AN AUTOMATED APPROACH FOR TRANSLATING ENGLISH TEXT TO PAKISTAN SIGN LANGUAGE

Sharoon Nasim[1], Sidra Minhas[2], and Nosheen Sabahat[3]

[1,2,3] Department of Computer Science, Forman Christian College, (A Chartered University) Lahore, Pakistan
Corresponding author: First Author (e-mail: sharoonnasim@fccollege.edu.edu.pk).

ABSTRACT: In Pakistan, the population of hearing-disabled children is on the rise. Unfortunately, the literacy rate of this population is far less than their counterpart peers due to a lack of knowledge and scarcity of resources for Pakistan Sign Language (PSL). However, the availability of an electronic repository of PSL (www.psl.org.pk) has opened new horizons for the well-being of the hearing-impaired population of Pakistan. In this work, we aim to propose a standardized system that focuses on converting written English sentences into PSL to enhance the readability and understanding of an international language. Currently, the system is focused on providing an interface for preschool children to enhance their confidence in understanding the English language. Our system has two input options: typed English text or images of English text. The English Input Text is Translated into PSL lexical after performing NLP and machine translation techniques such as tokenization, lemmatization, etc., this process also deals with out of vocabulary words, and the resulting system is developed in three phases that are provided in the methodology section. For translating alphabets and words, our system achieved an accuracy of 89.58% and 100% respectively in a 10-fold cross-validation setup. For the evaluation of English sentence translation, we selected 5 storybooks with 104 sentences from the Oxford Reading Tree for ages 3-4 years and converted them to PSL. Later, with the help of a domain expert, we obtained an accuracy of 81.25% and a Level of Understanding of 75.96% which was recognized as comparable and rated as user-friendly by the domain experts. The availability of an automated application to convert the international English language to PSL will eradicate the need for a third person for translation while enhancing the social inclusion of deprived children.

Index Terms-- Automated translation, Natural Language Processing (NLP), Pakistan Sign Language (PSL), Text recognition

I. INTRODUCTION

Communication is the basic need, and language is the medium through which anyone can communicate with their environment. People with hearing and speaking abilities can easily do so by learning the language of communication. Similarly, the deaf community uses Sign Language for this purpose, but the communication barrier occurs when people from different language backgrounds come across, like a Japanese and Urdu speaking person cannot communicate without a translator or until they know any mutual language like English. Google translate can solve this problem for the normal community by providing translation, but the problem is not solved for the deaf community, it has been observed that a hearing-impaired person who doesn't know English or Local speaking/writing language cannot communicate, their knowledge of the written language is very little that cannot become a bridge between such communication. Therefore, mostly they need a person who can help them as a translator.

Further, Pakistan Sign Language has two variations in hand gestures, It is two-handed or single-handed, PSL is two-handed and ASL (American Sign Language) uses single-handed signs, which causes communication challenges when two Sign languages have different signs for any word. For this reason, it is very important to document sign language. In Pakistan, this process has been carried out since 1987. Different institutes publish resources about Pakistan Sign Language that include Basics of Daily Conversation, Relationship terms, Time and Season, Body Anatomy, Numbers, and different Sign Language Stories. [8] All these efforts are made to help the deaf community to learn sign language. But now in the time of digitalization, it is not enough to just transform those handbooks into digital applications and web platforms, though it is more accessible and cheaper as compared with the publication expenses. The main objective of this research is to help children to learn and understand the English language in their Early life.

According to a report by the World Health Organization [1], there are around 430 million individuals who require rehabilitation for their 'impairing' hearing loss (432 million adults and 34 million children). It is estimated based on findings and observations that this number will increase up to 700 million or one in ten individuals, who will have a hearing problem by 2050 [1]. These numbers include the individuals that lose their hearing ability due to different factors along with those who are deaf by birth. Disability of any kind causes hindrances in normal life, however, being hearing-impaired results in a huge communication gap between the impaired and society which further causes a lack of productivity and mental health issues. The prominent ways of communication used by the Deaf and hard-of-hearing (DHH) people are lip-reading, sign language based on hand symbols and gestures, and with the help of an interpreter. Constant efforts are being made to improve this condition by incorporating lip-reading and sign language for better communication and to reduce the dependency upon a third person i.e. the interpreter.

According to the Family Education Services Foundation (FESF) [2], 70% of the DHH people prefer sign language over lip-reading. There are reportedly 143 sign languages around the world. Henceforth, a plethora of ICT interventions have been designed and implemented for the conversion of speech to sign language [3] and text to sign language [4]. Many solutions have been successfully implemented and are being used for the conversion of text to American Sign Language [3], and text to British Sign Language [3]. However, most of these solutions are inefficient and depend upon spelling words and converting each alphabet to the respective sign. Others that aim at sentence recognition and translation have a very low level of Understanding (LOU) [5]. However, the situation in developing countries like India, Pakistan, and Nepal is worse.

As reported in 2018, approximately 10 million citizens of Pakistan are hearing-impaired. Only 10% of these use PSL [6], whereas 90% of these have limited or no access to education. Approximately 1.2 out of every 1,000 children who are screened are impacted by moderate to profound congenital, bilateral hearing loss [7]. The Digital Pakistan Policy (2018) and ICT Rights of Persons with Disabilities Act (2020) pledge to provide equal rights of education and opportunities to the deaf, however, the evidence of this is scarce. A total of 91 schools and training centers exist for all disabilities, specifically for the deaf. There are less than 40 that are collectively run by the Directorate General, Provincial Governments, and private sectors like the Family Education Services Foundation (FESF) and the Deaf Reach Program [2][8-10].

