Developing a Prototype to Translate Pakistan Sign Language into Text and Speech While Using Convolutional Neural Networking

Maria Naseem*
Department of Sciences & Humanities, National University of
Computer and Emerging Sciences, Lahore, Pakistan

Summaira Sarfraz
Heads of the Department of Sciences & Humanities, National University of
Computer and Emerging Sciences, Lahore, Pakistan

Ali Abbas
Department of Sciences & Humanities, National University of Computer
and Emerging Sciences, Lahore, Pakistan

Ali Haider
Electronics Department, University of engineering and technology, Taxila, Pakistan

Abstract
The purpose of the study is to provide a literature review of the work done on sign language in Pakistan and the world. This study also provides a framework of an already developed prototype to translate Pakistani sign language into speech and text while using convolutional neural networking (CNN) to facilitate unimpaired teachers to bridge the communication gap among the deaf learners and unimpaired teachers. Due to the lack of sign language teaching, unimpaired teachers face difficulty in communicating with impaired learners. This communication gap can be filled with the help of this translation tool. Research indicates that a prototype has been evolved that can translate the English textual content into sign language and highlighted that there is a need for translation tool which can translate the signs into English text. The current study will provide an architectural framework of the Pakistani sign language to English text translation tool that how different components of technology like deep learning, convolutional neural networking, python, tensor Flow, and NumPy, InceptionV3 and transfer learning, eSpeak text to speech help in the development of a translation tool prototype.
Keywords: Pakistan sign language (PSL), sign language (SL), translation, deaf, unimpaired, convolutional neural networking (CNN).
DOI: 10.7176/JEP/10-15-18
Publication date:May 31st 2019

1. Introduction
Language is a medium through which people communicate with one another. They interact, communicate and share their ideas using language. Language is described as the process of human communication, both the spoken and written languages including the use of words in a structured and manner (Abbas& Sarfraz, 2018).

People communicate with each other through different languages. But many people interact with one another through sign language (SL) because they cannot speak. Signs show different ideas and thoughts but not using spoken or written words. For certain purposes, they use specific signs and optical images. As a spoken language, sign language also varies in different countries and regions. Many of the advanced countries have established their sign languages. These countries have all the essential mechanisms of sign language, such as formation, grammatical constructions, and clear signals. America, China, and Arabic countries are the most important countries which have their sign languages. But in Pakistan, this process is under investigation. In Pakistan, many institutions are working to develop more standard sign language, but still, they are unable to assemble a graded model (Khan et al., 2015).

As Sulman and Zuberi (2000) mentioned in their study that Anjuman Behbood-e-Samat-e-Atfal (ABSA), Karachi, Sir Syed Deaf Association (SDA), Rawalpindi, National Institute of Special Education (NISE), Islamabad and Pakistan Association of Deaf (PAD), Karachi are the four leading institutions playing a vital role in developing PSL. There are 72 million deaf people in the world. They want to interact with other people through a language (Abbas & Sarfraz 2018).

Many countries have developed instruments and software which facilitate the deaf people. Tools are accessible for translating speech and text into sign language. For instance, "Pro deaf" is a versatile application which converts the text to ASL. Foong and Wibowo (2008) developed another instrument named S2V (sign to voice). This tool converts the signs of the language into voices. This translation instrument will be beneficial for deaf people. The purpose of this work was to explore the problems and difficulties faced by teachers while they

communicate with them.

A lot of research has been done from the last three decades in the world of sign language and sign identification. From the previous few years, numerous methods of potential applications for HCI (human-computer interaction) have been suggested. Two patterns have been viewed for the categorization of sign language recognition. First one is a glove-based system. It can be used for measuring different parameters of gestures such as hand and position of the finger, angles, and identification of the tip, etc. But the major problem with the signer is inopportuneness. The way the interaction takes place between the signer and the system is sophisticated and less natural. The second category turned around to express vision-based hand recognition system with the use of machine vision and image processing techniques (Radpour, 2017) but these are different from each other in their models.

Recently, sign language became the limelight of the researchers and researches are being done on it which turns limited vocabulary of the gesture recognition for operations and has brought a gradual transition from isolated to continuous and static to dynamic. In the present situation, the human-machine interactive system assists in communicating between deaf and the hearing people in real-world situations. Many researchers have developed methods to improve the accuracy of recognition such as Artificial Neural Networks, HMM (Hidden Markov model), and Kinect platform, and have also evolved the effective algorithms for segmentation, recognition, classification and pattern matching (Vijay, 2012).

Between the community of deaf and the general public, a real-time sign language translator is an essential milestone for communication. Based on convolutional neural networking, the development and implementation of a fingerspelling translator of an American sign language (ASL) were presented. A pre-trained architecture of Google Net was applied and was trained on the dataset of ILSVRC2012 using the ASL datasets of Massey University and Surrey University to apply transfer learning to this task. A robust model was produced that classifies letters a-e correctly with first-time users (Garcia& Viesca, 2016).

According to Bheda (2017), convolutional neural networks have been hugely successful for image recognition and classification problems. This technique was implemented successfully for the recognition of human gestures in recent years. Notably, a lot of work has been done in sign language recognition using deep CNN (convolutional neural networking) with input recognition. This process was made much easier for developing characteristic depth and motion profile with the use of cameras that sense depth and contour, for each gesture of sign language (Agarwal & Thakur, 2013).

