e-ISSN: 2250-0758 | p-ISSN: 2394-6962
Volume-12, Issue-6, (December 2022)
https://doi.org/10.31033/ijemr.12.6.16

Real-Time Sign Language Translator

Rathnayake R.K.D.M.P.[1], Wijekoon W.M.S.T[2], Rajapakse K.G.[3], Rasanjalee K.A.[4] and Dilshan De Silva[5]
[1]Sri Lanka Institute of Information Technology, Malabe, SRI LANKA
[2]Sri Lanka Institute of Information Technology, Malabe, SRI LANKA
[3]Sri Lanka Institute of Information Technology, Malabe, SRI LANKA
[4]Sri Lanka Institute of Information Technology, Malabe, SRI LANKA
[5]Sri Lanka Institute of Information Technology, Malabe, SRI LANKA

[2]Corresponding Author: it19210452@my.sliit.lk

ABSTRACT

Sign language is so widespread that people with hearing/speech impairments are familiar with it, and others are unfamiliar with it. As a result, there is a significant communication gap between people with speech and hearing impairments and the rest of the general population. A human sign language interpreter is a common solution for bridging this gap. However, because the number of sign language interpreters is small in comparison to the number of deaf and mute people in the world, some deaf and mute people cannot afford to use a human interpreter all of the time when communicating with others. This communication must be automated so that the deaf-mute community does not rely on human interpreters. This paper focuses on developing a system that can translate American Sign Language into words/sentences and vice versa in real-time, with some extra features that will help remove the communication barrier between ordinary and hearing/talking impaired people. The main function is to detect and identify sign language performed by the user. Initially, the system is trained to detect and identify signs using object detection and motion-tracking techniques. A convolutional neural network model was trained on a manually created ASL data set. The identified signs are then translated into English to form grammatically correct sentences. A text-to-text transformer built on an encoder-decoder architecture is used to detect grammatical errors and provide correct sentences. To further improve effectiveness, the system incorporates a feature to translate an image containing English text into American Sign Language. Furthermore, the system consists of a voice-to-sign language translator and a virtual sign keyboard. The methodology has been explained in further sections.

Keywords-- American Sign Language, Convolutional Neural Network, Grammatically Correct Sentences, Object Detection, Real-Time, Speech and Hearing Impairments, Voice-to-Sign Language

I. INTRODUCTION

Our society contains a considerable population of hearing and speech-impaired people. According to the World Health Organization (WHO), the number of hearing-impaired populations has reached 400 million [19]. The need for a real-time sign language translator cannot be underestimated with such a high number of individuals. Several sign language translators are being used nowadays. Even though it is a question of whether they fulfill the needs of communication between deaf-mute people and ordinary people. The main form of communication is done verbally but unfortunately, not all of us can communicate verbally. Even so, we cannot exclude them from our society. These disabilities could occur because of several reasons. Some could be born with speech and hearing impairments. In contrast, others could lose their hearing ability due to exposure to loud noises, tumors, ear malformation, genetics, infectious diseases, use of drugs, certain accidents, or aging [20][21], etc. When a person is born with hearing impairments, it directly affects his speech ability as well. Therefore, people with hearing and speech impairment problems are considered a burden to society

The traditional way of communication between deaf-mute people and normal people is either using pen and paper to write down the idea or using body language. Using pen and paper to communicate an idea is time-consuming also it could get boring. Body language can sometimes be awkward and may not convey the correct message. As of today, we live in a busy world. Therefore, communication cannot be slowed down. With the emerging technologies, solutions should be built to reduce the gap between deaf-mute individuals and normal people.

Because of this communication barrier, normal people who can hear and speak find it difficult to interact with them. Therefore, people with disabilities are rejected from jobs [22]. They cannot access certain information. They may be excluded from certain social or public services and functions. There are several challenges they must face due to their inability to communicate like normal people.