To reduce the literacy gap created due to the non-availability of schools, online, automated platforms are required. In this research, we aim to develop a prototype to convert English text to PSL video lexicon. English text can be typed in or extracted from an image. While typed input will assist in live communication, image input will help in reading and understanding printed text like in storybooks, billboards, etc. For proof of concept, the current scope of alphabets, words, and sentences used is limited to ages 5 and below.

The rest of the paper is organized as follows. In Section II, previously done work in the domain of English to PSL translation is covered. In Section III, we describe the dataset extraction and representation whereas in Section IV, we explain the methodology followed for the development of this system. In Section V, we analyze the results and discussion while in Section VI, we present the conclusion along with remarks of a domain expert.

II. RELATED WORK

The hearing-impaired community of Pakistan is immensely stigmatized and hence given the least priority, however with staggering numbers of the deaf, efforts are being made for the well-being and inclusion of this portion of the community. Various studies conducted on PSL are summarized in Table 1. Haseeb, A. et al. [3] worked on converting English speech into PSL. They have used machine translation and some available tools (sharp NLP, .Net with C#, WordNet, Microsoft Active Media Library, Microsoft Access, and Windows forms for GUI) to create a prototype. their approach gets 78% accuracy on the specific set of sentences. But their work still requires improvement in some areas like logic engine, NLP technique, and dataset to get practical worth. Their research does not deal with the Out of Vocabulary (OOV) words. Furthermore, their work is limited to a specific area and only deals with available vocabulary. They have developed a prototype that is named PSLIM. [3] They have used the rule-based machine translation technique to transform the syntactic structure of English into PSL.

Another Research work was quite helpful [4] though this work was different it translated the Arabic Sign Language to Arabic text. They have carefully worked on the syntax and grammar of both languages. Their proposed framework applies the morphological, syntactic, and semantic analysis of an Arabic sentence to interpret it into a correct syntax and grammar for the sentence in Arabic Sign Language.

In 2018, Abass, A. and et al. developed a framework that serves the translation purpose, extensive work has been done by NLP, they have included English and Urdu subtitles along with the video. However, no work has been done on the Sentence Structure and Grammar of PSL [5].

Recently researchers from UMT Pakistan have developed a sentence corpus that contains 2000 sentences, sentences in these videos are manually translated by involving domain experts of PSL. The input sentence is morphologically, syntactically, and semantically analyzed. The output results in a sentence that satisfies the structure and grammar of PSL.

Their proposed approach uses the rule-based machine translation and their system resulted in a 0.78 (Bilingual Evaluation Understudy) BLEU score [12]. They have also provided a detailed analysis of the grammatical structure of English and PSL. However, most of these studies report their results on self-determined datasets and not standardized ones.

Farooq et. al (2021) [31] shed light on various approaches used to bridge the communication gap between deaf communities and the hearing world. The authors delve into methodologies like generating sign language animations from natural language text and recognizing gestures for real-time translation. However, the review also acknowledges limitations in these techniques, such as the vast number of sign languages with unique grammar and vocabulary, and the challenge of accurately capturing the nuances of facial expressions and body language that are crucial aspects of sign language communication. By identifying these limitations and highlighting ongoing research directions, the researchers have pave the way for future advancements in sign language translation, promoting greater inclusivity and accessibility for deaf individuals.

Kahlon and Singh (2021) [32] provide a comprehensive review of the field of machine translation from text to sign language. The authors meticulously examine the existing literature, highlighting the critical role of this technology in facilitating communication for the deaf and hard-of-hearing community. They underscore the importance of universal access to information, emphasizing how machine translation can bridge the gap between spoken/written language and sign language. The review covers various methodologies employed in the field, including rule-based, statistical, and neural network approaches. It also discusses the challenges faced, such as the lack of standardization in sign languages and the complexity of translating context and non-verbal cues. The authors further delve into the evaluation metrics used to assess the performance of these translation systems. This paper stands as a significant contribution to the field, offering valuable insights and directions for future research in machine translation for sign language. It underscores the need for more inclusive and accessible communication tools, paving the way for advancements that could profoundly impact the lives of millions around the world.

Attar et. al (2023) [33] present a systematic review of the state of automation in sign language. The authors provide an in-depth analysis of the current advancements and challenges in the field, emphasizing the importance of automation for enhancing accessibility for the deaf and hard-of-hearing community. The review covers a broad range of topics, including the various techniques used in sign language recognition, generation, and translation. It also explores the role of machine learning and artificial intelligence in advancing these technologies. The authors highlight the need for more research in low-resource languages, pointing out the significant gaps that exist in this area. They also discuss the evaluation metrics and datasets used in the field, providing a critical assessment of their adequacy and reliability. This comprehensive review serves as a valuable resource for researchers and practitioners in the field, offering a clear overview of the current state of the art and suggesting potential avenues for future research. It underscores the potential of automation in sign language to transform communication and information access for the deaf community.

The preliminary and most vital step for the betterment and inclusion of the deaf is to provide quality education to deaf children so they can be comparable to their peers. Consequently, the first step was to formulate a common and unanimously accepted PSL curriculum. In 2018, Technology-Based Deaf Education (TBDE) developed the first-ever Pakistan Sign Language (PSL) online dictionary with 7,000 words translated into their video lexicon being used nationwide. The dictionary is available in the form of a website, mobile application as well and a DVD for those with no access to the internet. The existence of such a dictionary opens doors to unlimited horizons for the betterment of the DHH people in terms of education and communication [2].