The technology of depth sensing use is quickly growing in popularity, and the other tools have been incorporated into the process that proved successful, developments such as custom designed and color gloves have been used to facilitate the process of recognition and to make the step more efficient feature extraction by making specific gestural units easier to classify and identify (Dong, Leu& Yin, 2015).

Until now, different methods of automatic sign language recognition were not able to make use of the depth sensing technology which is widely available today. Previous works made use of just a simple camera technology to generate datasets of simple images, without depth or contour information, only the pixels present but classifying images of ASL letter gestures using CNN has had some success (Garcia & Viesca, 2016) but using a Google Net architecture which is pre-trained.

The sings for all the alphabets from A to Z are being recognized using the architecture of combinational neural networks. The benefit of using this algorithm is high processing speed which can produce results in a real-time manner. The processing speed is also increased due to neural network architecture so that researchers can go for words and sentences gesture in the future (Mekala, Gao, Fan, & Davari, 2011).

In the field of deep learning with recent advances, neural networks can have far-reaching applications and implications for sign language interpretation as (Radpour, 2017) present a method to classify images of both the letters and digits in American sign language using deep convolutional networks and Abbas & Sarfraz (2018) have highlighted the similar gap and developed a prototype to convert texts and speech into Pakistani sign language in order to facilitate the communication between the deaf and the unimpaired people.

A lot of work has been done to overcome this gap but our enthusiasm for this project is to help people who are impaired to have a normal conversation with the unimpaired teachers in the classroom environment through a program that inputs sign language into speech and text form and to overcome the gap between the deaf and dumb and unimpaired teachers. As (Badhe & Kulkarni, 2015) mentioned in his study that the process would go the other way around. The current study aims to provide a basic program that translates words into Pakistani sign language. If the program is elaborated after our project, it can be a program that allows people to have a free conversation of whether a person is disabled or not. As (Badhe & Kulkarni, 2015) mentioned in their study, this kind of program is significant for it and can be even more elaborated to a program that translates conversation (sentences) into Pakistani sign language as well means as now we are working on words, but we can go for sentences gestures in future. So, the current study aims to remove a communication barrier between the deaf and normal people.

2. Literature Review

There is a massive gap between communication between normal and disabled people. So, sign language can help in communicating with disabled people. Different type of gestures is being used having numerous forms in sign language. Similarly, the sign languages vary in different areas and nearly 138 (Pakistan sign language, 2015) sign languages are acknowledged until these days. British American Sign Languages are based on the English language whereas the Chinese and Indian sign languages have also been developed. The grammar of the gesture-primarily based sign languages ranges from the grammar of written and spoken languages because the gesture-based languages are based on shapes and ideas, while on the other hand, spoken and written languages encompass words and grammar regulations. Both the languages have different grammatical structures for this particular reason (Debevc et al., 2014; Marschark et al., 2004). The information technology field has kept on influencing human life strongly. Different technologies, tools, and devices have been developed to assist humanity in solving various issues. Human beings have tried to bridge the communication gap between the deaf and normal individuals using information technology. The notion behind such IT-based tools facilitates the deaf to communicate in a better way with the unimpaired people and vice versa. There can be various situations in which such IT-based tools can be proved helpful to eliminate this communication gap (Khan et al. 2015).

Supreme urbanized nations have addressed the problems in their listening to-impaired people by throwing schemes related to information technology for minimizing the gap among the deaf and ordinary people (Abbas & Sarfraz, 2018). Pakistan communication via gestures is a phonetically underneath researched in the non-appearance of any organized data around the language substance, syntax, and devices and administrations for discussion. For this reason, the primary contribution of this study is to give awareness on the tasks for bridge this communication gap for Pakistani deaf community by the use of the prevailing literature, and to recommend an information technology dependent architectural framework to discover important mechanisms to construct programs which may also assist bridging the gap among the deaf and ordinary people of the country (Khan et al. 2015).

Modern-day technologies are all about cell phone computing, gesture-based environments, and cloud computing. The world is bounding into the gestural mechanism. IT-based companies like Microsoft, Google, and Leap Motion are introducing devices like Kinect, Google Glass and Leap Motion controller (Potter, 2013), so, development in the technology may be used to benefit deaf people.

For deaf children, communication is enormous scrap. It gets impossible for them to blend with society because they cannot communicate in a normal way. In the academic scenario, the learning environment for deaf learners is not always parallel compared to normal learners. One of the most effective methods for the deaf is communicating through Sign Language. The communication through signing can't be comprehended through the gesture-based interface by a specialist or other fellows. It creates difficulties between deaf and normal person communication (Mindess, 2014).

In upstanding frameworks and research for gesture-based communication Translation and acknowledgment gadget, each picture based and sensor-based strategies are utilized. Most recent studies concentrated more on the hand gesture identification system because of the software in HCI, Robotics, game-based learning, and sign language identification programs and systems. Different processes and algorithms from pc vision community have been used (Itkarkar & Nandy, 2014).