There are about 138-300 distinct sign languages currently used in the world [23]. Hand gestures and facial expressions are mainly used to perform the sign. Sign language does have its own grammar rules and sentence structures. The word order of a sentence in English and

e-ISSN: 2250-0758 | p-ISSN: 2394-6962
Volume-12, Issue-6, (December 2022)
https://doi.org/10.31033/ijemr.12.6.16

American Sign Language is different [24]. Further, Sign language has signs only for the root words in English. Also, there are no signs for prepositions in English [12]. These differences must be considered when building a sign language translator as it will improve the communication experience between a sign language user and a non-sign language user.

Further, people who have used sign language to communicate for most of their life may find it difficult to read. Sign language translators can support them by converting texts into sign language, especially texts on posters, banners, road signs, etc. As more people work with their smartphones, it is also beneficial that the translator can be used as both a web and a mobile application. These issues or aspects require more attention when discussing sign language, its users, and the need for a sign language translator.

II. LITERATURE SURVEY

As most people are aware, many people worldwide have done much research on sign language translation. One such research was Sign Language Interpreter Research Using Image Processing and Machine Learning by Omkar Vedak, Prasad Zavre, Abhijeet Todkar, Manoj Patil from Student, Department of Computer Engineering, Datta Meghe College of Engineering, Mumbai University, Airoli, India. There they implemented a system that translates Indian Sign Language into English. Here, the proposed system would capture images of stationary characters using a web camera. Then the images are processed to get enhanced features. Then an algorithm they wrote translates the sign language into English text. This text will later be converted into speech. However, the proposed system only had 26 characters pre-registered into the system, which limited them to very few hand characters, they also had to take about 4800 images for training and 1200 images for testing. And the accuracy of the system was about 88%. The proposed system was planned for implementation on mobile applications [1].

Mahesh Kumar NB [17] recognize gestures using Linear Discrimination Analysis (LDA). They used MATLAB to extract features from the dataset. [14] has conducted fundamental research on sign language datasets using the CNN algorithm to obtain satisfactory results from the training and testing of the dataset. The proposed system [18] uses selfie sign language to process the images and is tested using a method called 'Stochastic pooling'. The CNN model was used to train the dataset in [18]. It was tested using various window sizes and different batches of images. The output accuracy was 92.88%, higher than the other investigated methods.

According to research [15] done in Sri Lanka, to translate Sinhala sign language to text and audio outputs, they have used AI (Artificial Intelligence) and ML (Machine Language). Their development includes a feature to translate verbal Language to Sri Lankan sign language. The development of this translator is divided into 4 main parts. The initial step is to detect hand gestures. Pre-trained models are used to capture the hand signs. Classification and translation of detected hand signs is the next step. This step is performed using image classifiers. Afterward, the identified sign will be given to the user in the form of either a text or audio output. The ultimate step is converting text to sign. When a user enters an English text, it will be converted to sign language and display the relevant animated images of the sign. They have named this translator “Easy Talk,” and its sole purpose is to build efficient and effective communication between the differently abled community and normal people.

In the research [16] of Speech To Sign Language Translator For the Hearing Impaired by Ezhumalai P, Raj Kumar M, Rahul A S, Vimalanathan V, Yuvaraj A in Department of CSE, RMD. Engineering College. Using Natural Language Processing algorithms, they have implemented a system that converts speech or text to Sign language. The input audio is turned into text using speech recognition, which is then translated into sign language. Word segmentation and root word extraction are done using NLP techniques. This project will be translated the speech into Indian Sign Language to make it more accessible to Indians. They use python, OpenCV for image and video processing, and Pyaudio for speech recognition. And Tkinter for the GUI. They used similar technologies as our system, but the interfaces are not interactive as for modern technology.