TABLE I. SUMMARY OF RELEVANT WORK

No.	Author	Purpose	Data	Performance Metrics
1	Haseeb, A. A., Illyas, A.	Speech and text Translation into PSL.	Self-Collected	Accuracy 61% (untrained), 78% (trained)
2	Ali Abbas, Summaira Sarfraz	English text and speech to PSL to facilitate learning.	FESF Video Clips	Not Provided
3	Luqman, H., Mahmoud, S.A.	Converting the recognized Arabic Sign Language sentence into an Arabic sentence	Arabic Sign Language Corpus 2018	80% Manually, 0.39 BLEU and 0.45 TER (The Editing Required)
4	Nabeel Sabir Khan, Adnan Abid, Kamran Abid	Written English sentences for avatar-based PSL video	Self-collected 2000 sentences	The bilingual evaluation understudy (BLEU) is: 0.78

In this paper, we propose a framework for translating the international English language to PSL using the standardized dataset of PSL. Our system will take typed English text or extract English text from an image, look it up in the PSL dictionary, and display the corresponding PSL gesture. The major contribution of our work is:

a. Enhancing the usability of the system by incorporating text image inputs and employing Optical Character Recognition (OCR).
b. System evaluation at alphabet, word, and sentence level
c. Incorporating Natural Language Processing (NLP) for parsing, tokenization, and Syntactic arrangement of English sentence components into PSL grammar.
d. Written English sentence conversion to PSL by combining several video lexicons and conducting an understanding evaluation through a domain expert.
e. Utilizes a standardized dataset of PSL lexicon developed by FESF and Deaf Reach.
f. Provide an interactive and accurate application for pre-school and early years students.

III. DATASET

In this work, the dataset from the Pakistan Sign Language website (www.psl.org.pk) developed by FESF is used. FESF is a dedicated group in Pakistan working tirelessly towards creating educational opportunities for the most neglected members of civil society. Under their numerous initiatives is the Sign Language Development project which aims at deaf education and employment. In 2020, a PSL dictionary containing 5000+ vocabulary words from different categories was made available online. The available dataset contains 5348 video clips in 68 categories as shown in Table 2. Each video clip represents the PSL lexicon corresponding to a word. For the current study, a category of "Extras" is added to the vocabulary for placing the important words that are missing in the repository [2]. All these video clips follow the same structure in that each sign is repeated twice once with English subtitles and pronunciation and once with Urdu subtitles and pronunciation. To manage the dataset logically all the files are placed in different folders as per category. Further to the processing of the dataset, an SQLite database is maintained to perform all the tasks smoothly. This dataset belongs to an NGO (Deaf Reach), therefore using this dataset for any commercial purpose without taking proper permissions from the dataset owner can be an ethical and Intellectual Property concern. It is publicly available, but downloading the complete dataset in bulk is not possible we need to download each video one by one. Further due to small size of dataset it cannot be use directly for the deep learning or machine learning tools.

TABLE II
SUMMARY OF PSL VOCABULARY AND CATEGORIES

Category	CLIPS	Category	Clips
Adjective	464	Death And Funeral	14
Adverb	69	Drink	23
Airport	46	Extras	7
Alphabet	26	Family And Marriage	161
Appliances	33	Famous People	24
Around The House	35	Farming And Agriculture	52
Art	121	Flowers, Plants, and Trees	95
Banking	34	Food	47
Bathroom	27	Food Dishes	48
Beach	22	Fruits	32
Beauty	85	Geography	113
Bedroom	23	Government	57
Birds	37	Grammar	35
Body Parts	72	Health And Medicine	246
Brand Names	108	Holidays And Celebration	64
Burger King	92	Hygiene	17
Calendar And Time	53	Insects, Spiders and Reptiles	37
Clothes And Accessories	89	Jewellery	40
Classroom	58	Kitchen	81
Cleaning Product	19	Law And Order	61
Carpentry	66	Living Room	15
Colors	18	Mammals	64
Computers	63	Marine Life	25
Countries And Continents	38	Mathematics	86
Media	44	Office	46
Military	64	Pakistan Places	83
Music And Dance	31	Prepositions	38
Nouns	536	Professions	189
Numbers	32	Pronoun	36
Sewing	49	Science	133
Space	21	Sentences	43
Spices	31	Transport	105
Sports And Games	139	Vegetables	28
Weather	81	Verb	477

IV. METHODOLOGY

The complete implementation pipeline is engineered after analyzing all the existing systems and technologies that can be supportive of achieving the desired goal. Proper planning has been done to conduct the implementation into three phases. Figure 1 shows the summary of the implementation procedure. In phase 1, we check our system's efficacy in identifying and translating the 26 English alphabets, upper and lower case both to PSL whereas in phase 2, we quantify the conversion of words to the respective lexicon. In phase 3, we employ NLP methods to convert an English sentence to the respective representation in PSL and evaluate the accuracy and level of understanding of the sentences. Models used in this research are Tesseract OCR (primary) and EasyOCR (secondary) [17-19] for text recognition, this combination is used for handling of various fonts and single characters. For Text processing in Logic Engine Natural Language Toolkit (NLTK) and wordnet [23-30] is used for POS tagging, Lemmatization, Stemming, tokenization, Synonym finding, Similarity checking of two words.

