IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Sign Language To English Conversion Using Machine Learning

Mr.Navakishore[1], Pooja Sakinala[2], Ginjupally Sai Nikhil[3], Bhavani Prasad[4] , and Katla Srinadh[5]

[1]Associate Professor, Hyderabad institute of technology and management, Medchal, Telangana

[2]UG student, Hyderabad institute of technology and management, Medchal, Telangana

[3]UG student, Hyderabad institute of technology and management, Medchal, Telangana

[4]UG student, Hyderabad institute of technology and management, Medchal, Telangana

[5]UG student, Hyderabad institute of technology and management, Medchal, Telangana

Abstract: Developing a desktop app where a PC's webcam tracks a person making sign language (fingerspelling) gestures and instantly converts it to text and voice. In the developed ASL interpreter, the sign language signs will be taken in form of text , which will later be voiced out. In this way we are developing a finger spelling action language converter. For recognizing these gestures, Convolutional neural network (CNN) is of assistance. A CNN is known to be very resourceful in addressing computer vision challenges and upon adequate training can reach the required features with precision.

Index Terms -CNN,ASL,training

I. INTRODUCTION

Deaf people often resort to American Sign Language (ASL) which is a natural language that takes its roots from spoken languages despite having different linguistics based on body-grammar. There in America, people who cannot hear or see, it is an appealing source of nonsense. There is no sign language that is formal or informal. Different signal languages are based on certain geographic territories. For instance, British Sign Language (BSL) is a completely distinct language from American Sign Language (ASL), and people in the
USA who are used BSL would have a hard time understanding ASL. Some countries integrate features of ASL found in their native sign languages. Sign language is a form of communication between individuals with normal speech and hearing who use verbalisms devoid of phonation and audition. Approximately 360 millions of people across the globe are afflicted by hearing loss with 328000000 of them being adults and 32000000 being children. Any degree of hearing loss which causes difficulty in hearing sounds of more than 40 decibels in the better ear is termed as a disabling hearing impairment. Therefore, combined with an increase in the population of the deaf, there is also an increased need for interpreters. There arises a need for intervention in order to bridge the communication gap between the hard of hearing individuals and the hearing individuals. Among the many areas of research that people are venturing into with sign language, the translation of signed languages is the fastest developing area today. And it is the most anthropometric mode of communication for the hearing challenged individuals. A device that can detect hand movements can enable a deaf person to communicate with a hearing person without the need for an interpreter. The solution is aimed at an automatic transliteration from American Sign Language to text and voice outputs. In the proposed device, a massive set of samples has been made use of to understand isolated phrases from the same old American sign language, which may be a cause for concern regarding the use of a virtual camera. Based upon all the signs including their respective alphabets, the database

consists of one thousand specific images of gesture. Such studies are focused on some very rudimentary aspects of signal language and the intention is to convert them to text and audio. American Sign Language is a language one can see. In addition to the signing, the brain concepts ways the language is transcribed with the eyes. Hand shape, orientation, and movement, and facial expressions as well as movement of the frame all play a significant role in communicating information. Sign language is not a common language – both the complete U. It Has its own signal 6 language along with regions having dialects like how many languages are there in the world which is called the English language, the recognition rate by the ASL language as in compare to the proficient level is of ninety percent of the organizations in general employing Indian sign language. The beautiful Portions of India it [ISL] has a little variation in signing cultures but the structure remains constant at certain levels within the America. Indian deaf people have this perspective that this is much better than other sign languages because it comes naturally to them – they learn it through the environment and people around them. The stages in acquisition of sign language are in hierarchies just like oral languages, the children start out babbling with their hands. Given that India too doesn't have so many Training institutions per se for Indian Sign Language [except for ISLRTC which was set up last year: it might be the hope of ISL] there is an absence of comprehension regarding this language registered with a number of people.