III. METHODOLOGY

Since the proposed system includes several features to facilitate communication, five main components will ensure the system’s functionality and completion. They are;

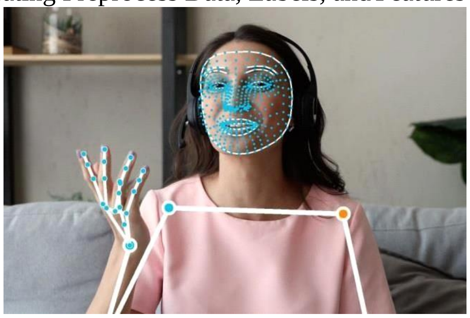
- Detecting and identifying Sign language.
- Predicting grammatically correct sentences using the words relevant to the given sign.
- Translating text in images to sign language.
- Translating voice to sign language.
- Virtual Sign Language Keyboard.

A. Detecting and identifying Sign language

The camera should first detect and identify Sign language for this function to work. The designed system will be programmed in this component to follow the tasks mentioned below.

- Key points Configuration

- Key point values Extraction
- Key point values Training
- Creating Preprocess Data, Labels, and Features



- Building and Training LSTM Neural Network

When configuring the motion detection key points and landmarks have to be implemented. For this, the best technology to use is Media Pipe Holistic. MP allows for capturing live perception of body poses, facial landmarks, and hand tracking. In this system, capturing the body poses an unnecessary. Hand tracking and facial landmark recognition are enough. So as mentioned before, configuring key points using MP holistic comes first. An example of a configured image with a Media Pipe can be seen in Fig.1. and Fig.2.

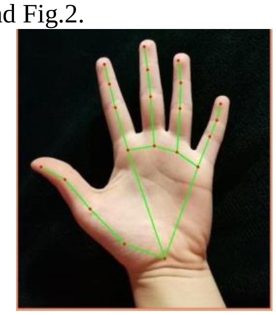


Figure 2: Hand landmarks

Once the MP Holistic configuration is done, the next step is to extract and train those values. After these steps, Pre-processed data or the hand sign word data set need to be created. These Datasets will be created manually. As the final step, LSTM Neural Network must be created and trained.

Once the above tasks have been completed, the system will be able to detect motions by using key points and will be able to identify them by referring to the data sets that have already been uploaded to the database. Since the data needs to be accessed from anywhere, a cloud database is best suited for this system.

B. Predict grammatically correct sentences using the words relevant to the given sign

The first step is to put the sign language word and the English sentence together. The words will be put together as one statement until a delay of more than five seconds. The sentence will then be passed through a transformer to make it grammatically correct. The task mentioned above is carried out using the T5 transformer model. T5 can be trained to convert text across different formats because it is a text-to-text model. After tweaking, the model also provides a decent level of accuracy.

Separate data sets have been made to address the differences between the two languages. They are,

- When asking questions, the question word ("why", "what", "how", "when", etc.) is signed at the end in ASL, whereas in English it is said at the front.
- ASL has signs only for the root words in English.
- Prepositions do not have signs in ASL
- The word order of an ASL sentence comes as an object, subject, verb whereas in English it is subject, verb, object

Each dataset contains more than 1000 data with a grammatically incorrect sentence as input and a corrected sentence as output. The T5 tokenizer tokenized both inputs and targets. The dataset was divided into training and test data. 80% of the dataset is taken for training and 20% for testing and evaluating the model. The data set was trained stepwise, gradually increasing the data and changing the training arguments to achieve better accuracy. The Seq2Seq trainer class is used to instantiate the model. The model is trained on data with a learning rate of 0.01 and a batch size of 64. The model is trained on data for 30 epochs and at the end, it retrieves the best one with less loss. Weights and biases are used to monitor model training performance in real-time.

With the increase of data and the fine-tuning of the model, there was a clear decrease in the loss and an increase in accuracy. The current loss is 0.16 which means there is 84% accuracy. The corrected sentence is displayed in text format in real-time. Users can select the narration option and make the system narrate the grammatical sentence as it will be giving a more natural experience of communicating. When the user clicks on the narration button the grammatically corrected sentence will be passed to the narrate method. The Speech Synthesis web speech API is used to implement the narration feature.

