Towards an Arabic Sign Language (ArSL) corpus for deaf drivers

Samah Abbas [Corresp., 1] , **Hassanin Al-Barhamtoshy** [2] , **Fahad Alotaibi** [3]

[1] Management Information Systems Department, Faculty of Economics and Administration, King Abdul Aziz University, Jeddah, Mecca, Saudi Arabia

[2] Information Technology Department, Faculty of Computing & Information Technology, King Abdul Aziz University, Jeddah, Mecca, Saudi Arabia

[3] Information System Department, Faculty of Computing & Information Technology, King Abdul Aziz University, Jeddah, Mecca, Saudi Arabia

Corresponding Author: Samah Abbas
Email address: sabbas@kau.edu.sa

Sign language is a common language that deaf people around the world use to communicate with others. However, normal people are generally not familiar with sign language (SL) and they do not need to learn their language to communicate with them in everyday life. Several technologies offer possibilities for overcoming these barriers to assisting deaf people and facilitating their active lives, including natural language processing (NLP), text understanding, machine translation, and sign language simulation. In this paper, we mainly focus on the problem faced by the deaf community in Saudi Arabia as an important member of the society that needs assistance in communicating with others, especially in the field of work as a driver. Therefore, this community needs a system that facilitates the mechanism of communication with the users using NLP that allows translating Arabic Sign Language (ArSL) into voice and vice versa. Thus, this paper aims to purplish our created dataset dictionary and ArSL corpus videos that were done in our previous work. Furthermore, we illustrate our corpus, data determination (deaf driver terminologies), dataset creation and processing in order to implement the proposed future system. Therefore, the evaluation of the dataset will be presented and simulated using two methods. First, using the evaluation of four expert signers, where the result was 10.23% WER. The second method, using Cohen's Kappa in order to evaluate the corpus of ArSL videos that was made by three signers from different regions of Saudi Arabia. We found that the agreement between signer 2 and signer 3 is 61%, which is a good agreement. In our future direction, we will use the ArSL video corpus of signer 2 and signer 3 to implement ML techniques for our deaf driver system.

Towards an Arabic Sign Language (ArSL) Corpus for Deaf Drivers

6   Samah Abbas [1], Hassanin Al-Barhamtoshy [2], Fahad Alotaibi [3]

8   [1] Management Information Systems Department, Faculty of Economics and Administration,
9   King Abdulaziz University, Jeddah, Saudi Arabia
10   [2] Information Technology Department, Faculty of Computing & Information Technology, King
11   Abdulaziz University, Jeddah, Saudi Arabia
12   [3] Information System Department, Faculty of Computing & Information Technology, King
13   Abdulaziz University, Jeddah, Saudi Arabia

16   Corresponding Author:
17   Samah A. Abbas [1]
18   Almuhammadia, Jeddah, Mecca, 23622, Saudi Arabia
19   Email address: sabbas@kau.edu.sa

Abstract

22   Sign language is a common language that deaf people around the world use to communicate
23   with others. However, normal people are generally not familiar with sign language (SL) and they
24   do not need to learn their language to communicate with them in everyday life. Several
25   technologies offer possibilities for overcoming these barriers to assisting deaf people and
26   facilitating their active lives, including natural language processing (NLP), text understanding,
27   machine translation, and sign language simulation. In this paper, we mainly focus on the problem
28   faced by the deaf community in Saudi Arabia as an important member of the society that needs
29   assistance in communicating with others, especially in the field of work as a driver. Therefore,
30   this community needs a system that facilitates the mechanism of communication with the users
31   using NLP that allows translating Arabic Sign Language (ArSL) into voice and vice versa. Thus,
32   this paper aims to purplish our created dataset dictionary and ArSL corpus videos that were done
33   in our previous work. Furthermore, we illustrate our corpus, data determination (deaf driver
34   terminologies), dataset creation and processing in order to implement the proposed future
35   system. Therefore, the evaluation of the dataset will be presented and simulated using two
36   methods. First, using the evaluation of four expert signers, where the result was 10.23% WER.
37   The second method, using Cohen's Kappa in order to evaluate the corpus of ArSL videos that
38   were made by three signers from different regions of Saudi Arabia. We found that the agreement
39   between signer 2 and signer 3 is 61%, which is a good agreement. In our future direction, we will
40   use the ArSL video corpus of signer 2 and signer 3 to implement ML techniques for our deaf
41   driver system.

Introduction

Deaf people use sign language to communicate with their peers and normal people who know their sign language. Therefore, sign language is the only way to communicate with deaf people, although it receives even less attention from normal people. Furthermore, sign language is not a unified language among deaf people around the world, as each country has its own sign language. For example, American Sign Language (ASL), British Sign Language (BSL), Australian Sign Language (Auslan), Indian Sign Language (ISL) and Arabic Sign Language (ArSL). In particular, the Arab countries that use ArSL, such as the Gulf, Al-Sham, and some North African Arab countries, have similarities and differences in sign language. The reason for this is the differences between in dialects. As a result, deaf people face problems communicating with the community in many aspects while they are working or practicing their daily lives, such as health, education, and transportation. One of the solutions for these problems is using a sign language interpreter, of their work or daily life, such as health, education, and transportation. One solution to these problems is the use of a sign language interpreter, a person who knows sign language and can interpret it to normal people. However, this solution is not optimal because of loss of privacy and professional independence (Forestal,2001; Broecker,1986). Researchers have therefore developed certain technologies that allow for the provision of a computer interpreter instead of a human. For example, sign language recognition technology, machine translation (MT). In addition, some Arab researchers and organizations, such as the Arab League Educational, Cultural and Scientific Organization (ALECSO), have made efforts to unify ArSL by introducing the first dictionary in 1999 (Al-Binali& Samareen,2009).