II. LITERATURE SURVEY

Paper [1] elaborates a hands free demonstration of Taiwanese data language having a wireless system for processing the data. To differentiate hand motion, they have inner sensors put into gloves to show the parameters as given by, posture, orientation, motion, defined of the hand in Taiwanese Sign Language could be recognize in no error. Flex inner sensor of a hand measures the gesture and the palm size is measured using a g sensor while movement is measured with a gyroscope. Input signals would have to be consider for testing for the sign to be legal or not periodicaaly. Since the signal that was sampled may sustain for a time altre than the time threshold, the transmitted legal gesture through a mobile device incorporating Bluetooth and used for sign differentiation and conversion. The performance in gesture recognition with the introduced system and algorithm is also good. As demonstrated the result get the accuracy of 94% with the concurrent architecture.

Realizing a sign language detector working in real time increases the efficacy of the community to able to reach out to even the most disabled persons such as those who can see or hear. Authors [2] have using machine learning algorithms presented the idea of a translation with skin colour tone to detect the ASL. They have created a skin tone-based segmentation which accurately detects skin tone colours and allows them to be assigned a voice for detection. They have utilized YCbCr color space as it is commonly used for color coding in video template purposes and gives good quality human skin tone differentiation. Further they have taken the CbCr plane to distribute the skin tone color. People from Different ethnicity have their tones different which is crafted in a model. The authors of the document [3] utilized a technique employing animation making which they termed synthetic animation making technique to convert Malayalam language to Indian sign language. HamNoSys is being used in this case as an intermediate representation for the sign language in this method. In this method, the application takes certain predetermined word groups say one or more and animates It. There exists an interactive system which converts the portion of words into HamNoSys system lingo. It is application which also curates everything that has been built as it has been in use by the Kerala's government in order to educate the people about sign language and its sensitization. Communicating with deaf individuals or even within the hearing community has become a problem in the recent times and as the society does not have proper such translator to put I plan it is akin to a day time id. Authors [4] have recommended ADA which is very much for the deaf and hearing aid users in the form of a communicate app. However, it is important to note that the level of complexity associated with developing an app for it is such that it no longer becomes a straightforward touch as there are a lot of factors to consider such as the design, memory usage among others. What their application does is that they take a picture of a sign gesture and later converts is to a meaningful word. At first, they have compared the gesture using histograms with respect to the sample test and other samples which are specifically related to BRIEF offered to lessen the processing time and the CPU burden. They have explained a process on which on their app, it's very easy to add up a gesture and store it in their database for further and expand detection set. So lastly, they came strong with having an app as a translator instead of several applications that are being used lately by the users.

Paper [5] is semi complete. Deaf persons communicate using sign language which in English is commonly referred to as ASL. Within their created device or application, there are numerous terms, for instance, a visible interface that allows a deaf individual to input the order of sign data, a translator that prepares the sequences in structured form for the Spanish language, a text converter to speech which simply means that all these bits are put together into a grammatically correct Spanish sentence of course. They have done a lot of work towards creating an interface that is visually friendly to the users who are impaired where they have come up with many ways of writing the sign language in real time. This is what they have used in an example of a translation system to describe as a final system model. Mostly deaf users drawn from cities like Madrid and Toledo are the sources of the test data provided as part of the initial data compilation that included all kinds of information including metrics. This Spanish corpus is the first of its kind designed for a general research that concentrates on a narrow field only. With more than 4000 Spanish language sentences in it, which are converted into LSE afterwards. After many versions finally, they have given a ready to use translator based on the domain of renewal of Id records and having used pressure in license.