Despite being a turf imprecisely discovered, projects have recently started appearing related to this topic. As Noberto et al. (2015) mentioned in his study that a sit-down of virtual sign project potentialities, there is an educational game in development which goal is to facilitate the sign language learning process. The game induces the user to learn sign language through an innovating and fun model while interacting with a virtual world, using gesture execution of sign language to concretize the common goals and accomplish the missions along with the game. This scheme becomes even a superior asset for the development of Portuguese sign language (Escudeiro et al., 2015). Indian Sign Language (ISL) Translation gadget for sign Language learning, the ISL translation system makes use of microphone or USB camera to get pictures or continuous video image (from normal people) which may be interpreted by the application. Attained expressions are expected to be a translation, scale, and rotation invariant. In this method, the steps of interpretation are the acquisition of pictures, detection of binarized type hand shape and function extraction. The GUI software is showing and sending the message to the receiver. This structure makes normal people speak effortlessly with deaf/dumb persons (Jose, Priyadharshni, Anand, Kumaresan & Kumar, 2013).

The final signing avatar device essentials can transform preferred spoken words to sign language inevitably. An inordinate deal of work has been positioned into this translation effort, with various levels of success, in different variations of sign languages, e.g., American, Greek, South African, Arabic, Spanish, Italian, Japanese, British, and the Netherlands (Clymer et al., 2012).

Sign language (SL) is a foundational medium between the individuals who are not able to hear well. This method is also called optical motion dialect. People who are not able to understand well use this method of sign language as the main channel for communication. Every country has its sign language. For example, China,

America India, and Pakistan have their sign languages which are known as Chinese sign language, American Sign Language, Indian sign language and Pakistani sign language. Many progressive countries give a lecture on this issue. They organize different project activities including information technology to eliminate the gap among a deaf and the normal people. Many surveys have been conducted on this issue in central and south Asia. However, in Pakistan, this method is under investigation because there is no structured or organized information about language grammar, contents and instruments for transmission. (Khan et al., 2015). But till now the main point of this research is to discuss the problems to make a way between the normal and deaf community and after using the literature. They suggest many items to build a bridge.

The rules of sign language are different from the rules of spoken and written languages. The sign language is based on shapes, and written language is based on word formation and some basic rules of grammar (Debevc et al., 2015). The information technology has the main impact on our lives a human makes many things that we use in our life.

In India, exceptional work has been done to make a way between abnormal people. They have facilitated them through different methods. The major hurdle that these people face is, they cannot communicate with other people. The virtual tongue or effective tongue has grown (Kumar et al., 2014).

Pakistani researchers are also working to create different instruments to facilitate the impaired people. For the transformation of American Sign Language into text form, a sensory glove was developed by (Mehdi & Khan, 2002). This system was created to make communication possible between the normal and affected people. This application is called "talking hands." Artificial neural networks were used to accept the signals from the sensory glove. These values are based on 24 English language alphabets, and two punctuation signs were introduced. Through this method, a deaf person can write a complete sentence (Mehdi & Khan, 2002, 2004). Bukhari et al. (2015) also have conducted a similar study, for communication between deaf and impaired people. These sensory gloves have sensors which recognize the signals of speech through finger movement. This instrument is very useful, and it converts the alphabets into text and speech (Razing & Latif (2016). Work has been conducted on to translate Pakistan sign language into text formation. They used a leap motion device technology which is based on two units; one trains the system for translation, and the other accumulates the information using leap motion device. This module is called "communication module."

There is also a work conducted by Fatima and Huma (2011). The instrument is based on signs. These signs present the text on a computer alliance. It is the visual based work and done first time in Pakistan without using any glove.

There are many progressed works in Pakistan for deaf people who want to talk with normal people. A fuzzy classifier is a tool created by Kauser et al. (2008). This tool identifies the signs of deaf people or method colored gloves are used to recognize the fingertips. The validity of this method is 95 percent.

Many researchers are making efforts to fill this gap between the deaf and normal people. Khan et al. (2015) foregrounded the issues taking the shape of a framework for gesture-based communication interpretation in Pakistan. They suggested an architectural framework that helps the deaf people to translate English or Urdu texts or speech into Pakistani sign language (Khan et al. 2015).

Badhe and Kulkarni (2015) conducted research which aimed to develop a tool to help people who are impaired or hard of hearing have a normal conversation with other people for this we are going to establish a program that input sign language in text form and to allow the other person to understand.

Garcia and Viesca (2016) applied a pre-trained architecture of Google Net which is trained on the dataset of ILSVRC based on convolutional neural networking and also used the ASL datasets of Massey University and Surrey University to apply transfer learning to this task. They produced a robust model that classifies letters a-e correctly with first-time users. For all ASL Letters, a fully generalizable translator can be produced.

The technology of depth sensing use is quickly growing in popularity, and the other tools have been incorporated into the process, that proved successful, developments such as custom designed, color gloves have been used to facilitate the process of recognition and make the step more efficient feature extraction by making specific gestural units easier to classify and identify (Dong, Leu & Yin, 2015).

Until now, different methods of automatic sign language recognition were not able to make use of the depth sensing technology which is widely available today. Previous works made use of just a simple camera technology to generate datasets of simple images, without depth or contour information, only the pixels present but classifying images of ASL letter gestures using CNN's have had some success Garcia & Viesca (2016) but using a GoogLeNet architecture which is pre-trained (Cui, Liu, & Zhang, 2017) have proposed a deep structure with recurrent convolutional neural network for constant sign language identification. We have designed a staged optimization procedure for educating our deep neural network structure.

Deep convolutional neural networks (CNN) have accomplished breakthroughs in gesture identification (Gupta et al., 2015). And sign recognizing (Natlia et al., 2014), and recurrent neural networks (RNNs) has also proven significant results while learning the dynamic temporal dependencies in sign recognizing (Pigou et al. 2018).