In Saudi Arabia, deaf people have some difficulty communicating with others. In particular, when deaf people are driving a vehicle and non-deaf people are sitting as passengers. The non-deaf passenger cannot understand the deaf driver. Also, a non-deaf passenger cannot describe the location needed by the deaf driver. However, there are many mobile applications that can be used to facilitate communication, such as Tawasol and Turjuman, but they still do not provide real-time translation (Team Mind Rockets,2017; Al-Nafjan, Al-Arifi & Al-Wabil,2015). In addition, non-deaf passengers or drivers do not necessarily download the sign language translation application and use it only once. In the area of deaf drivers, Saudi Arabia faces a lack of technology that can improve communication between deaf drivers and non-deaf passengers. To our knowledge, no research has yet been conducted to introduce a solution in the area of deaf drivers in Saudi Arabia. In addition, there is a lack of corpus of ArSL videos being constructed in the transportation domain, especially for deaf drivers.

The purpose of this research is to publish our dataset as a continuation of previous work (Abbas, 2020). Also, we aim to publish our corpus in order to implement ML as future work. This paper is organized as follows: The second section illustrates some research work done to create videos and images of the standard corpus of sign language in each country as a database. In order to conduct their experiments to implement some systems that help deaf people in specific areas. We have also mentioned some research done in the Arab countries. In order to create an ArSL corpus by focusing on certain areas. The third section illustrates the linguistic background of ArSL, speech recognition and machine translation (MT) systems, while the fourth section focuses on the design architecture of ArSL. In the fifth section, we discuss the processing and data collection modules and evaluation videos of the deaf driver corpus that necessitated the implementation of ArSL in the deaf driving context. Finally, we illustrate future research directions.

Research question

88   Our research aims to answer these two questions:

89   1. Can we create a data corpus and video corpus for deaf drivers?

90   2. Can we evaluate the video corpus created of deaf drivers?

91 Literature Review of ArSL Recognition and MT System

92   This literature review illustrates various standard sign language corpora for some countries of
93   the world. In fact, when we start designing and developing a system for deaf people, we need the
94   essential corpora of videos or images. In addition, these corpora are considered as a reference
95   database that must be rich with a large and complete volume of standard sign language. They
96   will help researchers to implement machine learning techniques and develop specific systems for
97   deaf people.

98   Starting with ASL, researchers developed approximately 2576 videos that included movements,
99   hand shapes, and sentences (Martínez et al.,2002). Another researcher collected and annotated
100   videos around a corpus of 843 sentences from Boston University and RWTH Aachen University
101   (Dreuw et al.,2008). In Britain, approximately 249 participants completed a conversational
102   dataset of the BSL video corpus (Schembri et al.,2018). In Indian sign language, researchers
103   implement 1440 gestural videos of IPSL by 9 Indian signers (Kishore et al.,2011). In Brazil, they
104   have created the LIBRAS-HCRGBDS database of the Libras video corpus which has
105   approximately 610 videos by 5 signers (Porfirio et al.,2013). In Korean Sign Language (KSL),
106   the researchers developed 6 thousand vocabulary words in the KSL corpus database (Kim,Jang
107   & Bien, 1996). In Arabic Sign Language (ArSL), 80 Arabic signers implemented a sensor-based
108   dataset for 40 sentences (Assaleh, Shanableh & Zourob, 2012). In addition, some researchers
109   have produced a corpus of 500 videos and images. It contains hand shapes, facial expressions,
110   alphabet, numbers. In addition, the movement in simple signs, continuous sentences and
111   sentences with lip movement were performed by 10 signers. These corpuses are called
112   SignsWorld Atlas. However, still the performed corpus does not cover all words and phrases as a
113   database (Shohieb, Elminir & Riad,2015). Recently, the researchers led their efforts to build the
114   KArSL database as a comprehensive ArSL reference database for numbers, letters and words.
115   They used Kinect devices to make these ArSL videos available to all researchers (Sidig et al.,
116   2021).

117   Specifically, in ArSL, researchers are dedicating their efforts to improving the focus of accuracy
118   in a specific domain, such as medicine and education. Some researchers used 150 video signs
119   from the Java programming domain in their experiment (Al-Barhamtoshy et al.,2019). In the
120   same educational field, two researchers introduced an intelligent system for deaf students based
121   on the image. This system helped the deaf student in the educational environment. Therefore,
122   they created the dictionary that needs it in life activities and academic environments (Mohammdi
123   & Elbourhamy, 2021). Some of them performed some sentences in Arabic language without
124   focusing on a specific field. Some focused on the jurisprudence of prayer and their symbols as
125   datasets (Almasoud & Al-Khalifa, 2012). Other researchers used 600 phrases in the health
126   domain (Luqman & Mahmoud, 2019). Similarly, some researchers have implemented their
127   research using the data created in the field of education (El, El&El Atawy,2014; Almohimeed,
128   Wald & Damper,2011). In our research, we focus on a different domain that can help deaf people
129   in the work environment, such as driving a cab. The corpus we created will help deaf drivers to
130   communicate with non-deaf passengers while deaf people drive their cars.

131 ArSL and Linguistic Background

132   ArSL shows huge complexities in phonology, morphology, and structure, which is not the case
133   for other sign languages. These complexities are explained below.

Phonology of ArSL

134
135 Phonemes are mental representations, which is just a way to empty what is inside the brain.
136 Phonemes are made up of four elements: 1) The shape of the hand. 2) Orientation of the hand. 3)
137 Position of the hand. 4) Direction of the hand in motion (Schechter, 2014). In phonology, these
138 four elements are known as hand features (MFs). In particular, in Sign Language, there are MFs
139 and also Non-Manual Features (NMFs) are involved. NMFs refers to emotional parts of the
140 body, for example, lip motion, facial expression, shoulder movement, head movement, eyelids
141 and eyebrows. In general, in ArSL, we use both MFs and NMFs to give the correct meaning,
142 called essential NMFs. On the other hand, if the signer uses just only the MFs, the meaning will
143 turn into another meaning that the signer did not mean to express (Johnston& Schembri, 2007).

Morphology and Structure of ArSL

144
145 The rules of grammar in ArSL are not the same as those in Arabic. The differences are: verb
146 tenses, differences between singular and plural, rules for prepositional and adverb and gender
147 signs. Regarding sentence structure, ArSL uses only the Subject-Verb-Object (SVO) structure
148 instead of the SVO, OVS and VOS structures (Luqman& Mahmoud, 2019).

