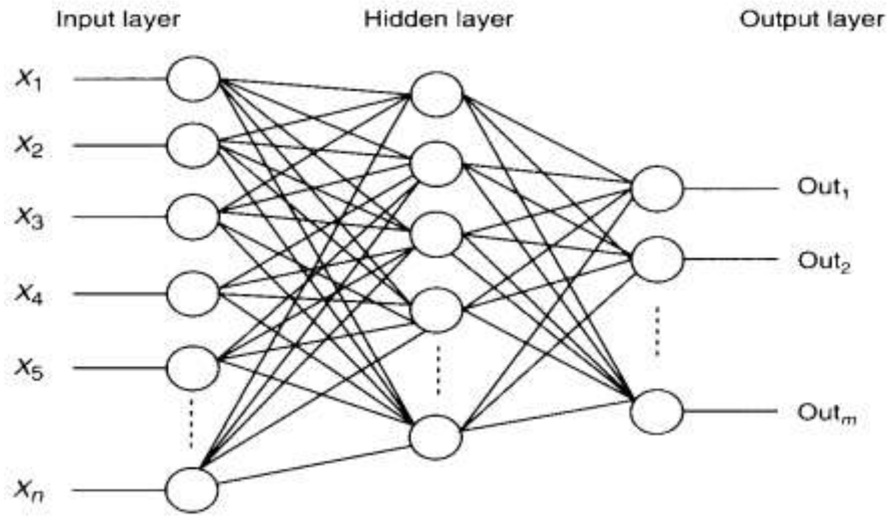

Figure 2.ANN Model for the given sign language data

IV. Experimental results and discussion

The data was split into training and test sets with 75:25 ratio. The data is fed to the ANN. The weights and biases are modified accordingly. The neurons activated using the sigmoid function remained the same but the hidden neurons are given a range of biases and weights in an attempt to push the output closest to the expected strings. The accuracy increased from iteration 1 through 4 where iteration 2 marked the beginning of a steep learning curve. The resultant accuracies are shown in Table 2. The plot of the same against the symbol length is shown in Figure 3.The test data is then fed to the ANN model. The accuracy corresponding to the length of words is shown in Table 3. The final graph that shows the capability of the ANN model is in Figure 4. From the table and the graph it is abundantly clear that the ANN model proposed in this paper achieves an accuracy closer to 100% (Lowest being 99.49).

Table 2: The result of a 4 iteration ANN training model for ASL translation (Data taken from :MSASL library) training set.

Word Length	Count	It_1	It_2	It_3	It_4
1	13	0.123083	0.21894	0.78549	0.963324
2	338	0.075934	0.221975	0.808319	0.979844
3	1691	0.068611	0.215689	0.797042	0.989155
4	3655	0.072413	0.214618	0.7962	0.989982
5	3272	0.071712	0.209656	0.792557	0.990826
6	2430	0.069649	0.213212	0.788642	0.989296
7	1478	0.072586	0.207934	0.796391	0.98798
8	1291	0.072352	0.206367	0.794487	0.989008
9	863	0.074028	0.201052	0.794087	0.98811
10	326	0.078177	0.207805	0.786988	0.989013
11	380	0.074121	0.193916	0.782961	0.983257
12	161	0.072394	0.199871	0.776029	0.986595
13	48	0.104953	0.210427	0.82828	0.990986
14	50	0.072724	0.22402	0.804592	0.983482
15	30	0.081439	0.243295	0.80906	0.986482
16	18	0.110806	0.178171	0.834464	0.982165
17	10	0.058138	0.139306	0.790245	0.999153

Table 3. The test results showing the accuracy against word length.