The authors[6] have created a system which operates continuously in that sign language gesture sequence is input in order to create a training set of the system and bring up a sign from the training set. They have presented a paradigm in which instance learning as density matrices algorithm is utilized to oversee the entire sentence construction and learning of the sign composed by several gestures in response to noisy inputs. The first one that they had applied to demonstrate a continuous stream of word items and representations is then used to assist in the identification of gestures. They have applied this set to a limited amount of automatic cut sign data which is known to them for training, recognition and storage of a particular cut sign information. Around thirty sign language cut data was saved that is derived from the proposed designs. Mexican Sign Language (LSM) network is a wholly deaf community in Mexico, and a collection of various hand movements and facial expressions forming certain images. The authors explained also [7] that the absence of an automated mechanism that can convert such gestures to an LSM is a serious hurdle in the assimilation of the hearing impaired into the society. Such a picture gives a completely different approach to rendering recognition of LSM alphanumeric characters which uses 3D Haar like features obtained from the depth images using Microsoft Kinect. The features are analyzed with a strong ensemble of algorithms. In order to demonstrate that the performance of our method is normal, we Took a complex and fast sign made of numbers and letters, and showed the results in comparison with the results of traditional 2D-hair-like features. Static static LSM signs as well as dynamic signs with enhanced accuracy than that achieved by our device are recognized by the apparatus with extensively used 2D features. As It has been preferred for a society that having a sign language for hearing impaired and deaf people to communicate. As coming of several technologies in the In the past, developing a translation tool that translates sign into the right sentence has been fairly simple and quite common as well. They have demonstrated this in their workAs they have illustrated in their work easiest MTS is based primarily on an Arabic sign language, which is embedded in the translation, aiming at fine etiquette of interactions, however they have also illustrated that their project scope lies beyond the application and the set of parameters used of that application. The final output of the application translates Arabic sign language in text format through a machine learning system, the authors established.

In research [9], the authors have developed a speech-to-sign language system for deaf individuals through various tests. The developed system even incorporates animated demonstrations and a translation module in order to enable speech recognition based on ASR systems. Further on, the following strategies for carrying out the translation process were employed: state transducer and phrase structure grammar. For evaluation purposes the following metrics were adopted: WER, evaluated first BLEU, last NIST. The following passage shows the manner in which we speech by machine translation corrector with all three systems outlined earlier. The research showed that a finite state transducer with
WER of 28.21% to 29.27% for ASR output was achieved. The next [1O] article represents the application of contemporary and novel as well as reconfigured communication systems which predominantly takes aim at the targeted audience of the disabled people as they have further clarified. The system consisted of the following two scenarios such that first translates speech in Spanish to sign language in Spanish without the aid of a translation.

The article in [11] discusses the way deaf people interact with other members of the society. It also shows how different parts of the body such as the arms, hands, fingers and even facial expressions are put to use. Motion capture has helped in the communication of these movements as well as in the depiction of sign language. The dactyl technology modeling alphabet employs a 3-d model constructed of various segments. This was developed with respect to the efficiency of the realisation on the Ukrainian sign language. This is based for the universal character and designed for further model.

As times are improving and developing sign language emerges as one the best medium to communicate with the deaf community by having the supportive system to be in a help to society. The project they have demonstrated [12] here are trying to make the communication easy by having the sign mainly having dynamic and static in ISL are being converted into the speech. A placed sensor of glove with sensor of flex help to design the orientation of hand and following actions. Using wireless transmission which converts it to the further bits of speech as the output. In this project they studied about LSTM networks which for long time formed dependencies. The result of this projects leads to a success rate of 98% accuracy which could able to identify the 26 gestures.