ArSL and Design Architecture

149
150 New technologies that support communication have a significant impact on human life. For
151 deaf people, the developers and researchers have tried to use some new technologies to make the
152 life of deaf people easier by developing automated systems that can help them to communicate in
153 various aspects of their life with others. In this section, we will illustrate the brief explanation of
154 some of the techniques used in order to implement the automated system for better
155 communication between the community and among themselves.

Machine Translation

156
157 Machine Translation (MT) is a standardized name for the system that relies on computer analysis
158 to translate between two languages. It is used for text and speech using artificial intelligence (AI)
159 and natural language processing (NLP). For example, translating from source "English" to target
160 "Arabic" (Pradhan, Mishra& Nayak, 2020; Verma, Srivastava& Kumar, 2015). MT is also used
161 to translate text or speech into video sign language or avatar. There are several approaches to MT
162 that can be used depending on what we need to translate and what constitutes a better translation.
163 One of these approaches is direct translation, without regard to grammar rules. To improve the
164 quality of MT, the rule-based approach was introduced, which consists of parsing the source and
165 target language. Another approach is corpus-based, which deals with massive data containing
166 sentences. There are also knowledge-based approaches, which take into account the
167 understanding of the source and target text in the context of linguistic and semantic knowledge.
168 Finally, there is google translation which is developed by Google (Verma, Srivastava& Kumar,
169 2015).

Arabic Speech Recognition

170
171 Speech Recognition (SR) is a computerized system that converts speech into text or signs.
172 This system is used to communicate between humans and machines. It is also known as
173 automatic speech recognition. Machine-based SR is a complicated task due to differences in
174 dialects, contexts and speech styles. To reduce this complexity in SR, the system can exploit the
175 repetition and structure of the token speech signal as multiple sources of knowledge. SR sources
176 are constructed based on knowledge of phonology, syntax, phonemes, grammar and phrases
177 (Katyal, Kaur& Gill, 2014; Chow et al., 1987). In addition, the SR system has several classifiers.
178 1) The speech utterance, which is composed of separate, connected, continuous and spontaneous

179   speech. 2) The speaker model, which contains one of two dependents that was designed for a
180   specific speaker or independently for different speakers. 3) The size of vocabulary (Radha&
181   Vimala).
182   In terms of the SR process, there are four steps in which SR can be implemented. Analyze the
183   speech signal, then extract the feature using different techniques to identify the vector, such as
184   MFCC (Mel-frequency cepstral coefficient). Then, we build a model using different techniques
185   like HMMs with the training dataset. In the last step, we test the model in the matching setting,
186   taking the dissection and measuring the performance based on the error rate (Saksamudre&
187   Deshmukh, 2015; Yankayiş, Ensari& Aydin) as shown in Figure 1.

188   **ArSL Gestures Recognitions**

189   Gesture recognition is defined as the ability of the computer to understand gestures and
190   execute commands based on the gestures made. The first gesture recognition system was
191   introduced in 1993 as a kind of user interface for perceptual computing that helps capture human
192   gestures and transfer them into commands using a computerized system. It is used with many
193   technologies, especially in the field of games, such as X-box, PlayStation, and Wii Fit. These
194   games use Just Dance and Kinect Sports, which recognizes the hand and certain body parts
195   (Schechter, 2014; Darrell& Pentland, 1993).
196   In the sign language domain, the gesture recognition system uses the following processes: 1)
197   Recognize the deaf signs. 2) Analyze the sign. 3) Converting that sign into meaningful text
198   (words or phrases), voice, or expressions that non-deaf can understand. In addition, there are two
199   main methods for gesture recognition in ArSL, namely vision-based devices (video or image)
200   and wearable-based devices. Each of them has advantages and disadvantages. One advantage of
201   wearable devices is that there is no need to search for background and lighting. The disadvantage
202   is that they interfere with movement. On the other hand, the advantage of vision-based
203   technology is the ease of movement, while the disadvantage is the effect of changing the
204   background and lighting (Ambar et al., 2018; Paudyal et al., 2017).
205   In addition, each has different processes and techniques. In the vision-based process, one or
206   more cameras are the main tools that must be available to use this method. On the other hand, the
207   wearable-based method depends on certain types of equipment and computers mainly. In terms
208   of process, the wearable-based method spells the alphabet by reading the specific information in
209   each sensor of the finger or glove joint. However, the vision-based method has certain steps,
210   which are the follows:
211   • Image capture, which consists in using a camera to collect data (building the corpus) and
212   analyze the collected images.
213   • Pre-processing, which consists of preparing the images and identifying the information
214   according to color (segmentation).
215   • Feature extraction uses some techniques like root mean square (RMS) to identify the
216   feature vector.
217   Classification that classifies based on the feature vector to build the model (Mahmood&
218   Abdulazeez, 2019; Derpanis, 2004). The vision-based gesture recognition process is shown in
219   the Figure 2.

220   **Deaf Driver Corpus**

221   For creating a deaf driver corpus, we divided the processes that we will use for this creation
222   into 4 modules. The four modules are preprocessing, recording, assessment, and validation
223   modules. These high-level and low-level approaches are represented in (Figure 3) and (Figure 4).
224

Data Collection, Creation and Annotation

In this section, we explain our data collection, data determination (deaf driver terminologies) and data set creation and processing to create our video corpus through the preprocessing and recording module.

In the preprocessing module, we first collected data at two levels: a word or phrase level dataset and a sentence level dataset to create a small Arabic dictionary. This dictionary is divided into eight sections (categories). 1) Welcoming "Salam Alaikum and How are you? ". 2) Directions "Left and Right". 3) Place "School and Deaf Association ". 4) Traffic and Transportation " Driver's license and Traffic lights". 5) Sentences used by deaf drivers when they need to talk with their passengers, e.g., "We have arrived" 6) Sentences used by passengers when they need to talk with their deaf drivers, e.g., "I do not have cash to pay the amount ". 7) General Words " No and Yes and In and On ". 8) Amount "Dollar, Riyal and 2 Riyals until 100 Riyals " (Abbas, 2020).