This section provides a detailed explanation of the proposed framework and its implementation. Input and the resulting output were planned in the beginning that a text is recognized from an image and after processing the text a video is generated that provides PSL interpretation of the recognized text. Techniques used in both input and output phases are improved throughout the implementation, but this paper states the finalized approach used in both phases. Further implementation of processing the recognized text is explained in three phases. Each phase is planned in a process to improve the final product, therefore, the complexity of each phase increases with the complexity of the task.

A) INPUT PHASE

a) IMAGE PROCESSING AND ROI SELECTION

Reading an image is the first component of the Input Phase, therefore OpenCV is applied to read an image, then ROI (Region of Interest) is selected from the image by creating a bounding box, the image is cropped according to the bounded box, then cropped image is converted to a grayscale image, it utilizes less memory. In the end, the image is converted into a binary image, this process optimizes the performance of OCR. OpenCV also takes a picture from the camera and then processes that image in the same way [12-14].

b) TEXT RECOGNITION

The processed image is input for the next component which is Text Recognition [15] [16] OCR is a technique that helps to read text from images. It was a challenging task to detect text from different types of images and text fonts, Tesseract OCR is adopted and maintained by Google since 2006 [17]. Briefly, in the initial step of Tesseract OCR, an input picture is processed through the "Adaptive Thresholding" procedure in this the image is changed into Binary Image. Later Binary Image is processed through "Connected Component Analysis" in which text/word parts are outlined. Then these outlined characters are passed to Pass-1. In Pass-1 detected text/words are additionally processed into an adaptive classifier which thinks about the data as trained data.

Now the recognized text is again processed through Pass-2 but this will utilize the Adaptive Classifier for the text recognition.

The explanation for passing through Pass-2 is to know the context of text from Pass-1 [16-19]. Tesseract OCR is very efficient, but it has some shortcomings, it is unable to read text from images that contain a single alphabet or text in an uncommon font size. Therefore, another secondary OCR is employed to deal with these shortcomings. After comparisons between famous and powerful OCR engines, EasyOCR [13][15][17][20] is employed as a secondary OCR.