Foong (2018) provides a system prototype which is capable of automatically identify sign language to assist normal people in speaking more efficiently with the hearing or speech impaired people. The sign to Voice machine prototype, S2V, was developed using Feed Forward Neural network for two-series signs detection. Different sets of universal hand gestures have been captured from the video camera and applied to train the neural network for classification purpose. The experimental results have proven that the neural network has accomplished excellent result for sign-to-voice translation.

In the present situation, the human-machine interactive system assists communication between deaf and the hearing people in real-world situations. Many researchers have developed methods to improve the accuracy of recognition such as Artificial Neural Networks, HMM, and Kinect platform, and have also evolved the effective algorithms for segmentation, recognition, classification and pattern matching (Paranjape Ketki Vijay, 2012).

After researching, we came to know that there are already many existed apps where people can input a sentence, or a phrase and a virtual robot translate those phrases and sentences into sign language with the hand gesture, such as "pro deaf translator" on Android's Play Store. With time iPhone users "ASL translator" where we see that androids and apple apps differ, another app in the app store for when a text is input there are videos for the deaf to watch, so for our research, we will keep these differences in our mind while making other apps.

3. Architectural Framework

This research presents an architectural framework of Pakistan sign language recognition tool developed to convert the sign language to text and voice. By using this particular tool, the sign language can be converted the sign into text by using conventional picture classification program that utilizes Google's Machine Learning library, TensorFlow and a pre-prepared Deep Learning Convolutional Neural Network model called Inception. An architecture of this recognition has been presented in Figure 1. The system accepts the input in the form of a picture from camera and translates sign into Text as well as voice.
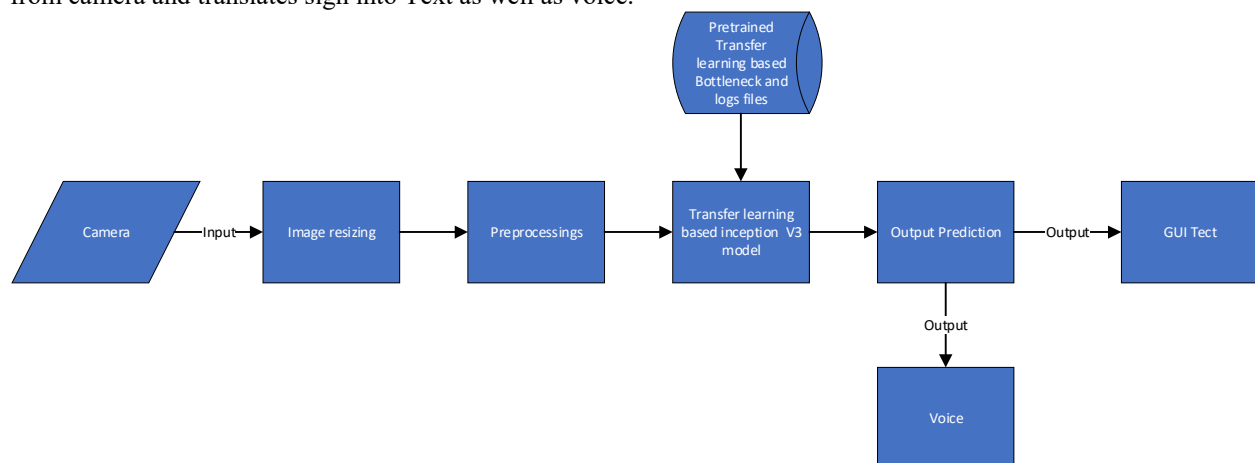

Figure 1

3.1 Tool Specifications

There are unique technological tools and structures used to build up the Pakistan sign Language Recognition Tool.

3.1.1 Programming language

Python is a general-purpose, versatile and popular programming language. Python is object-oriented language making it very attractive for accelerated Application Development. Python's simple syntax value to understandability and therefore lessens the cost of program preservation. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms.

3.1.2 OpenCV

OpenCV (Open Source Computer Vision) is a library of programming functions mainly aimed at real-time computer vision. In simple language, it is a library used for Image Processing. It is primarily used to do all the operation related to Images. The library has more than 2500 optimized algorithms, which includes a comprehensive set of both classic and state-of-the-art computer vision and machine learning algorithms.

3.1.3 TensorFlow

TensorFlow is a framework that represents complex computations as graphs; this makes it easier for analysis of models; multi-dimensional arrays called Tensors are used to do the same. Tensor Flow was used for training the PSL dataset based upon deep learning. It is used for research as well as development proposes at the same time.

3.1.4 NumPy

NumPy is a Python package which stands for 'Numerical Python.' It was used for carrying out numerical

computations. It offers a wide range of logarithmic processing function such as differentiation, integration. It supports for large, multi-dimensional arrays and matrices, along with an extensive collection of high-level mathematical functions to operate on these arrays.