C. Translating text in images to sign language

This component can directly convert an image containing English text/numbers into American Sign Language symbols (gestures performed by a hand). This function consists of two main processes; Detect and Recognize texts/numbers for a given input image and translates the extracted texts from the image into ASL symbols.

a) Text detection and recognition from a given input image

This process is mainly related to Optical Character Recognition (OCR). In this process, deep learning OCR algorithms were used.

e-ISSN: 2250-0758 | p-ISSN: 2394-6962
Volume-12, Issue-6, (December 2022)
www.ijemr.net
https://doi.org/10.31033/ijemr.12.6.16

i. Data processing

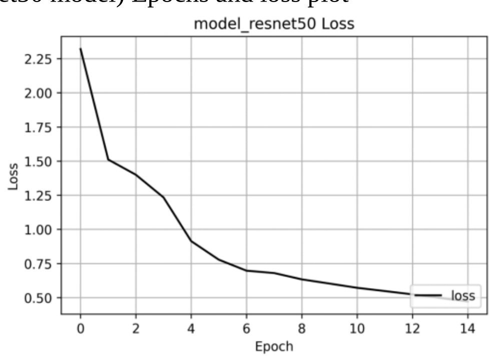
Although many data sets are freely available for text detection and recognition, due to unstructured text, different orientations, and other factors, not all datasets will work well with all deep learning models. ICDAR 2015 Dataset is used for this task because it is readily available and has enough images for non-commercial use. These images contain English text. The dataset includes 1500 images, 1000 for training, 500 for testing, and ground truth data for each set. Nearly 2,077 instances of cropped text and more than 200 samples of irregular text are also included in the dataset.

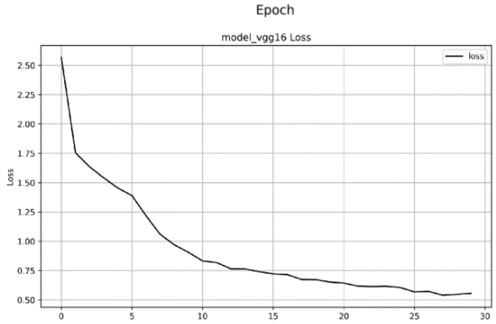
Pre-processing techniques were applied to the data set to improve accuracy and reduce the complexity of the dataset. The images in the dataset were resized into the same dimensions (720 x 180) and converted from BGR to RGB because they were in different sizes and color formats.

ii. Training text detection models

In this step, the ResNet50 model and the VGG16 model on ImageNet data were used as the text detection models. Both models (VGG16 model and ResNet50 model) were trained for 30 epochs with Adam optimizer with a batch size of 10 and a learning rate of 0.001.

Fig.3 below shows the Model-1 (VGG16 model) Epochs vs loss plot, and Fig.4 shows the Model-2 (ResNet50 model) Epochs and loss plot





iii. Model analysis and Model quantization

As shown in Fig.3 and Fig.4, the total loss for Model-2 (ResNet50 model) is lower than Model-1 (VGG16 model), which means that Model-2 has higher accuracy for text detection. Since the ResNet50 model performs better than the VGG16 model, the ResNet50 model was chosen for text detection.

The process of model quantization aids in lowering the computational and memory costs of using neural networks. After performing model quantization, the ResNet50 model size was reduced from 105.43 MB to 48.34 MB.

iv. Final implementation of text detection and recognition

The Tesseract and east_resnet(ResNet50) models have been combined to create the final text detection and recognition pipeline. The model is now prepared to detect and identify text when an image is uploaded. The model is saved as a pickle file using the Python pickle module.