III. DESCRIPTION OF THE METHOD

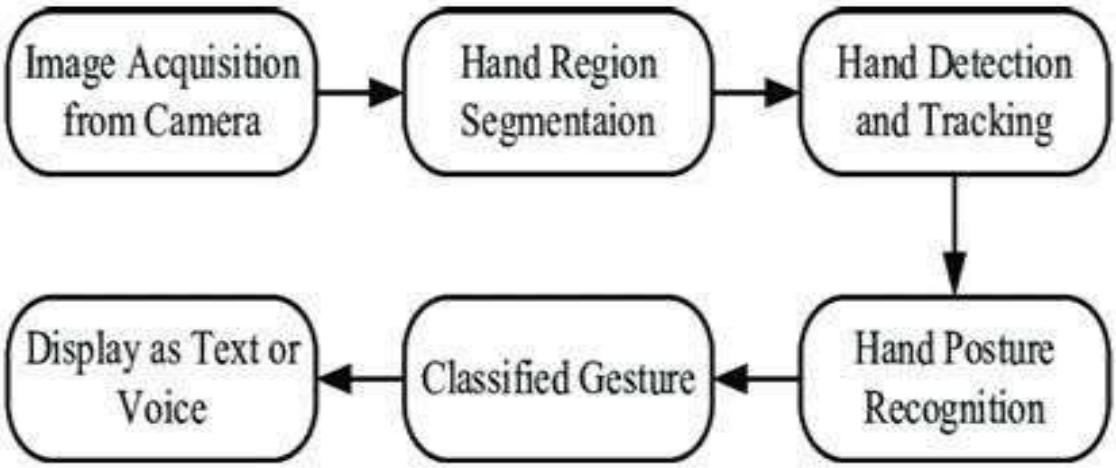

Fig 3.1 Flow Chart of the Project

Message conversion in the real time as signs and through speech: 1. Detection of sign language for male or female. 2. Employing the machine vision for converting the images to words training phase. 3. Word construction 4. Sentence construction 5. Content construction 6. Audio production A. Schematic representation of the process This is a schematic representation of the steps which contribute to achieving the aim of the project. The outline of these stages has been provided in greater details below. 1. Image Acquisition The images of the moving hands and faces are available from the internet camera. During the whole signing, an OpenCV stream video with the camera is on. Some frames of the stream are taken out and examined as black and white images measuring 50*50. This measurement is maintained throughout the work as the entire available data set is of that same size. 2. Hand Region Segmentation & Hand Detection and Tracking Scanning of the Hand Region for Hand Gestures. This step occurs in the preprocessing stage of the image, which is prepared for running into the model to get the prediction. The areas of the image with hand motion are enhanced. This applications so betting the odds in favor of winning many times. 3. Hand Posture Recognition Images after preprocessing are given to the keras CNN. The CNN model that has been trained already provides predicted label. Probability is assigned to each of the gesture labels. The maximum probability value is assumed to be the predicted label. 4. Text to Speech synthesis The recognized gesture helps in building up the word in the model. Using pyttsx3 library, the recognized words are turned into the corresponding speech.

The speech synthesis in text-to-speech is very basic but it surprisingly enhances the experience by giving it a touch of an oral dialogue. B. Convolutional Neural Network for Speech and Language Processing.

C. The functioning of Cnn Architecture

The CNN model for this project consists of 11 layers. There are 3 convolutional layers. The first convolutional layer, which is responsible for identifying low-level features such as (for example) lines, accepts as its input an image of size 50*50 in the grayscale imagery. 16 filters of size 2*2 are applied in this layer consequently an activation map of 49*49 is obtained for each filter which implies output is equal to 49*49*16. A rectifier linear unit (relu) layer is also included in order to remove any negations on the map and in its definite presence, all the areas are left at fill of 0s. After that, a maxpooling layer is applied whereby the activation is reduced to 25*25 by taking only the maximum values in 2*2 areas of the maps. This step enhances the chances of obtaining the required feature. This is also from the head of second convolution layer, whose objective is to find features such as, for example, angles, and curves. The layer has 32 filters measuring 3*3 that produce an activation map which is 23*23, which means output is equal to 23*23*32. A maxpooling layer again reduced the activation map to 8*8*32 by selecting the maximum in every 3*3 area of the map. The third convolution layer works on the input where high level features such as gestures and shapes are incorporated. 64 filters 5*5 in dimension makes the output to be 4*4*64. A maxpooling layer makes the map to be 1*1*64. The map is then changed to a 1 dimension array of 64 elements length. A dense layer brings the map to a dimension of 128 sizes. A dropout layer drops out random elements from the map to avoid overfitting. In the end, a dense layer brings down the map to an array of 44 numbers which correspond to classes available. Each class has a corresponding probability of prediction associated with it.