The total words and phrases (sentences) is 215. Some of these words and sentences fall under the general communication, collected from the Saudi Sign Language Dictionary, 2018 edition (Saudi Association for Hearing Impairment, 2018). Some of them are collected in the contextual domain of the normal conversation that was done between Saudi cab drivers and passengers. We mean in the contextual domain is each country has its own contextual domain in the payment process. For example, Saudi Arabia has its own curranty which is Riyal. Table 1 shows a part of the Arabic dictionary we created.

In the recording module, we made our videos at a rate of about 30 FPS (frames per second), by a camera for our ArSL corpus. These videos were taken by one of the ArSL expert signers who is not deaf. Figure 5 shows an ArSL corpus captured by a non-deaf expert signer. In addition to that, we made video captures with three expert signers from different occasions in Saudi Arabia. Two of them are totally deaf, while the third is hard of hearing. Figure 6 represents an ArSL corpus captured for each deaf expert. To record 215 words containing simple phrases and signs, we took about forty-five minutes continuously with the non-deaf expert signer. However, the three deaf expert signers took approximately fifty minutes. The expert signers used only one hand if it was appropriate for the context of the deaf driver, unless the sign required the use of both hands. Next, we segmented the video using VEGAS (Video Production Software - Unleash Your Creativity | VEGAS, n.d.) video editing software (we segmented each sign containing words or phrases that matched our dictionary into a single video, with a total number of 215 videos). To support future work, we added Arabic audio to each video and labeled it with Arabic text that refers to the same recorded ArSL.

Dataset Evaluation

These collected and created data will be used for the evaluation and validation of our ArSL corpus. In this section, we have explained the evaluation module while the validation will be done in the future work.

In the evaluation module, we evaluated the video of each corpus generated based on our created dictionary using two techniques. First, we used a human expert evaluation technique. Second, we used a statistical method of measuring the agreement between three deaf experts signing in ArSL that we recorded.

In the first technique, which is a human expert evaluation, we made the evaluation based on the views of four participants who are experts in ArSL. One of them is completely deaf and the second is hard of hearing. Two other participants are not deaf, they work in a deaf club and are experts. We used the quantitative approach (questionnaire) and divided it into two sections. The

271 first section was a demographic questionnaire (gender - age - education level - whether he/she is
272 deaf or not). In the second section, we introduced an evaluation of the video based on the word
273 (phrase) or related sentence. We attached each video to each related phrase or word. Participants
274 were asked to evaluate the 215 sign videos to see if the video recorded for each word or phrase
275 was correct or not. If not correct, it meant that the video was not related to that particular word or
276 phrase. We asked the four participants to choose one of the types of correction they had to make
277 for each video (add - replace - delete). The evaluation method is explained in (Table 2) using
278 some videos evaluated by one of the experts.

279   The evaluation results of the deaf experts based on the word error rate (WER) for each
280 category (section) such as welcome, directions, and place are explained in Figure 7. This means
281 that for each category, the evaluation result of these videos was wrong. Also, they should be
282 replaced with correct videos.

283   As can be seen in Figure 7, first, the hospitality category has a high percentage of WERs, at
284 50%. Second, the location category has a percentage of about 36%. Thus, we need to reduce the
285 WER in our video corpus that is related to these two categories (sections) in order to improve the
286 communication between drivers and deaf passengers and also to describe the correct location for
287 the passenger's destination.
288   The total WER of our video corpus was 10.23% as shown in Table 3.
289   To resolve the corpus of erroneous ArSL videos, we again captured these videos based on the
290 corrected signs that the raters explained to us.
291   In the second technique of measuring the agreement between the signs of each of the two deaf
292 experts, we used Cohen's Kappa criterion using Eq. (1).

293
294
295 $$K = \frac{P_0 - P_e}{1 - P_e} \qquad (1)$$
296

297   Whereas *P0* represent the number of observed proportional agreement between two variables
298 by using Eq. (2).
299
300 $$P0 = \frac{1}{n} \; \Sigma_{i=1}^{g} fii \qquad (2)$$
301

302   Number of agreements that expected by chance represented by *Pe* and the formula represents as
303 Eq. (3).
304
305 $$P_e = \frac{1}{n^2} \; \Sigma_{i=1}^{g} fi + f + i \qquad (3)$$

306
307   Where (fi+) is the total of row, and (f+i) is the total of column (Viera& Garrett,2005).
308   When we use Cohen's Kappa (*K*) , we can achieve one of the six types that represents in the
309 Table 4.
310   For more details, the evaluation was done by a non-deaf coder who annotated each of the two
311 ArSL video corpora on the basis of true and false independently. First, we measured the
312 agreement between signer 1 and signer 2. Next, we measured the agreement between signer 1
313 and signer 3. Finally, we measured the agreement between signer 2 and signer 3. Therefore, by

314 applying Cohen's Kappa statistical method to measure the agreement between each of the two
315 videos in the ArSL corpus, we found the result represented in Table 5.
316   As we can see in (Table 5) The agreement between signer 2 and signer 3 reached about 60%,
317 which is better than the others. The reason is that they are from the same school, the Deaf
318 Association. However, signatory 1 is a volunteer who is not from their school. It should be noted
319 that the first two are completely deaf, while the second is hard of hearing.   In the validation
320 module, in the future work, we will implement the ML technique using python as a programming
321 language. To do this, we will use the ArSL video corpus that was made by signer 2 or signer 3.
322 Which one got the best agreement based on Cohen's Kappa method for evaluation. Next, we will
323 divide our data into training and testing datasets to measure accuracy and error rate.