Finally, the float16 quantized model was selected after model quantization. Furthermore, we need to combine this model (the text extraction model) with the next stage, which is to convert the extracted text into ASL symbols.

b) Translation of extracted text into ASL symbols

Now, the text will be detected and identified when a user uploads an image. The next step is to implement a method for converting those extracted texts into American Sign Language symbols.

i. ASL word dataset creation and ASL font creation

ASL has signs for words, phrases, alphabets, and numbers. In this step, the ASL sign data set containing the ASL images for the relevant word was created.

There are no signs for every word in American Sign Language. For example, let's take a person's name, "John". In that case, there is no sign for "John" in ASL. Therefore, we require a method of showing that name in ASL. In order to accommodate this, an ASL font for English alphabetic characters was created. The ASL font was created using the images of hand signs in TrueType Font file (ttf) format. The created ASL font contains signs for 0 - 9, a-z / A-Z. Additionally, it contains punctuation marks to enhance the readability or the clarity of the translation of text images to ASL.

ii. Final implementation of extracted text into ASL conversion

In this step, the extracted text from the input image will be obtained using the previously trained and quantized float16 model. The text is then stored in an array list in alphabetical or word format. After that, the pattern recognition/matching parameters obtained from the extracted text are compared with the ASL signs in the database. The corresponding result will be shown after the correct values have been matched. The system returns images that describe the motion if there is a corresponding sign for the extracted text in the database. At the same time, other words that do not have a corresponding sign in

e-ISSN: 2250-0758 | p-ISSN: 2394-6962
Volume-12, Issue-6, (December 2022)
https://doi.org/10.31033/ijemr.12.6.16

the database are directly translated using the ASL alphabet (ASL font).

D. Translating voice to sign language

In this component the first part is about voice-to-sign translation. From the microphone voice input is received by the system. This component used a web kit speech recognition, a JavaScript web API that recognizes speech in web pages. The speech recognition features in Chrome may be finely controlled and flexible thanks to this API, which enables the program to process human speech into text format.

The next step is pre-processing the words that are in the text format. English contains its word format, grammar, and many more, while ASL only includes the core words with signs. Because of that, it needs pre-processing when converting text to ASL. For this part, the NLTK library was used. The first step of pre-processing is tokenizing the sentence. In this process, sentences will split into words. After that, stop words are removed, and the next step is to lemmatize to get the word's root form. After lemmatization, words are classified into tenses.

Now the last step is to show the signs by searching the database. it will show a sequence of images extracted from the database. If the word is not in the DB the alphabetical signs are displayed on the screen.

E. Virtual Sign Language Keyboard

The system also has another feature for sign language users to type in what they are trying to say via a virtual keyboard. The keyboard is the most used input item when communicating on the internet. Nowadays, the most keyboard is customized according to the various languages to ease of use.

In this, we're implementing a sign language keyboard where the normal keys are replaced with the ASL alphabetical signs with the backspace, clear, and space buttons. And the key pattern is normal as the physical keyboard.

First, through the camera, it will identify the hand and landmarks of the hand by using the CVzone package. After that, it will calculate the gap between the tip of the index and the thumb finger. If that gap is less than the assigned value, it will detect as a pressed key and the color will change to show the key is selected. The implemented sign keyboard is displayed while the key detection process is running. And the output is shown at the bottom of the keyboard. Clear and the backspace button is also implemented for error correction.

IV. RESULTS AND DISCUSSION

After following various methodologies to implement the components of the proposed system, a comparison was done with each methodology to determine the overall best ways to implement the components. After the comparisons, the methodologies which showed the best test results were chosen.

- ***Detecting and Identifying Sign Language***

Two models were tested to implement this component. The first model used TensorFlow Object detection picture in picture frame detection and the second model used the action detection Long Short-Term Memory model (LSTM).

Five signs from ASL were taken for testing purposes, each performed ten times to check the detection system's accuracy. After that, the two models were compared.