Length	Count	Accuracy
1	1	1.0000
2	76	0.9964
3	412	0.9973
4	1001	0.9971
5	832	0.9983
6	627	0.9981
7	376	0.9980
8	325	0.9969
9	235	0.9979
10	104	0.9981
11	98	0.9994
12	43	0.9981
13	10	0.9999
14	13	0.9994
15	9	0.9949
16	8	0.9998
17	2	0.9986

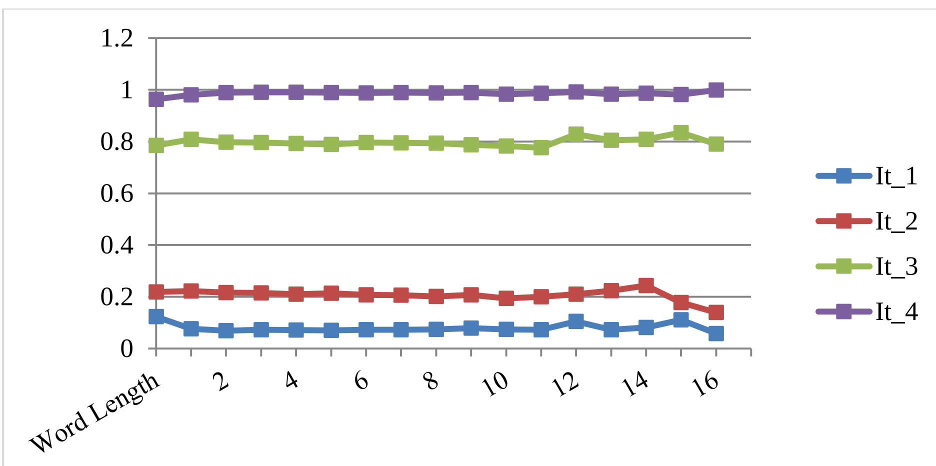

Figure 3. Plot of the iterative improvement in accuracy of the ANN model for the training set.

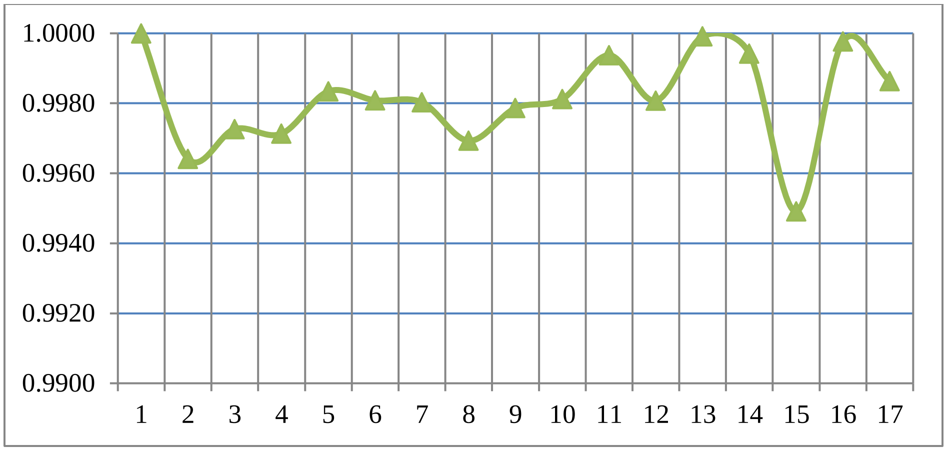

Figure 4. The accuracy of the ANN model on the test set

V. CONCLUSIONS

This paper has presented a model for translating sign language symbols into text using artificial neural network as the learning system. The experiment has shown optimal results and the accuracy obtained is commendable. It is intended to continue the research on other less popular sign languages to further improve the model. The improved model would have all the ANN training modules standardized to suit multiple sign languages. Data cleaning and other preprocessing tasks will come to the foreground with other languages and that would be a part of the improved model as well.

REFERENCES

[1] Chanda, P. ; Comput. Eng. Dept., Chiang Mai Univ., Chiang Mai, Thailand ; Auephanwiriyakul, S. ; Theera-Umpon, N. ,"Thai sign language translation system using upright speed-up robust feature and c-means clustering" , Fuzzy Systems (FUZZ-IEEE), 2012 IEEE International Conference , 2012.