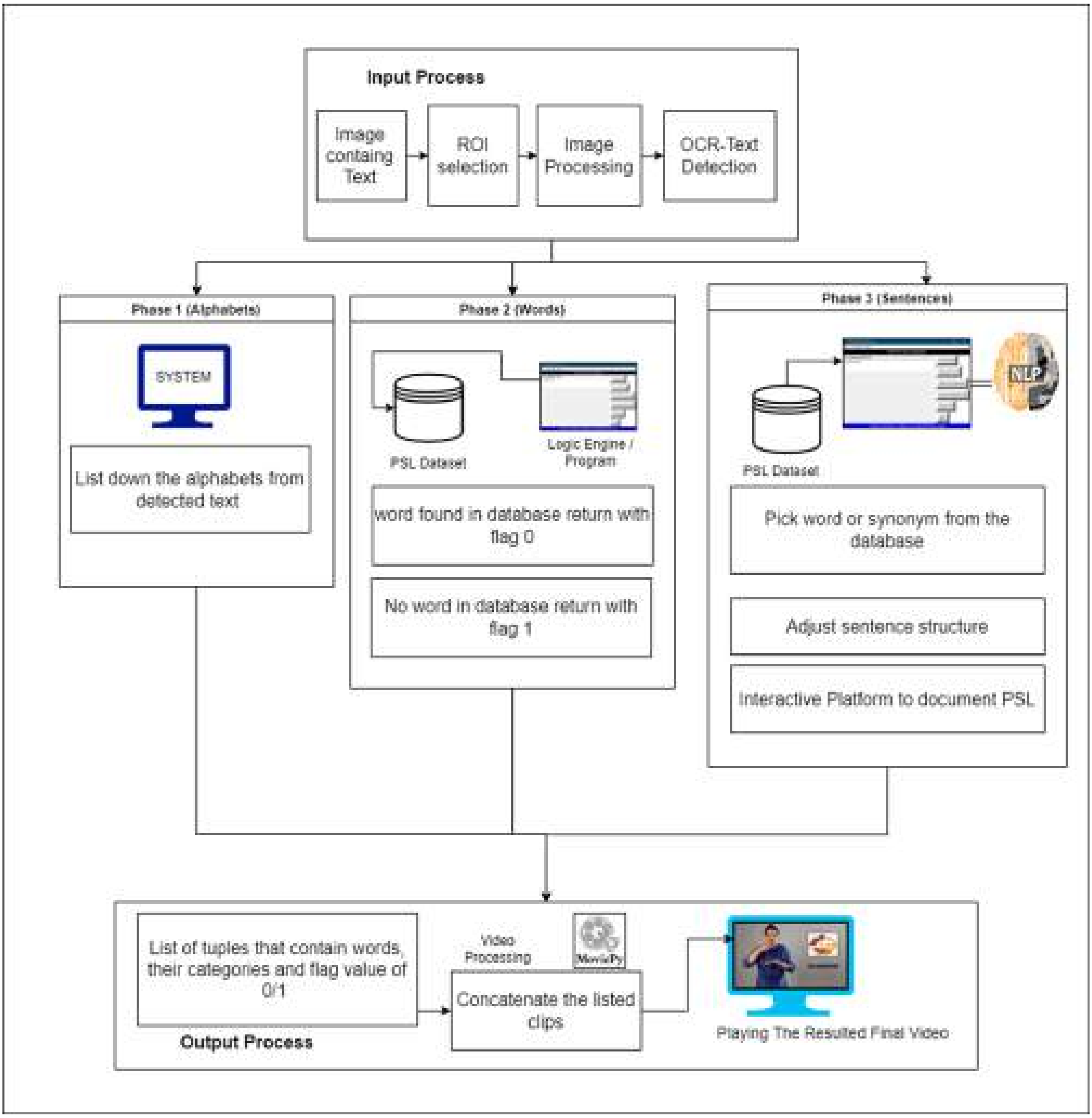

FIGURE 1. System Overview; describing all the phases of development

Text recognition is divided into two stages in which Tesseract OCR performs as the primary OCR and EasyOCR as the Secondary OCR, if Tesseract doesn't recognize anything then the processed image is given to the EasyOCR. The provided framework for input is finalized in phase three of processing whereas it works for all three phases. Before processing, recognized text is shown to the user in GUI, the user can verify whether a correct text is recognized or not. During implementation, different techniques were employed and compared, but only the efficient and best resources are mentioned.

B) OUTPUT PHASE

This phase takes a list containing tuples of a filename, category name, and flag value, if the flag value is 0 it means that the word exists in the database and has a video clip in the repository but if it is 1 that means the word is not found. This module deals with both situations. If a word exists then it resizes the video and splits the video in half as each video contains each sign twice, and it depends on the input to concatenate the video clips in sequence or play the single video clip. Unlike when the flag value is 1 it means that the word is not found in that case, therefore the word is spelled with the PSL alphabets, and all alphabets of a word are concatenated in a sequence and returned as a single video. Later it is played or concatenated with other word clips. After creating the video, it adds subtitles to let the user know about the complete word/sentence, after performing all these steps it plays the video [21][22].

C) PROCESSING PHASE

The final processing phase is improved in three phases as per the complexity and availability of the dataset.

a) PHASE 1:

In Phase 1, the objective is to read text from an image, and as a result, the system returns a video of PSL alphabet signs.

INPUT	An image that contains the English Alphabet/ Word
OUTPUT	Video Clip of PSL Alphabet or Sequence of Alphabets for word
DATASET	26 Video Clips of Alphabet

i. Logic Engine:

After reading the text with OCR, the text is converted into lowercase as there is one sign for both upper and lower-case alphabets. If there is only one alphabet, then the filename from the file is given to the Video Processing Module but if there is more than one alphabet then the filename of each alphabet is given in the form of a list.

b) PHASE 2:

In Phase 2, the objective is to read text from an image, and as a result, the system returns a video of PSL signs video.

INPUT	Image that contains English Alphabet/ Word
OUTPUT	Video Clip of PSL Sign or Sequence of PSL signs for word, If any word is not found in the vocabulary, then return a sequence of alphabets for that word
DATASET	Most Important Categories (Nouns, Verb, Adjectives, School, Pronoun, etc.)

i. Database Processing:

As PSL signs are increased with their categories, so in this phase, the sqlite3 database is established which keeps the filename, its directory (category), and the associated word. And before saving it into the database every text is transformed into lowercase. This database is designed in such a way that it can be updated with the newly added vocabulary video files.

ii. Logic Engine:

After reading the text with the help of OCR, the text is converted into lowercase, then carriage return "\n" is removed further punctuations are removed except (?, .) After processing text is tokenized, then these tokens are searched from the database by using SQL queries. It forwards the filename and category name to the next component, if a word is not found in the database, then it returns a flag value that is used by the video processing module.

c) PHASE 3:

In Phase 3, the objective is to read the English sentence text from an image and to return a PSL video clip sequence that makes sense and is understandable to the user.

INPUT	An image that contains English sentence text.
OUTPUT	Video Clip of PSL Signs that provide understanding to the user about written text
DATASET	5348 Video Clips from 68 Categories

word	category	syn	status	sentence
patch	Health and Medicine	bandage	None	he put a patch on it
bricks	Carpentry	brick	A	he put it on the bricks
pour	Geography	stream	None	the water poured out
roses	Flowers, Plants and tree	rose	A	she wanted to water the ro
said	Verb	say	A	said mum
pool	Geography	pond	None	the children got the paddli
biff	Verb	lick	None	biff turned on the tap
and	Prepositions	with	A	NULL
pop	Verb	start	None	up pops the big carrot
loads	Nouns	burden	None	loads of toads
toads	Insects_Spider_and_Reptiles	toad	A	loads of toads
chip	Verb	check	None	dad got biff and chip
load	Verb	stretch	None	a coach load of men ran up
men	Family and marige	man	A	a coach load of men ran up

FIGURE 2. Database keeping a record of synonyms and Out of Vocabulary words

i. Database:

In this phase, another table is added to the database to keep a record of synonyms that are used in place of words that are out of vocabulary (OOV). This table is useful to increase the responsiveness of the system. Five fields are kept in this table word (filename), category of synonym (directory name), synonym, status (this records the approval of PSL developers), and sentence (this records the sentence for which the word is not found). Later, the PSL Expert can review this table to approve or disapprove the suggested synonyms.

```
If  Aux "have" is in the start of sentence

>>> [sub,verb,obj,Prep,Neg,qes,wh,time]

If there is preposition in the beginning of the sentence

>>> [Prep,sub,Neg,verb,obj,qes,wh,time]

otherwise this structure for all other cases

>>> [sub,time,obj,Prep,Neg,verb,qes,wh]
```

ii. Logic Engine:

This part improves the logic engine of phase 2, before converting text into tokens, contractions are replaced with the appropriate words “don’t” and change into “do not”. Later text is chunked with the help of context-free grammar [23-26]. Chunking [24][27] returns a tree that contains the POS tag for now leaves of the trees are listed. These tags and their explanation is provided in Table III. This list is further processed to arrange the tokens according to the grammatical structure of PSL.