324 Conclusions and Future Work

325   Through this research, we have reviewed the previous studies that have been conducted in the
326 field of ArSL corpus generation as a standard sign language used by different countries. We have
327 also illustrated the different areas in which researchers are conducting their efforts to create a
328 data dictionary and annotated ArSL corpus, especially in Saudi Arabia. We have clarified the
329 difficulties encountered in translating ArSL from a grammatical, semantic and syntactic
330 perspective. How they affect the accuracy of translation and recognition. Finally, we described
331 the ArSL dictionary for deaf drivers and explained the data collection processes to construct the
332 videos in our corpus. These were recorded using a single camera and then verified using two
333 methods. First, using the evaluation of 4 participants, experts in ArSL, two of whom were deaf.
334 Second, by using Cohen's Kappa statistical method to measure the agreement between each of
335 the two videos in the ArSL corpus recorded by three signers.
336   The created ArSL corpus provides opportunities to test various feature extraction methods and
337 recognition techniques. Extending and validating the dataset using machine learning (ML) will
338 be the focus of future work. In addition, this corpus will be used to design our proposed system
339 to facilitate communication with deaf drivers.

340 References

341 Abbas SA. 2020. Improving Arabic Sign Language to Support Communication between Vehicle Drivers and
342   Passengers from Deaf People. *King Abdulaziz University : Scientific Publishing Centre*, Saudi Arabia.
343 Al-Barhamtoshy HM, Abuzinadah NE, 3, A, 4, TF, 5 and AA, Allinjawi AA. 2019. Development Of An Intelligent
344   Arabic Text Translation Model For Deaf Students Using State Of The Art Information Technology. *Bioscience*
345   *Biotechnology Research Communications* 12:338–345. DOI: 10.21786/bbrc/12.2/17.Al-Binali M, Samareen S.
346   2009. Grammar of the Unified Qatari Arabic Sign Language. *Dar Al-Sharq, Doha,* Qatar, In Arabic.
347 Alfi AEEE, Basuony MMRE, Atawy SME. Intelligent Arabic text to Arabic Sign Language Translation for Easy Deaf
348   Communication
349 Almasoud AM, Al-Khalifa HS. 2012. SemSignWriting: A Proposed Semantic System for Arabic Text-to-SignWriting
350   Translation. 2012. DOI: 10.4236/jsea.2012.58069.
351 Almohimeed A, Wald M, Damper RI. 2011. Arabic Text to Arabic Sign Language Translation System for the Deaf
352   and Hearing-Impaired Community. In: *Proceedings of the Second Workshop on Speech and Language*
353   *Processing for Assistive Technologies*. Edinburgh, Scotland, UK: Association for Computational Linguistics,
354   101–109.
355 Al-Nafjan A, Al-Arifi B, Al-Wabil A. 2015. Design and Development of an Educational Arabic Sign Language
356   Mobile Application: Collective Impact with Tawasol. In: Antona M, Stephanidis C eds. *Universal Access in*
357   *Human-Computer Interaction. Access to Interaction*. Lecture Notes in Computer Science. Cham: Springer
358   International Publishing, 319–326. DOI: 10.1007/978-3-319-20681-3_30.
359 Ambar R, Fai CK, Wahab MHA, Jamil MMA, Ma'radzi AA. 2018. Development of a Wearable Device for Sign
360   Language Recognition. *Journal of Physics: Conference Series* 1019:012017. DOI: 10.1088/1742-
361   6596/1019/1/012017