3.1.5 InceptionV3 and transfer learning

Inceptionv3 model has been pre-trained for the ImageNet Large Visual Recognition Challenge using the data from 2012, and it can differentiate between 1,000 different classes, like Dalmatian, dishwasher, etc. We would use this pre-trained model for sign recognition. This model is trained with large dataset extending up to millions of pictures using multi-layers' convolution neural network (CNN) and artificial intelligence. Usually, each layer has its feature extraction as could be seen in figure2. Signals travel from the first layer (input) to the last one (output), possibly after traversing the layers multiple time. Edges in the earlier segments, Shapes in the middle layer and some high-level data specific features in the later layers. Hence inceptionV3 could apply to a different but related problem. These models could be used to Pakistan sign language recognition just by retraining the top upper layer only and still achieving the remarkable accuracy. Transfer learning is an exploration problem in machine learning that centers around putting away information picked up while solving one problem and applying it to an alternate yet a related issue. Inception v3 architecture model trained on ImageNet images was prepared a new top layer that can recognize other classes of Pakistan sign language. As figure 2 explained multi-layers' convolution neural network (CNN) and artificial intelligence. Usually, each layer has its feature extraction.
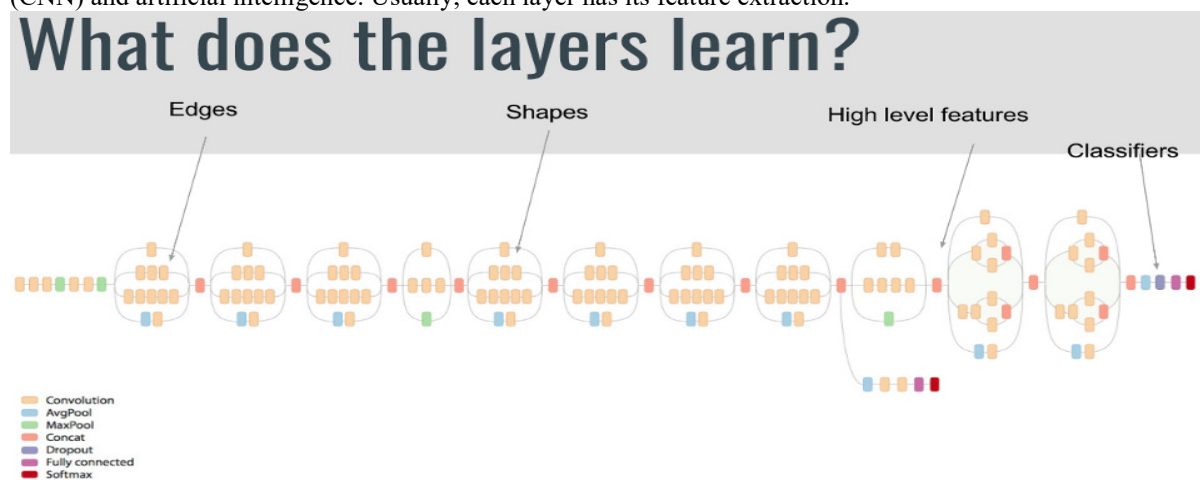
Figure 2

3.1.6 Espeak Text to Speech

ESpeak is a compact open source software speech synthesizer for English and other languages, for Linux and Windows. It converts the text to speech.

3.2 Transfer Learning Process

The inceptionv3 model's top layer was retrained with the 15 different signs and about with 21,000 pictures using, and the bottleneck file created as a result was graphed via tensor flow and saved for further recognition purposes. Whenever we are having the dataset of millions of pictures, then we try to train the model from scratch but in the case of a few thousand pictures transfer learning is suitable.

3.3 Training Process

3.3.1 Dataset

The dataset was made by making a video and then extracting its frame via Matlab's image processing toolbox. The dataset of each sign was about 1500 to 2000 picture. The overall number of the picture was about 21000 as below figure 3 explained.
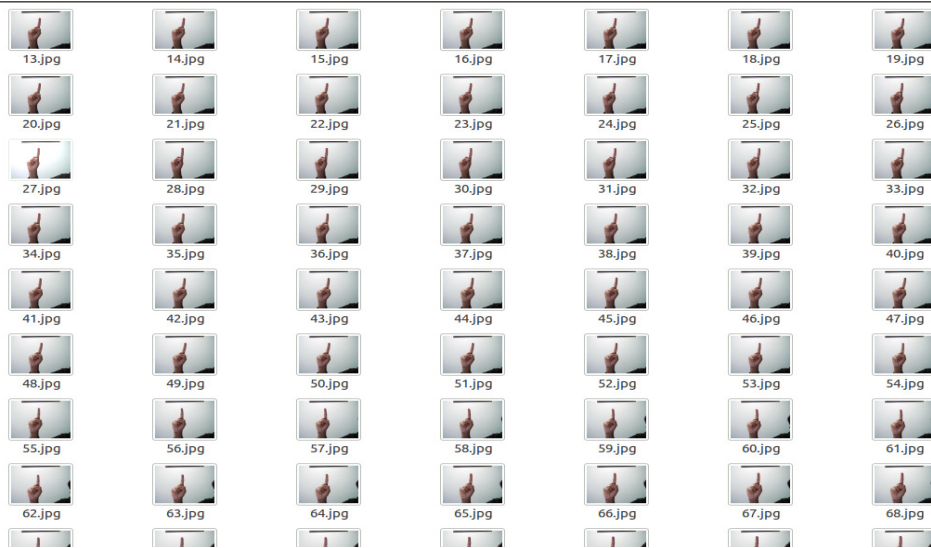
Figure 3

3.3.2 Training

The training was based on transfer learning technique of deep learning. It was used to retrain the pertained inception V3 deep learning model. The feature from each picture was extracted, and the essential elements were saved as bottlenecks to save computational expense. These trained models were used to generate TensorFlow graph which would be used in the recognition stage as in figure 4.