TABLE III
RESULTING SENTENCES FOR THE VIDEO PROCESSING MODULE

English.	Resulted Structure for PSL	Urdu
You are going to school	You now school toward go	تم ابھی سکول کی طرف جا رہے ہو۔
Tom and Ifra tug the big carrot	Name tom and name ifra big carrot tug	ٹوم اور افرا نے بڑی گاجر کھینچی۔

Resulting POS TAGS:
[(’you’, ’PRP’, 0), (’now’, ’JJ’), (’school’, ’NN’, 4), (’to’, ’TO’, 3), (’go’, ’VBG’, 2)]
[(’name’, ’JJ’, -1), (’tom’, ’NN’, 0), (’and’, ’CC’, 1), (’name’, ’JJ’, -1), (’ifra’, ’NN’, 2), (’big’, ’JJ’, 5), (’carrot’, ’NN’, 6), (’tug’, ’VBP’, 3)]

Words changed: “To” with “toward” as it is an OOV word. Similarly, “and” with “with” and “tug” with “pull”

Arranging grammatical structure is an important task. This has followed a rule-based approach that uses flags for phrase detection. This part iterates over the list of tokens, it separately identifies the Subject, Object, Verb, Negation, Preposition, Question Mark, and Time. Along with this lemmatization is also performed on verbs because our dataset contains the first form of the verb. A dictionary of helping or auxiliary verbs is generated to identify the tense of the sentence. For now, only three tenses: present, past, and future are processed and saved as Time. The lists of Subjects, Objects, and Verbs are searched from the dataset. If they do not exist in the database a proceeding token is added “name” for subject/ object, and “work” for a verb. This will help the user to understand that the next word is name or work. A rule is defined to arrange the syntax of the output list.

iii. Video Fetching Logic:

This part also searches for matching words like “sign language” which is a phrase that contains two words but there is only one single video file with this complete word “sign language”. Instead of joining two video clips, this part searches for the matching cases from the database to bring the appropriate video clip from the database for this Like query is used and some words that are found are handled to avoid those words from the tokens list. Secondly, in this phase to deal with out-of-vocabulary words, a Natural Language toolkit has been utilized, which uses the famous WordNet developed by Princeton University [28][29].

FIGURE 3. Rules that are used to adjust the sentence structure.

To deal with the words that are OOV, those words are processed to replace them with their synonyms that are available in the database. To find the synonyms, WordNet from NLTK [30] is used to list the synonyms. Each word from the list of synonyms is searched from the database and a new list is created that contains synonyms that are found in the database. Now the proceeding step is to find the most relevant synonym. For this purpose, Wu-Palmer Similarity [30] is utilized to return a score that represents how similar two-word senses are. This score is based on the depth of the two senses. semantic similarity has been applied to the synonyms available in the dataset, in this part, both words are converted into class, after that, Wu-Palmer similarity can be measured and the available synonym that has maximum similarity will be selected, to avoid irrelevant synonyms a measure of 0.25 score has been set to assure that word is relevant.

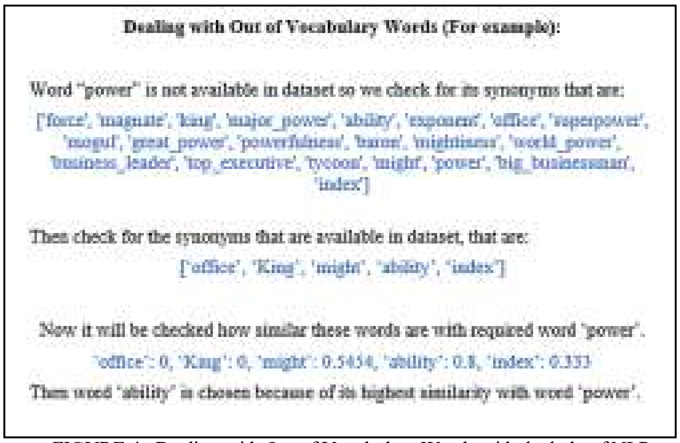

FIGURE 4. Dealing with Out of Vocabulary Words with the help of NLP

Word and its chosen synonym are saved in the database and if it is approved by the developing team, in the future it will directly replace the OOV words with their synonyms. This phase outputs the file path along with the flag to the Video Processing Module. Flag value tells the video process module whether the word exists in the repository or not, during this phase database is designed to provide an interactive approach that can handle the problem of documenting the PSL, it is designed in a way that it can serve as a universal platform for PSL dictionary, further, this approach will help the PSL instructors or Sign language interpreters.

V. RESULT AND ANALYSIS

Testing has been done in three phases; Testing has provided confidence that assures that this product met the claimed functionality. Detailed results are provided for each phase.

A) ACCURACY TESTING FOR ALPHABETS (PHASE 1)

As the system is designed to read the alphabet and return the alphabet Signs from PSL vocabulary, it is simple for the program to return the alphabet PSL signs but the vital part in this phase is the Text detection of how the system can read the alphabet from the images. For testing the accuracy of the system for this purpose “English Typed Alphabets and Numbers Dataset” [14] is used. 10 common fonts are tested for 52 alphabets including upper and lower case, randomly a font is selected and passed through the Image Processing and Text detection component, this process is repeated in 10-folds each time random fonts are tested. The average accuracy for this whole test is 89.585.