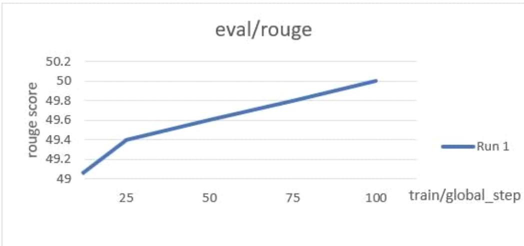
Table I: TF Object Detection Model Accuracy Vs LSTM Model Accuracy

Sign	TF Object Detection Model	LSTM Model
Sign 01	60%	90%
Sign 02	50%	100%
Sign 03	60%	80%
Sign 04	40%	80%
Sign 05	70%	100%

As seen above in Table I, the LSTM model with the action detection model is far more accurate than the TensorFlow Object detection picture in the picture frame detection model. So based on the test results, the LSTM model is better for Sign Language Detection and Identification.

- ***Predicting grammatically correct sentences using the words relevant to the given sign***

e-ISSN: 2250-0758 | p-ISSN: 2394-6962
Volume-12, Issue-6, (December 2022)
https://doi.org/10.31033/ijemr.12.6.16

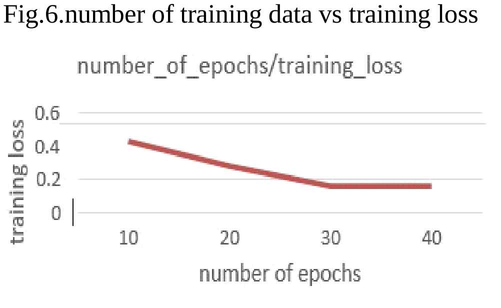


The ROUGE score was used as a metric for model evaluation. Below is an image taken from W&B showing the rouge score with one round of practice. The rouge score reached 50 after 1 epoch.

Figure 5: Accuracy with one round of training

According to the below graph, the increase in training data resulted in a decrease in training loss.

Fig.6.number of training data vs training loss



In conclusion, increasing the number of training data and epochs resulted in a lower training loss. Also changing the number of epochs after a certain number (number of epochs=30) has not made much difference. Therefore, the number of training data can be increased to increase the accuracy further.

C. Translating text in images to sign language

As mentioned in the methodology section, the East_resnet50 model and the East_vgg16 model were trained to detect text in an image. The loss function was considered to select the best model among them. Analyzing the results of east vgg16 and east resnet50 models, the total loss for resnet is approximately 0.15, which indicates that the resnet model is the better detection model.

Both models were trained using a dataset of 1000 image data. The dataset is trained over 30 epochs with a learning rate of 0.001 and a batch size of 10. Higher batch sizes typically do not lead to higher accuracy, and other aspects like learning rate and optimizer performance will also be crucial. Therefore, the model can be trained more effectively if the learning rate decreases with the batch size.Fig.8 and Fig. 9 show how the learning rate and batch size affect the model's accuracy.

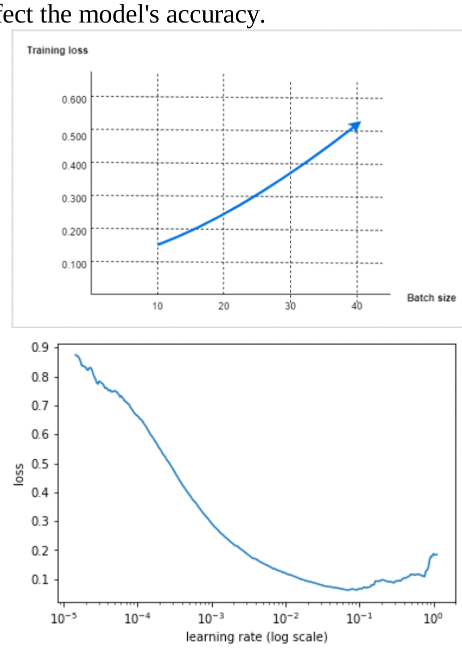


Figure 7: Number of epochs vs training loss

D. Translating voice to sign language

This function takes some time for the voice-to-text and text-to-sign translation based on the internet connection. The connection is categorized into fast (10 Mbps), average (7 Mbps), and slow (3 Mbps).