Figure 4

3.3.3 Recognition

The recognition was based on the capturing of the picture of the sign, and then that would be fed to the deep learning classifier model, and the best-suited result would be generated as explained in figure 5.
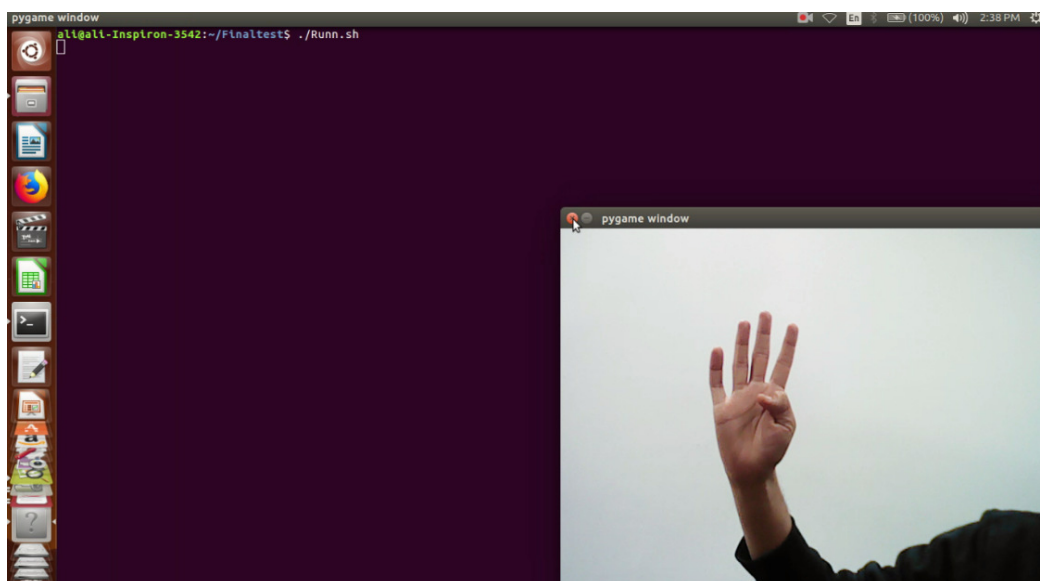
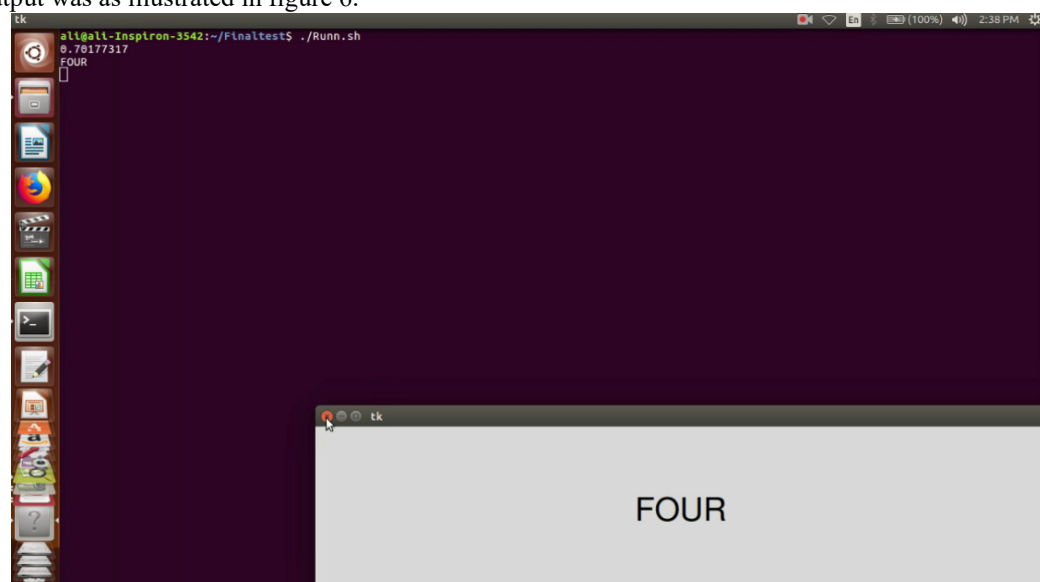*Figure 5*

and Output was as illustrated in figure 6.

Figure 6

4. Conclusion

The current research provides a literature review of work done by different researchers in the world especially in Pakistan and to overcome the communication gap between the deaf and unimpaired people. The literature review also shows that a lot of work has been done to overcome this gap but the current study aims develop a prototype in order to help people who are impaired to have a normal conversation with the unimpaired teachers in the classroom environment through a program that inputs sign language into speech and text form and to overcome the gap between the deaf and dumb and unimpaired teachers. The process will go the other way around.

Our goal is to provide a basic program that translates Pakistan sign language into words and speech. If our application is elaborated after our project, it can be a program that allows people to have a free conversation of whether a person is disabled or not. This kind of program is significant for it can be even more elaborated to a program that translates conversation (sentences) into Pakistani sign language as well means as now we are working on words, but we can go for sentences gestures in future.