B) ACCURACY TESTING FOR WORDS (PHASE 2)

In this phase word reading and returning relevant signs from PSL vocabulary is the task, though returning words is a simple task and well tested, while the word reading from the image document is the main task. The accuracy of this task has been tested by using the "East Ridge PTA Spelling Bee" word list. This list contains 180 words, all the words have been read, and primary OCR is well capable of detecting text from the documents. The text detection testing phase has achieved 100% accuracy in reading the text. This accuracy can vary if a word has been inappropriately bounded by the user.

TABLE IV
SUMMARY OF ACCURACY (PHASE 1)

Iterations	Accuracy (x)	Mean	89.5852187
1	84.3137	Known Variance	18.207289
2	92.3076	Standard Deviation	4.267
3	84.6153	Observations	10
4	92.3077	Hypothesized Mean	90
5	92.3077	Significance level	0.5
6	92.3077	z	-0.30739396
7	92.3077	P(Z<=z) one-tail	0.379271762
8	92.3077	z Critical one-tail	0
9	80.7692	P(Z<=z) two-tail	0.758543525
10	92.3077	z Critical two-tail	0.67448975

C) ACCURACY TESTING FOR SENTENCES (PHASE 3)

In this phase, the task of sentence reading and returning a correct and sensible sequence of PSL video clips is a complex process because it requires domain expertness and correct sentence structure that the user can understand. For this testing phase, I have used books from the Oxford Reading Tree series. For testing purposes, a domain expert from DWAF [8] evaluated the videos and provided a level of understanding score for each sentence and categorical label. Up till now, 104 sentences have been evaluated and a score of understanding is 395 out of 520 which results in a 75.96% score. The categorical result provides 69 "yes" and 31 "partial".

The formula for calculating the categorical accuracy score:

$$\frac{\left(No.of\ Yes + \frac{No.of\ Partials}{2}\right)}{Total\ No.of\ Sentences} * 100$$

TABLE V
SUMMARY OF TESTING IN PHASE 3

Total Sentences	104
Points achieved	395

No. of Yes (weight 1)	69	69
No. of Partials (weight 0.5)	31	15.5

Total Sentences	104
Total Points for 104 sentences	520
Level of Understanding in Videos	**75.96**

No. of Yes (weight 1)	69	69
Sum after applying weightage		84.5
Categorical Accuracy Score		**81.25**

VII. CONCLUSION

This research is focused on investigating the ways that will help the DHH people to understand the written English language. It has been observed that even after learning complete sign language still, deaf people can only communicate with those who can understand sign language, they find difficulty in understanding written language. That is why they are unable to do written communication because the normal community that does not know sign language prefers to have written communication. This research has resulted in a platform that can enable the DHH people to learn and understand English, further, they would be able to do written communication. Lastly, this is an interactive platform that can communicate with the concerned departments to let them know about the missing words, so they can add those words to the vocabulary. Detailed work has been done on grammar and sentence structure to provide sensible output results. This platform has included Text recognition that will provide extensive usability to read text from images. Testing and evaluation of the program have proved that the resulting videos provided much sense and understanding to users. This research has resulted in incomparable results that have provided more functionality and scope. In addition, it is the first approach that uses the officially developed PSL dataset.

DISCUSSIONS & FUTURE WORK

Research work always has room for improvement, this research has room for improving the NLP engine, and work can be done to improve the sentence structure of the resulting video clips. Furthermore, while finding the synonyms more semantic analysis can be done to use the words that are suitable for the context of the sentence, for example, the "play" word has a different meaning in different situations like "playing the game" or "playing music". Further avatar can be used so one avatar is doing all the signs to provide a more interactive experience to the user. More optimizations and additions can be done to the database so, the program will become more efficient and responsive as the system is used with time. The database of 2000 sentences mentioned in [11] can be utilized for future work. After achieving more accuracy than 90%, a mobile application can be made to provide a handy tool for the DHH people.

ACKNOWLEDGMENTS

I extend my deepest gratitude to my advisor Dr. Sidra Minhas for her invaluable guidance, unwavering support, and profound expertise throughout this research endeavor. Her mentorship and insights have been instrumental in shaping the direction and outcomes of this work. I am profoundly grateful to Dr. Nosheen

Sabahat for her dedication, encouragement, and motivation, which have undoubtedly enriched the quality of this research work.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest to report regarding the present study.

FUNDING STATEMENT

The author(s) received no specific funding for this study.

REFERENCES

[1] World Health Organization WHO, "World Health Organization Statistics about deaf and hearing loss," [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearingloss.

[2] Family Educational Services Foundation, "Pakistan Sign Language – About Us," [Online]. Available: https://www.psl.org.pk/about.

[3] A. A. Haseeb and A. Illyas, "Speech Translation into Pakistan Sign Language.," 2012.

[4] H. Luqman and S. Mahmoud, "Automatic translation of Arabic text to Arabic sign language," Univ Access Inf Soc, vol. 18, pp. 939-951, 2019.

[5] A. Abbas and S. Sarfraz, "Developing a Prototype to Translate Text and Speech to Pakistan Sign Language with Bilingual subtitle: A Framework," journal of Educational Technology, vol. 47, no. 2, pp. 248-266, 2018.

[6] "Sign Language Accessibility for The Deaf In Pakistan," 2018. [Online]. Available: https://www.yesprograms.org/stories/sign-languageaccessibility-for-the-deaf-in-pakistan.