Table II: Time For The Voice-To-Text And Text-To-Sign Translation

Translation	***Fast***	***average*** **slow**	
Voice-to-text translation	3.37s	4.05s	7.00s
Text to sign translation	2.02s	3.45s	6.25s

Table II above shows that with the average speed, the final output can be shown in 8 seconds, the slowest speed shows the output in 13 seconds, and the fast speed shows the output in 5 seconds. Most of the time the average speed is used so the time is reasonable to show the output.

V. CONCLUSION AND FUTURE WORK

The system contains all the functions for successful two-way communication for sign language users and Natural Language (English) users. The system can translate sign language to English text and predict grammatically correct sentences using those texts, narrate the predicted sentences, translate voice to sign language, and finally Convert English texts in an image to sign language. The system is designed as a web application; therefore, it resolves many device compatibility issues, OS compatibility issues, etc.

However, compared to the ASL dictionary, the designed system contains fewer signs than the ASL dictionary. Hence giving more space to increase the built-in system sign language database, and with the user's suggestions and ideas about the system, some minor changes will happen in the planned future maintenance periods.

ACKNOWLEDGMENT

We are very thankful to the Sri Lanka Institute of Information Technology, Malabe, Sri Lanka, for providing the necessary facility and resources to conduct the research.

REFERENCES

[1] Vedak, O., Zavre, P., Todkar, A. & Patil, M. (2019). Sign language interpreter using image processing and machine learning. Available at: https://www.irjet.net/archives/V6/i4/IRJET-V6I4413.pdf.
[2] Apoorv, S., Kumar, S. & Prabha, S. (2020). Indian sign language interpreter using image processing and machine learning. Available at: https://iopscience.iop.org/article/10.1088/1757-899X/872/1/012026. [Accessed 24 January 2022].
[3] Jiangbin Zheng, Zheng Zhao, Min Chen, Jing Chen, Chong Wu, Yidong Chen, Xiaodong Shi & Yiqi Tong. (2020). An improved sign language translation model with explainable adaptations for processing long sign sentences. *Computational Intelligence and Neuroscience, 2020, Article ID 8816125*, 11. DOI: 10.1155/2020/8816125
[4] Z. A. Memon, M. U. Ahmed, S. T. Hussain, Z. A. Baig & U. Aziz. (2017). Real time translator for sign languages. *International Conference on Frontiers of Information Technology (FIT)*, pp. 144-148. DOI: 10.1109/FIT.2017.00033.
[5] T. Vichyaloetsiri & P. Wuttidittachotti. (2017). Web service framework to translate text into sign language. *International Conference on Computer, Information and Telecommunication Systems (CITS)*, pp. 180-184. DOI: 10.1109/CITS.2017.8035336.
[6] P. S.G., J. J., S. R., S. Y., P. G. S. Hiremath & N. T. Pendari. (2018). Dynamic tool for american sign language finger spelling interpreter. *International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, pp. 596-600. DOI: 10.1109/ICACCCN.2018.8748859.
[7] https://www.youtube.com/watch?v=IDghHm_w4k4. [Accessed 30 June 2022].
[8] Medium. (2022). NLP: Building a grammatical error correction model—deep learning analytics. Available at: https://towardsdatascience.com/nlp-building-a-grammatical-error-correction-model-deep-learning-analytics-c914c3a8331b.
[9] Linguisticsociety.org. 2022. What is Sign Language? | Linguistic Society of America. Available at: https://www.linguisticsociety.org/content/what-sign-language#:~:text=ASL%20has%20many%20ways%20 of,different%20from%20that%20of%20English.
[10] Programmersought.com. (2017). An efficient and accurate scene text detector - programmer sought. Available at: https://www.programmersought.com/article/56252830 437
[11] Zhou, X., Yao, C., Wen, H., Wang, Y., Zhou, S., He,W. & Liang, J. (2017). EAST: An efficient and accurate scene text detector. Available at: https://doi.org/10.48550/arXiv.1704.03155.
[12] takelessons. (2019) ASL sentence structure: Word order in american sign language L ASL lessons, YouTube. Available at: https://www.youtube.com/watch?v=nzrbvMeoBnE. (Accessed: November 7, 2022).
[13] P, E., M, R., A S, R., V, V. & A, Y. (2021). Speech to sign language translator for hearing impaired. Available at: https://www.turcomat.org/index.php/turkbilmat/article/download/4679/3906.
[14] A.Memon, Z., M, U., S, T., Abbas Baig, Z. and Aziz, U. (2017). Real time translator for sign languages. Available at: https://www.researchgate.net/publication/322583431_Real_Time_Translator_for_Sign_Languages.
[15] D. Manoj Kumar, K. Bavanraj, S. Thavananthan, G. M. A. S. Bastiansz, S. M. B. Harshanath & J. Alosious. (2020). EasyTalk: A translator for sri lankan sign language using machine learning and artificial intelligence. *2nd International Conference on Advancements in Computing*