362 Assaleh K, Shanableh T, Zourob M. 2012. Low Complexity Classification System for Glove-Based Arabic Sign
363   Language Recognition. In: Huang T, Zeng Z, Li C, Leung CS eds. *Neural Information Processing*. Lecture Notes
364   in Computer Science. Berlin, Heidelberg: Springer, 262–268. DOI: 10.1007/978-3-642-34487-9_32.
365 Broecker B. 1986. Speaking the Language of Sign: The Art and Science of Signing (review). *American Annals of the*
366   *Deaf* 131:199–200. DOI: 10.1353/aad.2012.0792.
367 Chow Y, Dunham M, Kimball O, Krasner M, Kubala G, Makhoul J, Price P, Roucos S, Schwartz R. 1987. BYBLOS:
368   The BBN continuous speech recognition system. In: *ICASSP '87. IEEE International Conference on Acoustics,*
369   *Speech, and Signal Processing*. 89–92. DOI: 10.1109/ICASSP.1987.1169748.
370 Derpanis K. 2004. A Review of Vision-Based Hand Gestures. *Department of Computer Science York University*.
371 Dreuw P, Neidle C, Athitsos V, Sclaroff S, Ney H. 2008. Benchmark Databases for Video-Based Automatic Sign
372   Language Recognition. :6.
373 Forestal LH. A Study of Deaf Leaders' Attitudes Towards Sign Language Interpreters and Interpreting. Ph.D. Thesis.
374   United States -- New York: New York University
375 Katyal A, Kaur A, Gill J. 2014. Automatic Speech Recognition: A Review. *International Journal of Engineering*
376   *and Advanced Technology (IJEAT) ISSN* 2249-8958.
377 Kim J-S, Jang W, Bien Z. 1996. A dynamic gesture recognition system for the Korean sign language (KSL). *IEEE*
378   *Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 26:354–359. DOI: 10.1109/3477.485888.
379 Kishore PVV, Kumar PR, Kumar EK, Kishore SRC. 2011. Video audio interface for recognizing gestures of indian
380   sign. *International Journal of Image Processing (IJIP)* 5:479.
381 Luqman H, Mahmoud SA. 2019. Automatic translation of Arabic text-to-Arabic sign language. *Universal Access in*
382   *the Information Society* 18:939–951. DOI: 10.1007/s10209-018-0622-8.
383 Mahmood MR, Abdulazeez AM. 2019. Different Model for Hand Gesture Recognition with a Novel Line Feature
384   Extraction. In: *2019 International Conference on Advanced Science and Engineering (ICOASE)*. 52–57. DOI:
385   10.1109/ICOASE.2019.8723731.
386 Martinez AM, Wilbur RB, Shay R, Kak AC. 2002. Purdue RVL-SLLL ASL database for automatic recognition of
387   American Sign Language. In: *Proceedings. Fourth IEEE International Conference on Multimodal Interfaces*.
388   167–172. DOI: 10.1109/ICMI.2002.1166987.
389 Mohammdi H, Elbourhamy D. 2021. An intelligent system to help deaf students learn Arabic Sign Language.
390   *Interactive Learning Environments*. DOI: 10.1080/10494820.2021.1920431.
391 Porfirio AJ, Wiggers KL, Oliveira LES, Weingaertner D. 2013. LIBRAS Sign Language Hand Configuration
392   Recognition Based on 3D Meshes. In: *2013 IEEE International Conference on Systems, Man, and Cybernetics*.
393   1588–1593. DOI: 10.1109/SMC.2013.274.
394 Pradhan I, Mishra SP, Nayak AK. 2020. A Collation of Machine Translation Approaches with Exemplified
395   Comparison of Google and Bing Translators. In: Gunjan VK, Garcia Diaz V, Cardona M, Solanki VK, Sunitha
396   KVN eds. *ICICCT 2019 – System Reliability, Quality Control, Safety, Maintenance and Management*.
397   Singapore: Springer, 854–860. DOI: 10.1007/978-981-13-8461-5_97.
398 Radha V, Vimala C. 2012. A review on speech recognition challenges and approaches. *doaj. org* 2:1–7.
399 Saksamudre S, Deshmukh R. 2015. Isolated Word Recognition System for Hindi Language. *International Journal of*
400   *Computer Sciences and Engineering* 3:110–114.
401 Saudi Association For Hearing Impairment. 2018. Saudi Sign Language Dectionary, Riyadh.
402 Schechter S. 2014.What is gesture recognition? Gesture recognition defined. *Available at*
403   *https://www.marxentlabs.com/what-is-gesture-recognition-defined/* (accessed August 4, 2021).
404 Schembri A, Fenlon J, Rentelis R, Reynolds S, Cormier K. 2013. Building the British Sign Language Corpus.
405   *Language Documentation & Conservation* 7:136–154.
406 Shohieb SM, Elminir HK, Riad AM. 2015. SignsWorld Atlas; a benchmark Arabic Sign Language database. *Journal*
407   *of King Saud University - Computer and Information Sciences* 27:68–76. DOI: 10.1016/j.jksuci.2014.03.011.
408 Sidig AAI, Luqman H, Mahmoud S, Mohandes M. 2021. KArSL: Arabic Sign Language Database. *ACM*
409   *Transactions on Asian and Low-Resource Language Information Processing* 20:14:1-14:19. DOI:
410   10.1145/3423420.
411 Team Mind Rockets. 2017. Mind Rockets Inc, Assistive Technologies for the deaf. Available at
412   *http://mindrocketsinc.com* (accessed 08 Nov 2018).
413 Video Production Software — Unleash Your Creativity | VEGAS. Available at
414   *https://www.vegascreativesoftware.com/au/?AffiliateID=152&phash=CWd33ae7JRr95cbA&gclid=CjwKCAj*
415   *wgISIBhBfEiwALE19SU084G0eMYydNHuMFAyKgKFC8UTxNc9eZJqkoAgZENGvBd26E9EAZBoCV-*
416   *8QAvD_BwE* (accessed July 28, 2021).

417   Viera AJ, Garrett JM. 2005. Understanding Interobserver Agreement: The Kappa Statistic. *Family Medicine*:4.
418   YANKAYIŞ M, Ensari T, Aydin N. 2018. Performance Evaluation of Feature Extraction and Modeling Methods for
419   Speaker Recognition. 4.
420
421

Figure 1

A diagram for recognizing spoken the Arabic language to Arabic written text (Abba, 2020)

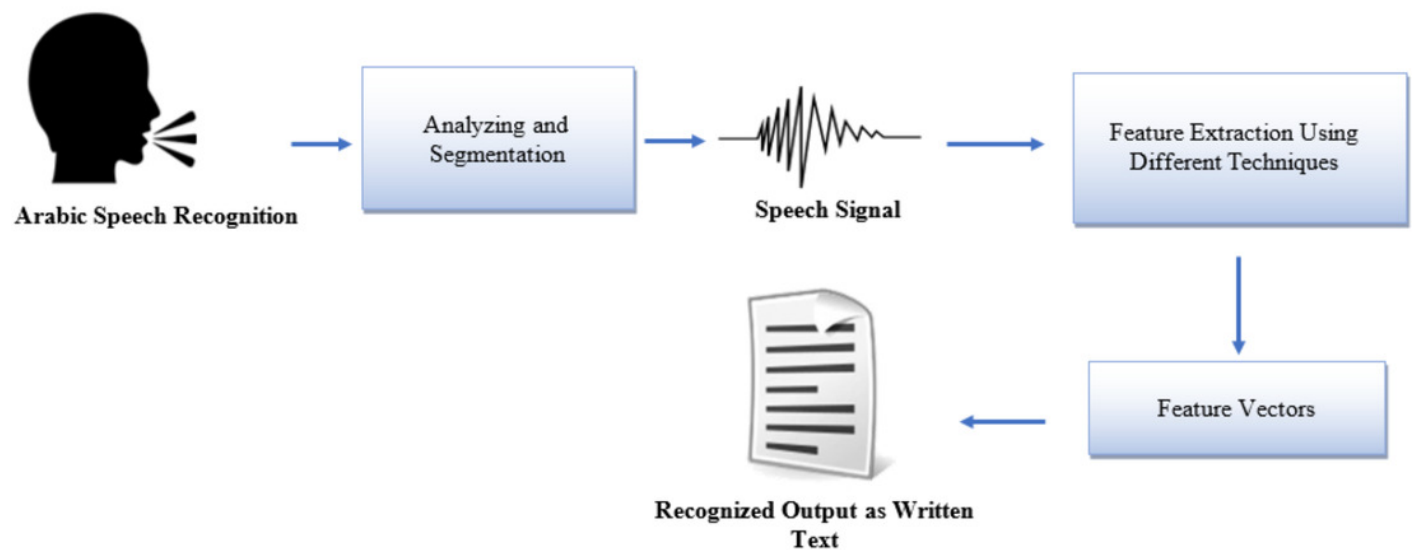



Figure 2

A diagram for the processing of ArSL gesture recognition Adapted from (Derpanis, 2004)