[7] C. P. Nephrology, "Auditory Screening," ScienceDirect, 2008. [Online]. Available: https://www.sciencedirect.com/topics/medicineand-dentistry/auditory-screening.

[8] Deaf Welfare Awareness Foundation, "Deaf Welfare Awareness Foundation," 2021. [Online]. Available: http://www.dwaf.org.pk/.

[9] "Institutes for Hearing Impaired Children," Special Education Department Government of Punjab, [Online]. Available: https://sed.punjab.gov.pk/hic.

[10] "Deaf and Dumb Schools," KPK Government Report, [Online]. Available: https://kp.gov.pk/page/deafanddumbschools.

[11] N. S. Khan, A. Abid and K. Abid, "A Novel Natural Language Processing (NLP)–Based Machine Translation Model for English to Pakistan Sign Language Translation," Springer Science+Business Media, LLC, part of Springer Nature 2020, 2020.

[12] OpenCV, "OpenCV-About," [Online]. Available: https://opencv.org/about/. [Accessed 7 June 2021].

[13] ProgrammerSought, "keras-ocr is a packaged and flexible version of the CRAFT text detector and Keras CRNN recognition model - Programmer Sought," [Online]. Available: https://www.programmersought.com/article/23613451910/.

[14] E. Sawrna, "English Typed Alphabets and Numbers Dataset," 25 Sep 2018. [Online]. Available: https://www.kaggle.com/passionoflife/english-typedalphabets-and-numbers-dataset?select=EnglishFnt.tgz.

[15] G. Krishnan, "Which is the best OCR tool in the market?" 2019. [Online]. Available: https://www.ideas2it.com/blogs/ocr-tool-comparison/.

[16] D. Sporici, E. Cus, nir and C. Boiangiu, "Improving the Accuracy of Tesseract 4.0 OCR Engine Using Convolution-Based Preprocessing.," 2020.

[17] C. Liao, "“OCR Engine Comparison — Tesseract vs. EasyOCR," 2020. [Online]. Available: https://medium.com/swlh/ocr-engine-comparisontesseract-vs-easyocr-729be893d3ae.

[18] T. Han and A. Hickman, "“Our Search for the Best OCR Tool, and What We Found," [Online]. Available: https://source.opennews.org/articles/somany-ocr-options/.

[19] M. Audichya, Saini and J. Kumar, "A Study to Recognize Printed Gujarati Characters Using Tesseract OCR," International Journal for Research in Applied Science & Engineering Technology, vol. 5, no. 9, pp. 1505- 1510, 2017.

[20] "Keras-ocr.readthedocs.io. 2021. keras-ocr — keras ocr documentation.," 2021. [Online]. Available: https://keras-ocr.readthedocs.io/en/latest/. [Accessed 7 June 2021].

[21] A. Majumdar, "Python moviepy: Automate Video Editing in Python! - JournalDev," 2021. [Online]. Available: https://www.journaldev.com/46531/python-moviepy-video-editing. [Accessed 17 Jan 2021].

[22] N. Kross, ""Editing Video with Python + MoviePy", Engineering Education (EngEd) Program," 2021. [Online]. Available: https://www.journaldev.com/46531/python-moviepy-video-editing.

[23] Pythonprogramming.net, "Python Programming Tutorials-Regular Expressions with re Python Tutorial," 2014. [Online]. Available: https://pythonprogramming.net/regular-expressions-regex-tutorialpython-3/. [Accessed 9 June 2021].

[24] Pythonprogramming.net, "Python Programming Tutorials-Chunking with NLTK," 2015. [Online]. Available: https://pythonprogramming.net/chunking-nltk-tutorial/.

[25] N. Bachani, ""Chunking in NLP: Decoded," 2020. [Online]. Available: https://towardsdatascience.com/chunking-in-nlp-decoded-b4a71b2b4e24. [Accessed 9 June 2021].

[26] S. Bird, E. Klein and E. Loper, "Natural Language Processing with Python," Sebastopol: O’Reilly Media, p. chapter 7, 2009.

[27] G. Rachiele, "Tokenization and Parts of Speech (POS) Tagging in Python’s NLTK library," 19 2 2018. [Online]. Available: https://medium.com/@gianpaul.r/tokenization-and-parts-of-speechpos-tagging-in-pythons-nltk-library-2d30f70af13b.

[28] G. A. Miller, "WordNet: A Lexical Database for English," Communications of the ACM, vol. 38, no. 11, pp. 39-41, 1995.

[29] C. Fellbaum, "WordNet: An Electronic Lexical Database.," MA: MIT Press., Cambridge, 1998.

[30] T. Simpson and T. Dao, "WordNet-based semantic similarity measure," [Online]. Available: http://www.codeproject.com/KB/string/semanticsimilaritywordnet.aspx.

[31] U. Farooq, M.S.M. Rahim, N. Sabir, A. Hussain, and A. Abid, 2021. Advances in machine translation for sign language: approaches, limitations, and challenges. Neural Computing and Applications, 33(21), pp.14357-14399.

[32] N.K. Kahlon, and W. Singh, 2023. Machine translation from text to sign language: a systematic review. Universal Access in the Information Society, 22(1), pp.1-35.

[33] A. Kumar, V. Goyal, L. Goyal, 2023. State of the art of automation in sign language: A systematic review. ACM Transactions on Asian and Low-Resource Language Information Processing, 22(4), pp.1-80.