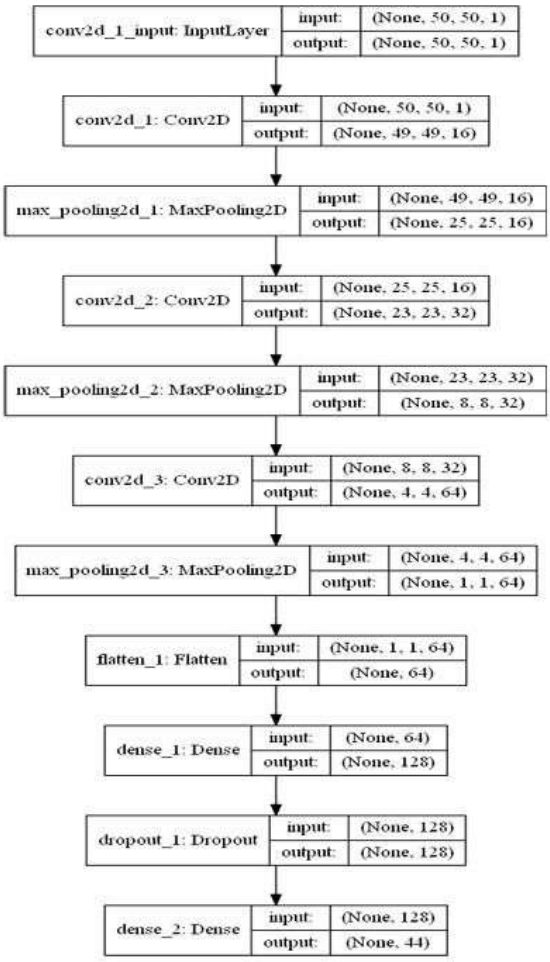

Fig 3.2 The CNN Architecture for the project

D. Identification of Letters and numerical

To find the bounding boxes of different-shaped items, we applied Gaussian foreground background subtraction which applied to model each pixel of the background into a mixture of K Gaussian models (k where k is from 3 to 5). The probable background colors are those that tend to stay for longer, the more static ones. On those nonbackground pixels, we draw a square envelope. After Getting all the pose and background, a Convolutional NN model has been built using those images to separate the pose elements from the background. These maps of features also show that the CNN can learn 'invisible' shapes of a few of the pose cues that are present in the training set without having to see them.

Algorithm

Real time sign language conversion to text and Start

S1: Set the hand histogram to adjust with the skin complexion and the lighting conditions.

S2: Apply data augmentation to the dataset to expand it and therefore reduce the overfitting. S3: Split the dataset into train, test and validation data sets.

S4: Train the CNN model to fit the dataset. S5: Generate the model report which includes the accuracy, error and the confusion matrix. S6: Execute the prediction file – this file predicts individual gestures, cumulates them into words, displays the words as text, relays the voice output. Stop