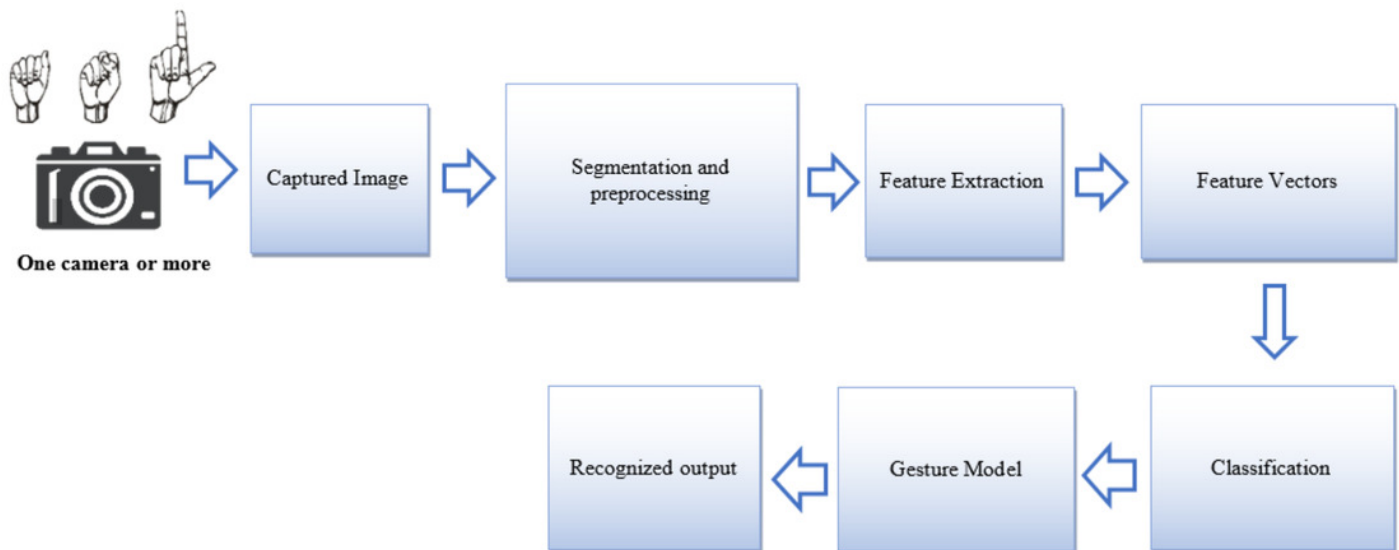



Figure 3

High-level approach of our dataset (corpus) collection (Abba, 2020)

Figure 4

Low-level approach of our dataset (corpus) collection (Abbas,2020)

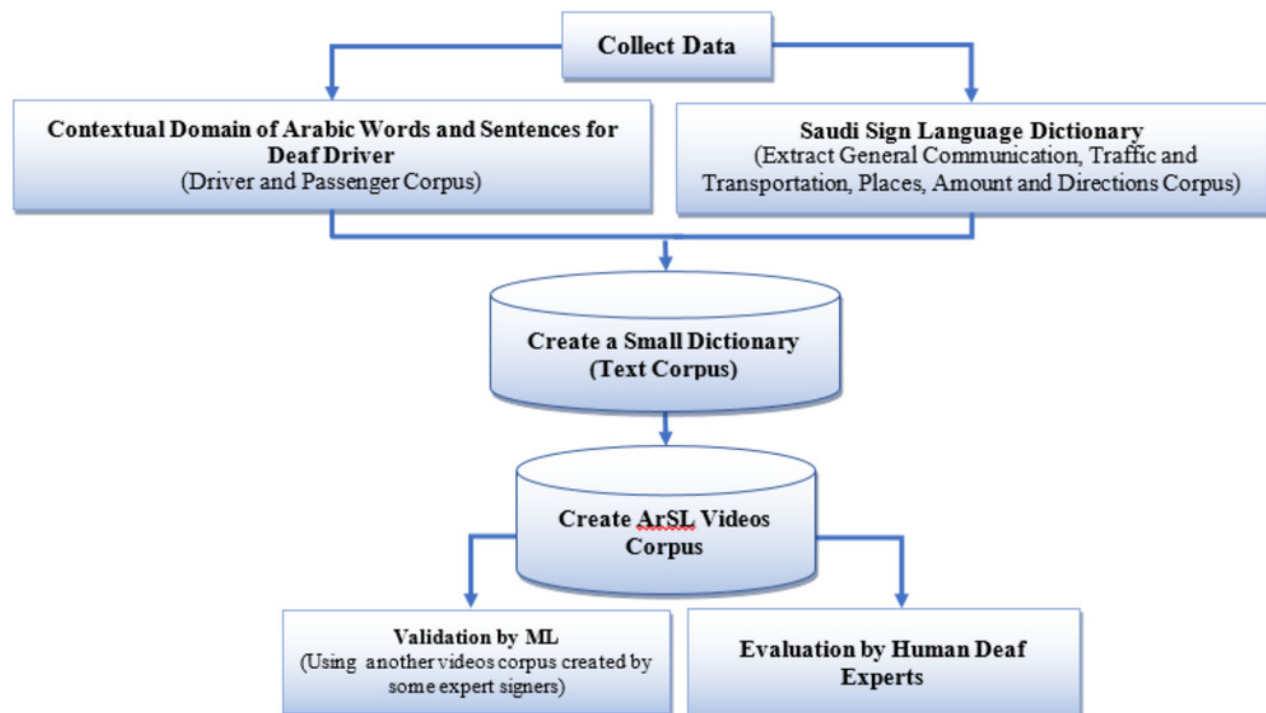



Figure 5

One Captured Video from our ArSL Video Corpus (Abbas, 2020)