The current research has highlighted the gap and evolved a prototype to translate English texts and speech into Pakistani sign language to bridge the communication gap between deaf and unimpaired people. In the future, a fully functional tool can be developed to translate text and speech into PSL and make the framework supportive for various platforms which are web and mobile-based technologies. In the field of education, to determine the feasibility of this tool, surveys and interviews will be conducted from the teachers teaching deaf learners in separate

classes or in the mixed classroom where deaf and unimpaired learners study together and also from the experts. So, the main goal of the current study is to minimize the communication hurdle between the deaf and the normal people. In the future, the researcher will develop a fully functional tool to translate PSL sentences into speech and text and make the framework supportable for multiple platforms, i.e., web and mobile base. Currently developed sign translation tool can only translate few words of Pakistan sign language into text and speech because it is a prototype whereas there is a need to create an app or handheld device to overcome the communication barrier while making this sign translation tool user-friendly. In the development of the handheld version of the tool, the system on chip (SoC) devices are being considered. The other models are also being trained to be compatible with low-cost computing devices with dependable accuracy.

References

Abbas, A., & Sarfraz, S. (2018). Developing a Prototype to Translate Text and Speech to Pakistan Sign Language with Bilingual Subtitles: A Framework. *Journal of Educational Technology Systems*, *47*(2), 248-266.

Agarwal, A., & Thakur, M. K. (2013, August). Sign language recognition using Microsoft Kinect. In *Contemporary Computing (IC3), 2013 Sixth International Conference on* (pp. 181-185). IEEE.

Badhe, P. C., & Kulkarni, V. (2015, November). Indian sign language translator using gesture recognition algorithm. In *Computer Graphics, Vision and Information Security (CGVIS), 2015 IEEE International Conference on* (pp. 195-200). IEEE.

Bheda, V., & Radpour, D. (2017). Using deep convolutional networks for gesture recognition in American sign language. *arXiv preprint arXiv:1710.06836*.

Bueno, J. (2010). Requisitos para um ambiente de comunicação como ferramenta de apoio à alfabetização bilingue de crianças surdas.

Bukhari, J., Rehman, M., Malik, S. I., Kamboh, A. M., & Salman, A. (2015). American Sign Language translation through sensory glove; SignSpeak. International Journal of U- and E-Service, Science and Technology, 8(1), 131–142. doi:10.14257/ ijunesst.2015.8.1.12

Caliwag, A., Angsanto, S. R., & Lim, W. (2018, July). Korean Sign Language Translation Using Machine Learning. In *2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN)* (pp. 826-828). IEEE.

Clymer, E., Geigel, J., Behm, G., & Masters, K. (2012). Use of signing avatars to enhance direct communication support for deaf and hard-of-hearing users. *Unpublished Manuscript, NTID Center on Access Technology, Rochester Institute of Technology, National Technical Institute for the Deaf*.

Cui, R., Liu, H., & Zhang, C. (2017, July). Recurrent convolutional neural networks for continuous sign language recognition by staged optimization. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Debevc, M., Milošević, D., & Kožuh, I. (2015). A comparison of comprehension processes in sign language interpreter videos with or without captions. *PloS one*, *10*(5), e0127577.

Dong, C., Leu, M. C., & Yin, Z. (2015). American sign language alphabet recognition using microsoft kinect. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (pp. 44-52).

Escudeiro, P., Escudeiro, N., Reis, R., Lopes, J., Norberto, M., Baltasar, A. B., ... & Bidarra, J. (2015). Virtual Sign–A Real Time Bidirectional Translator of Portuguese Sign Language. *Procedia Computer Science*, *67*, 252-262.

Farhadi, A., Forsyth, D., & White, R. (2007, June). Transfer learning in sign language. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on* (pp. 1-8). IEEE.

Fatima, A., & Huma, K. (2011). Image based Pakistan sign language recognition system

Foong, O. M., Low, T. J., & Wibowo, S. (2008). Hand gesture recognition: sign to voice system (S2V). *World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering*, *2*(6), 1794-1798.

Garcia, B., & Viesca, S. (2016). Real-time American Sign Language recognition with convolutional neural networks. *Convolutional Neural Networks for Visual Recognition*.

Gavrila, D. M. (1999). The visual analysis of human movement: A survey. *Computer vision and image understanding*, *73*(1), 82-98.

Halim, Z., & Abbas, G. (2015). A Kinect-based sign language hand gesture recognition system for hearing-and speech-impaired: a pilot study of Pakistani sign language. *Assistive Technology*, *27*(1), 34-43.

Itkarkar, R. R., & Nandy, A. K. (2014). A Study of Vision Based Hand Gesture Recognition for Human Machine Interaction. *International Journal of Innovative Research in Advanced Engineering (IJIRAE) ISSN*, 2349-2163.

Jose, M. J., Priyadharshni, V., Anand, M. S., Kumaresan, A., & Kumar, N. M. (2013). Indian Sign Language (ISL) Translation System For Sign Language Learning. *International Journal of Innovative Research and Development*, *2*(5).

Kausar, S., Javed, M. Y., & Sohail, S. (2008, August). Recognition of gestures in Pakistani sign language using

fuzzy classifier. In Proceedings of the 8th conference on Signal processing, computational geometry and artificial vision (pp. 101-105). World Scientific and Engineering Academy and Society (WSEAS).

Khan, N. S., Abid, A., Abid, K., Farooq, U., Farooq, M. S., & Jameel, H. (2015). Speak Pakistan: Challenges in Developing Pakistan Sign Language using Information Technology. *South Asian Studies*, *30*(2), 367.

Kožuh, I., Hintermair, M., & Debevc, M. (2014, July). Examining the characteristics of deaf and hard of hearing users of social networking sites. In *International Conference on Computers for Handicapped Persons* (pp. 498-505). Springer, Cham.