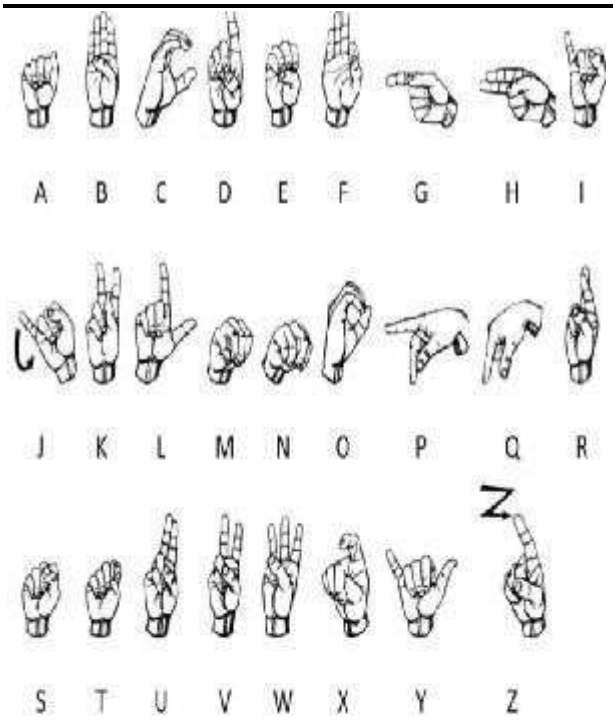

Fig 3.3 The Gesture Symbols for ASL Alphabets that will be in the training data

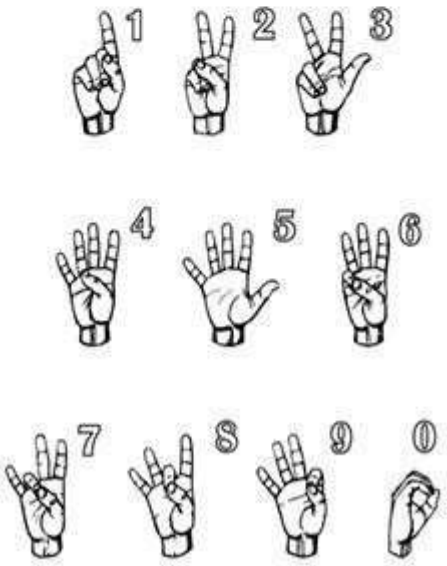

Fig 3.4 The Gesture Symbols for ASL Numbers that will be in the training data

IV. CONCLUSION

returns and payments, reducing the time and effort required compared to traditional methods. This efficiency can lead to faster processing and fewer errors in filing. This software has mainly two phases one is the design phase and another is the coding phase, in designing uses XML language, and coding uses Java language. Android Studio is the official IDE for Android app development and is widely used by developers worldwide. It continues to evolve with regular updates and improvements to support the latest Android SDKs and tools.

4.1 How to Create GST App in Android Studio?

This project serves as a brief introduction on how CNN architectures can be applied to computer vision ever so effectively. A finger spelling sign language translator is achieved with a success of 95%. It is possible the project looks into into other sign languages as it will only involve the development of the corresponding dataset and retraining the CNN. People often use sign languages in contexts rather than spell languages, hence, this approach would only be appropriate in solving a portion of the Sign Language Translation problem. The main aim has been achieved, that is, the interpreter is no longer necessary. There are a few finer points that need to be considered when we are running the project. The thresh needs to be monitored so that we don't get distorted grayscales in the frames. If for any reason this problem is experienced, we either have to clear the histogram or find better-lit areas. We could also put on gloves to do away with the varying shades of complexion that the signer may have. In this project, we could make a precise prediction as soon as we began experiments with the glove. Another concern which people may have will be the ability to perform the required ASL gestures. Improper gesture positioning will result in failure of prediction. This project can be further improved in certain ways for instance it can be developed as a web or mobile application for easy access to the users and also the present project does not support any native sign languages other than ASL but with sufficient amount of data and training it can be updated to support other native sign languages. This project does create a finger spelling language translator, but it is worth noting that each sign language can also be contextually used where every sign can be an object, a verb, etc, therefore interpreters have to understand contextual signing called.

REFERENCES

[1] L. Kau, W. Su, P. Yu and S. Wei, "A real-time portable sign language translation system," 2015 IEEE 58th International Midwest Symposium on Circuits and Systems (MWSCAS), Fort CO, 2015, pp. 1-4, Collins, 10.1109/MWSCAS.2015.7282137.

[2] S. Shahriar et al., "Real-Time American Sign Language Recognition Using Skin Segmentation and Image Category Classification with Convolutional Neural Network and Deep Learning," TENCON 2018 - 2018 IEEE Region 10 Conference, Jeju, Korea (South), 2018, pp. 1168-1171, doi:

10.1109/TENCON.2018.8650524. doi:

[3] M. S. Nair, A. P. Nimitha and S. M. Idicula, "Conversion of Malayalam text to Indian sign language using synthetic animation," 2016 International Conference on Next Generation Intelligent Systems (ICNGIS), Kottayam, 2016, pp. 1-4, doi: 10.1109/ICNGIS.2016.7854002.