e-ISSN: 2250-0758 | p-ISSN: 2394-6962
Volume-12, Issue-6, (December 2022)
https://doi.org/10.31033/ijemr.12.6.16

(ICAC), pp. 506-511. DOI: 10.1109/ICAC51239.2020.9357154.

[16] Ezhumalai P, et. al. (no date) Speech to sign language translator for hearing impaired, *Turkish Journal of Computer and Mathematics Education*. Available at: https://www.turcomat.org/index.php/turkbilmat/article/view/4679. (Accessed: January 22, 2022).

[17] Mahesh Kumar. (2022). *NB recognize gestures using Linear Discrimination Analysis (LDA). In that, it mentioned to extract features from the dataset they used MATLAB*. [Accessed 23 January 2022]

[18] van der Straaten, T., Briaire, J., Dirks, E., Soede, W., Rieffe, C. & Frijns, J. (2021). The school career of children with hearing loss in different primary educational settings—a large longitudinal nationwide study. Available at: https://academic.oup.com/jdsde/article/26/3/405/6232578 [Accessed 31 January 2022].

[19] M. Taskiran, M. Killioglu & N. Kahraman. (2018). A real-time system for recognition of american sign language by using deep learning. *41st International Conference on Telecommunications and Signal Processing (TSP)*, pp. 1-5. DOI: 10.1109/TSP.2018.8441304.

[20] Martel, J. (2022). Hearing and speech impairment resources. Available at: https://www.healthline.com/health/hearing-or-speech-impairment-resources#hearing-impairment. [Accessed 25 January 2022].

[21] Ardiansyah, A., Hitoyoshi, B., Halim, M., Hanafiah, N. & Wibisurya, A. (2022). Systematic literature review: american sign language translator. Available at: https://www.sciencedirect.com/science/article/pii/S1877050921000442. [Accessed 11 February 2022].

[22] V. N. T. Truong, C. Yang & Q. Tran. (2016). A translator for American sign language to text and speech. *IEEE 5th Global Conference on Consumer Electronics*, pp. 1-2. DOI: 10.1109/GCCE.2016.7800427.

[23] The different types of sign language in the UK: Calvert Exmoor. (2022) Calvert Trust Exmoor. Available at: https://calvertexmoor.org.uk/news/different-types-sign-language-uk/ (Accessed: November 7, 2022).

[24] American Sign Language. (no date) National Institute of Deafness and Other Communication Disorders. U.S. Department of Health and Human Services. Available at: https://www.nidcd.nih.gov/health/american-sign-language#:~:text=ASL%20is%20a%20language%20completely,word%20formation%2C%20and%20word%20order. (Accessed: November 7, 2022).