Kute, R. S., Vyas, V., & Anuse, A. (2019). Component-based face recognition under transfer learning for forensic applications. *Information Sciences*, *476*, 176-191.

Marschark, M., Convertino, C., McEvoy, C., & Masteller, A. (2004). Organization and use of the mental lexicon by deaf and hearing individuals. *American Annals of the Deaf*, *149*(1), 51-61.

Marschark, M., Sapere, P., Convertino, C., Seewagen, R., & Maltzen, H. (2004). Comprehension of sign language interpreting: Deciphering a complex task situation. *Sign Language Studies*, *4*(4), 345-368.

McKim, C. A. (2017). The value of mixed methods research: A mixed methods study. *Journal of Mixed Methods Research*, *11*(2), 202-222.

Mehdi, S. A., & Khan, Y. N. (2002). Sign language recognition using sensor gloves. In Proceedings of the 9th International Conference on Neural Information Processing, 2002. ICONIP'02 (Vol. 5, pp. 2204–2206). Piscataway, NJ: IEEE

Mekala, P., Gao, Y., Fan, J., & Davari, A. (2011, March). Real-time sign language recognition based on neural network architecture. In *System Theory (SSST), 2011 IEEE 43rd Southeastern Symposium on* (pp.

Mindess, A. (2014). *Reading between the signs: Intercultural communication for sign language interpreters*. Nicholas Brealey.

Molchanov, P., Gupta, S., Kim, K., & Kautz, J. (2015). Hand gesture recognition with 3D convolutional neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (pp. 1-7).

Molchanov, P., Tyree, S., Karras, T., Aila, T., & Kautz, J. (2016). Pruning convolutional neural networks for resource efficient transfer learning. *CoRR, abs/1611.06440*.

Neverova, N., Wolf, C., Taylor, G. W., & Nebout, F. (2014, September). Multi-scale deep learning for gesture detection and localization. In *Workshop at the European conference on computer vision* (pp. 474-490). Springer, Cham.

Norberto, M. T. VirtualSign Translator in Serious Games.

Norberto, M., Lopes, J., Escudeiro, P., Escudeiro, N., Reis, R., Barbosa, M., ... & Baltazar, A. B. (2015). Virtual Sign—Using a Bidirectional Translator in Serious Games. *China-USA Business Review*, 261.

P. Parmar, A. P., Chitaliya, N. G., Gesture Recognition System for Indian Sign Language on Smart Phone. International Journal of Advanced Research in Computer and Communication Engineering. 5(2), 376-379, 2016.

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health and Mental Health Services Research*, *42*(5), 533-544.

Paranjape Ketki Vijay, N. N. (2012). *Recent Developments in Sign Language Recognition: A Review.* Pune-48, India: Vishwakarma Institute of Information Technology.95-199). IEEE

Pigou, L., Van Den Oord, A., Dieleman, S., Van Herreweghe, M., & Dambre, J. (2018). Beyond temporal pooling: Recurrence and temporal convolutions for gesture recognition in video. *International Journal of Computer Vision*, *126*(2-4), 430-439.

Pinker, S. (2003). *The language instinct: How the mind creates language*. Penguin UK.

Potter, L. E., Araullo, J., & Carter, L. (2013, November). The leap motion controller: a view on sign language. In *Proceedings of the 25th Australian computer-human interaction conference: augmentation, application, innovation, collaboration* (pp. 175-178). ACM.

Quigley, S. P., & Paul, P. V. (1984). ASL and ESL?. *Topics in early childhood special education*, *3*(4), 17-26.

Radpour, V. . (2017). *Using Deep Convolutional Networks for Gesture Recognition in American Sign Language .* New York: State University.

Raziq, N., & Latif, S. (2016). Pakistan sign language recognition and translation system using leap motion device. In Advances on P2P, Parallel, Grid, Cloud and Internet 18 Journal of Educational Technology Systems 0(0) Computing (pp. 895–902). Cham, Switzerland: Springer. doi:10.1007/978-3-31949109-7_87

Sankar Kumar, S., Jenitha, J., Narmadha, I., & Suganya, A. (2014). An embedded module as "Virtual Tongue" for voiceless. International Journal of Information Sciences and Techniques,

Shaha, M., & Pawar, M. (2018, March). Transfer Learning for Image Classification. In *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)* (pp. 656-660). IEEE.

Sulman, D. N., & Zuberi, S. (2000). Pakistan sign language–a synopsis. Pakistan

Vijay, P. K., Suhas, N. N., Chandrashekhar, C. S., & Dhananjay, D. K. (2012). Recent developments in sign language recognition: A review. *Int J Adv Comput Eng Commun Technol*, *1*, 21-26.

Wasim, M., Siddiqui, A. A., Shaikh, A., Ahmed, L., Ali, S. F., & Saeed, F. (2018). Communicator for Hearing-Impaired Persons using Pakistan Sign Language (PSL). *INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS, 9(5), 197-202.*

Wium, A. M., & Louw, B. (2018). Mixed-methods research: A tutorial for speech-language therapists and audiologists in South Africa. *South African Journal of Communication Disorders*, *65*(1), 1-13.

Young, T. Y. (1994). *Handbook of pattern recognition and image processing (vol. 2): computer vision*. Academic Press, Inc.

Zohrabi, M. (2013). Mixed Method Research: Instruments, Validity, Reliability and Reporting Findings. *Theory & practice in language studies*, *3*(2).