[4] M. Mahesh, A. Jayaprakash and M. Geetha, "Sign language translator for mobile platforms," 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Udupi, 2017, pp. 1176-1181, doi: 10.1109/ICACCI.2017.8126001.

[5] S. S Kumar, T. Wangyal, V. Saboo and R. Srinath, "Time Series Neural Networks for Real Time Sign Language Translation," 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, 2018, pp. 243-248, doi: 10.1109/ICMLA.2018.00043.

[6] D. Kelly, J. Mc Donald and C. Markham, "Weakly Supervised Training of a Sign Language Recognition System Using Multiple Instance Learning Density Matrices," in IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 41, no. 2, pp. 526-541, April 2011, doi:

10.1109/TSMCB.2010.2065802.

[7] J. Jimenez, A. Martin, V. Uc and A. Espinosa, "Mexican Sign Language Alphanumerical Gestures Recognition using 3D Haar-like Features," in IEEE Latin America Transactions, vol. 15, no. 10, pp. 20002005, Oct. 2017, doi: 10.1109/TLA.2017.8071247.

[8] M. Mohandes, M. Deriche and J. Liu, "Image-Based and Sensor-Based Approaches to Arabic Sign Language Recognition," in IEEE Transactions on Human-Machine Systems, vol. 44, no. 4, pp. 551-557, Aug. 2014, doi: 10.1109/THMS.2014.2318280.

[9] R. San Segundo, B. Gallo, J. M. Lucas, R. Barra-Chicote, L. F. D'Haro and F. Fernandez, "Speech into Sign Language Statistical Translation System for Deaf People," in IEEE Latin America Transactions, vol. 7, no. 3, pp. 400- 404, July 2009, doi: 10.1109/TLA.2009.5336641.

[10] V. Lopez-Ludena, R. San-Segundo, R. Martin, D. Sanchez and A. Garcia, "Evaluating a Speech Communication System for Deaf People," in IEEE Latin America Transactions, vol. 9, no. 4, pp. 565570, July 2011, doi: 10.1109/TLA.2011.5993744.

[11] I. Krak, I. Kryvonos and W. Wojcik, "Interactive systems for sign language learning," 2012 6th International Conference on Application of Information and Communication Technologies (AICT), Tbilisi, 2012, pp. 1 3, doi: 10.1109/ICAICT.2012.6398523.

[12] E. Abraham, A. Nayak and A. Iqbal, "Real-Time Translation of Indian Sign Language using LSTM," 2019 Global Conference for Advancement in Technology (GCAT), BANGALURU, India, 2019, pp. 15, doi: 10.1109/GCAT47503.2019.8978343.

[13] Pigou L., Dieleman S., Kindermans PJ., Schrauwen B. (2015) Sign Language Recognition Using Convolutional Neural Networks. In: Agapito L., Bronstein M., Rother C. (eds) Computer Vision - ECCV

2014 Workshops. ECCV 2014. Lecture Notes in Computer Science, vol 8925. Springer, Cham

[14] Bheda, Vivek and Dianna Radpour. “Using Deep Convolutional Networks for Gesture Recognition in American Sign Language.” ArXiv abs/1710.06836 (2017): n. pag.

[15] M.V, Beena. “Automatic Sign Language Finger Spelling Using Convolution Neural Network : Analysis.” (2017).

[16] Tolentino, Lean Karlo S. et al. “Static Sign Language Recognition Using Deep Learning.” International Journal of Machine Learning and Computing 9 (2019): 821-827.

[17] Umang Patel and Aarti G. Ambedkar, “Moment Based Sign Language Recognition for Indian Language” 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA).

[18] Bhargav Hegde, Dayananda P, Mahesh Hegde, Chetan C, “Deep Learning Technique for Detecting NSCLC”, International Journal of Recent Technology and Engineering (IJRTE), Volume-8 Issue-3, September 2019, pp. 7841-7843