[2] Chanda, P. ; Comput. Eng. Dept., Chiang Mai Univ., Chiang Mai, Thailand ; Auephanwiriyakul, S. ; Theera-Umpon, N.,"Thai sign language translation system using upright speed-up robust feature and dynamic time warping",Computer Science and Automation Engineering (CSAE), 2012 IEEE International Conference on (Volume:2),2012

[3] Grif, Mikhail G. ; Department of Automatic Control Systems, Novosibirsk State Technical University, Novosibirsk, Russia ; Korolkova, Olga O. ;Demyanenko, Yelena A. ; Tsoy, Evgeniy B., "Computer sign language translation system for hearing impaired users", Strategic Technology (IFOST), 2012 7th International Forum on , 2012

[4] Zhen Liu ; Grad. Sch. of Eng., Nagasaki Inst. of Appl. Sci., Nagasaki, Japan ; Xiaomei Zhang ; Kato, J.,"Research on Chinese-Japanese Sign Language Translation System" , Frontier of Computer Science and Technology (FCST), 2010 Fifth International Conference on, 2010.

[5] Tolba, M.F. ; Fac. of Comput. & Inf. Sci., Ain Shams Univ., Cairo, Egypt, "Plenary talk II: Recent developments in sign language recognition systems", Computer Engineering & Systems (ICCES), 2013 8th International Conference on ,2013

[6] Tariq, M. ; Dept. of Comput. Sci., Kinnaird Coll. for Women, Lahore, Pakistan ; Iqbal,A. ; Zahid,A. ; **etal** "Sign language localization: Learning to eliminate language dialects", Multitopic Conference (INMIC), 2012 15th International,2012

[7] Othman, Achraf ; Research Laboratory of Technologies of Information and Communication & Electrical Engineering LaTICE-GE, Tunis National Higher School of Engineering (ENSIT), 5, Avenue Taha Hussein, B.P. 56, Bab Menara, 1008, University of TUNIS, Tunisia ; Hamdoun, Raouia, "Toward a new transcription model in XML for Sign Language Processing based on gloss annotation system" , Information and Communication Technology and Accessibility (ICTA), 2013 Fourth International Conference on, 2013.

[8] Lopez-Ludena, V. ; Grupo de Tecnol. del Habla, Univ. Politec. de Madrid, Madrid, Spain ; San-Segundo, R. ; Martin, R. ; Sanchez, D., "Evaluating a Speech Communication System for Deaf People", Latin America Transactions, IEEE (Revista IEEE America Latina) (Volume:9 , Issue: 4), 2011

[9] Madhuri, Y. ; Dept. of Biomed. Eng., SRM Univ., Kattankulathur, India ; Anitha, G. ; Anburajan, M.,"Vision-based sign language translation device", Information Communication and Embedded Systems (ICICES), 2013 International Conference on, 2013

[10] Trottier-Lapointe, W. ; PolyProject, Tech. Soc. of Ecole Polytech. de Montreal, Montréal, QC, Canada ; Majeau, L. ; El-Iraki, Y. ; Loranger, S. ,"Signal processing for low cost optical data glove", Information Science, Signal Processing and their Applications (ISSPA), 2012 11th International Conference on,2012

[11] Verma, H.V. ; Dept. of Comput. Sc. & Eng., Jaypee Inst. of Inf. Technol., Noida, India ; Aggarwal, E. ; Chandra, S. ,"Gesture recognition using kinect for sign language translation", Image Information Processing (ICIIP), 2013 IEEE Second International Conference on,2013

[12] Lombardo, V. ; Dipt. di Inf., Univ. degli Studi di Torino, Torino, Italy ; Battaglino, C. ; Damiano, R. ; Nunnari, F. ,"An Avatar-based Interface for the Italian Sign Language", Complex, Intelligent and Software Intensive Systems (CISIS), 2011 International Conference on,2011