Figure 6

Three Captured ArSL Video Corpus

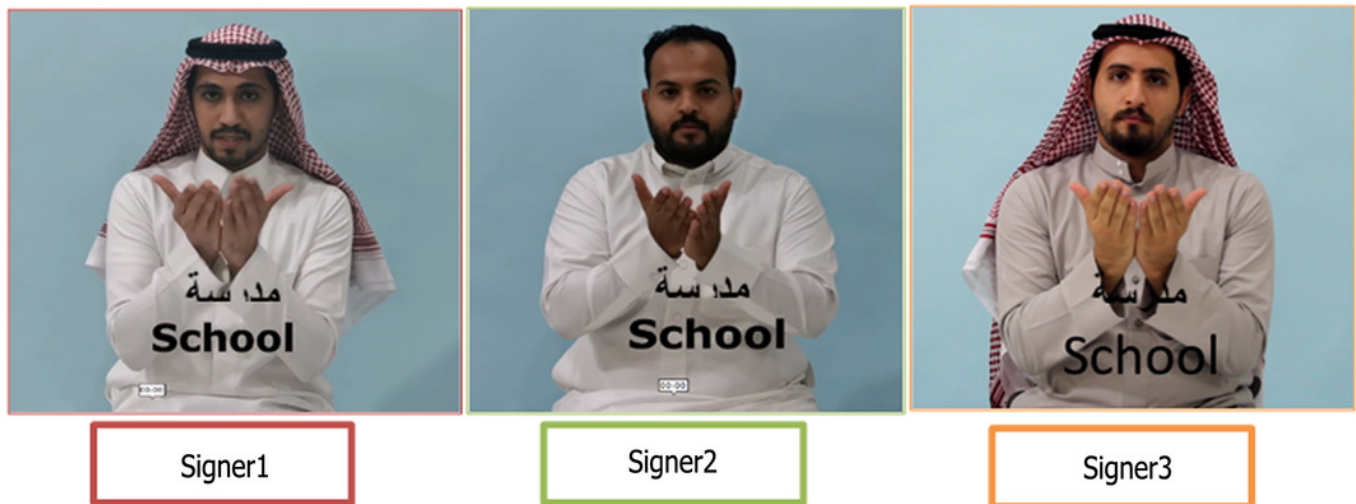
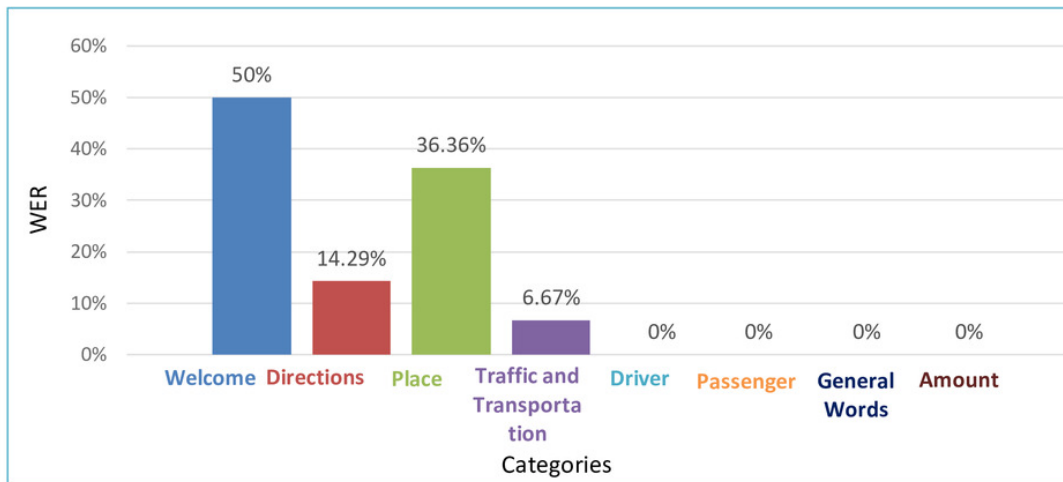



Figure 7

The Deaf Experts' Evaluations Results for Each Category Using (WER) (Abbas,2020)

An Example of our Created Dictionary (Abbas, 2020)

2
3

Word / Phrase / Sentences
Section 1: Welcome
1.Salam Aleikum (peace be upon you)
2.How are you?
Section 2: Directions
1. Left
2. Right
Section 3: Place
1. School
2. Deaf Association
Section 4: Traffic and Transport
1. Driving License
Section 5: Driver
1. We have arrived
Section 6: Passenger
1. I do not have cash to pay the amount

An example of Evaluation with one Expert (Abbas,2020)

2
3
4
5
6
7

Section One: Welcome				
Correction if the translation is wrong: The video of sign needs.			**Video evaluation based on the related word (phrase) or sentences.** × or √	**Word \ phrase \ sentences**
Deleting	**Replacing**	**Adding**		
			√	1-Salam aleikum
			√	2-Waleikum salam
	√		×	3-Good morning

Total Word Correct and Error Rate for our Created Video Corpus (Abbas,2020)

2

	Correct	Wrong
Signs' Videos Corpus	89.8%	10.2%

Categories of Cohen's Kappa interpretation (Viera&Garrett,2005)

Kappa (K)	Type of agreement
$K \leq 0$	No agreement
$0.01 \leq K \leq 0.20$	Poor agreement
$0.21 \leq K \leq 0.40$	Fair agreement
$0.41 \leq K \leq 0.60$	Moderate agreement
$0.61 \leq K \leq 0.80$	Good agreement
$0.81 \leq K \leq 1.00$	Perfect agreement

The Result of evaluation using Cohen's Kappa

2
3

	Number of observed agreements	Number of agreements expected by chance	Kappa	Agreement
Signer1 & Signer2	78.14%	63.97%	39%	**Fair Agreement**
Signer1 & Signer3	78.14%	66.39%	35%	**Fair Agreement**
Signer2 & Signer3	*87.91%*	*69.03%*	*61%*	**Good Agreement**