[13] Rakun, E. ; Fac. of Comput. Sci., Univ. of Indonesia, Depok, Indonesia ; Febrian Rachmadi, M. ; Andros ; Danniswara, K. ,"Spectral domain cross correlation function and generalized Learning Vector Quantization for recognizing and classifying Indonesian Sign Language", Advanced Computer Science and Information Systems (ICACSIS), 2012 International Conference on ,2012

[14] Hatami, N. ; Dipt. di Autom. e Inf., Politec. di Torino, Turin, Italy ; Prinetto, P. ; Tiotto, G.,"Sign Language synthesis using hand motion acquisition", Design & Test Symposium (EWDTS), 2010 East-West, 2010

[15] Grif, M.G. ; Dept. of Autom. Control Syst., Novosibirsk State Tech. Univ., Novosibirsk, Russia ; Demyanenko, Y.A. ; Korolkova, O.O. ; Tsoy, Y.B. ,"Development of computer sign language translation technology for deaf people", Strategic Technology (IFOST), 2011 6th International Forum on (Volume:2) ,2011

[16] Eunseok Jeong; Dept. of Electr. & Electron. Eng., Yonsei Univ., Seoul, South Korea ; Jaehong Lee ; DaeEun Kim " Finger-gesture Recognition Glove using Velostat (ICCAS 2011) , Control, Automation and Systems (ICCAS), 2011 11th International Conference on,2011

[17] Barczewska, K. ; Dept. of Autom. Control & Biomed. Eng., AGH Univ. of Sci. & Technol., Kraków, Poland ; Drozd, A., "Comparison of methods for hand gesture recognition based on Dynamic Time Warping algorithm", Computer Science and Information Systems (FedCSIS), 2013 Federated Conference on,2013

[18] Majeau, L. ; PolyProject, Tech. Soc. of Polytech. of Montreal, Montreal, QC, Canada ; Borduas, J. ; Loranger, S. ; El-Iraki, Y. ,"Dataglove for consumer applications",Fibre and Optical Passive Components (WFOPC), 2011 7th Workshop on ,2011

[19] Haoyum Xue ; Sch. of Eng. & Design, Brunel Univ., Uxbridge, UK ; Shengfeng Qin ,"Mobile motion gesture design for deaf people", Automation and Computing (ICAC), 2011 17th International Conference on, 2011

[20] Boulares, Mehrez ; Research Laboratory of Technologies of Information and Communication & Electrical Ingineering (LaTICE), Ecole Supérieure des Sciences et Techniques de Tunis, 5th Av. Taha Hussein, B.P. 56, Bab Mnara 1008, TUNISIA ; Jemni, Mohamed ,"Toward HMM based machine translation for ASL", Information and Communication Technology and Accessibility (ICTA), 2013 Fourth International Conference on ,2013

[21] Mohandes, M. ; Dept. of Electr. Eng., King Fahd Univ. of Pet. & Miner., Dhahran, Saudi Arabia ; Junzhao Liu ; Deriche, M., "A survey of image-based Arabic sign language recognition", Multi-Conference on Systems, Signals & Devices (SSD), 2014 11th International, 2014

[22] Dhahran 31261, Saudi Arabia (e-mail: mohandes@kfupm.edu.sa; mderiche@kfupm.edu.sa; liujunzhao@kfupm.edu.sa). ; Deriche, M. ; Liu, J. ,"Image-Based and Sensor-Based Approaches to Arabic Sign Language Recognition",Human-Machine Systems, IEEE Transactions on (Volume:PP , Issue: 99),2014

[23] Rakun, E. ; Fac. of Comput. Sci., Univ. Indonesia, Depok, Indonesia ; Andriani, M. ; Wiprayoga, I.W. ; Danniswara, K., "Combining depth image and skeleton data from Kinect for recognizing words in the sign system for Indonesian language (SIBI [Sistem Isyarat Bahasa Indonesia])" , Advanced Computer Science and Information Systems (ICACSIS), 2013 International Conference on, 2013

[24] M N Sowjanya, S N Thimmaraju A Comparison of Methods Used to Convert Sign Languages to Relative Formal Languages, JARDCS, Vol. 11, 01-Special Issue, 2019