AUTOMATIC ARABIC SIGN LANGUAGE RECOGNITION: A REVIEW, TAXONOMY, OPEN CHALLENGES, RESEARCH ROADMAP AND FUTURE DIRECTIONS

Ahmad Sami Al-Shamayleh[1], Rodina Ahmad[2], Nazean Jomhari[3], Mohammad A. M. Abushariah[4]

[1,2,3] Department of Software Engineering, Faculty of Computer Science and Information Technology, University of Malaya, 50603, Kuala Lumpur, Malaysia

[4] Department of Computer Information Systems, King Abdullah II School of Information Technology, The University of Jordan, 11942, Amman, Jordan

E-mail: ashamayleh@siswa.um.edu.my[1], rodina@um.edu.my[2], nazean@um.edu.my[3], m.abushariah@ju.edu.jo[4]

DOI: https://doi.org/10.22452/mjcs.vol33no4.5

ABSTRACT

Sign language is still the best communication mean between the deaf and hearing impaired citizens. Due to the advancements in technology, we are able to find various research attempts and efforts on Automatic Sign Language Recognition (ASLR) technology for many languages including the Arabic language. Such attempts have simplified and assisted the interpretation between spoken and sign languages. In fact, the technologies that translate between spoken and sign languages have become popular today. Being the first comprehensive and up-to-date review that studies the state-of-the-art ASLR in perspective to Arabic Sign Language Recognition (ArSLR), this review is a contribution to ArSLR research community. In this paper, the research background and fundamentals of ArSLR are provided. ArSLR research taxonomies, databases, open challenges, future research trends, and directions, and a roadmap to ArSLR research are presented. This review investigates two major taxonomies. The primary taxonomy that is related to the capturing mechanism of the gestures for ArSLR, which can be either a Vision-Based Recognition (VBR) approach or Sensor-Based Recognition (SBR) approach. The secondary taxonomy that is related to the type and task of the gestures for ArSLR, which can be either the Arabic alphabet, isolated words, or continuous sign language recognition. In addition, less research attempts have been directed towards Arabic continuous sign language recognition task compared to other tasks, which marks a research gap that can be considered by the research community. To the best of our knowledge, all previous research attempts and reviews on sign language recognition for ArSL used forehand signs. This shows that the backhand signs have not been considered for ArSL tasks, which creates another important research gap to be filled up. Therefore, we recommend more research initiatives to contribute to these gaps by using an SBR approach for signers' dependent and independent approaches.

Keywords: ***ArSLR, alphabet sign language, isolated words recognition, continuous sign recognition, hand gestures, deaf community, hearing impaired.***

1.0 INTRODUCTION

For hearing impaired to indulge in social activities and opportunities, Sign Language (SL) is considered as the most natural and expressive way. Due to the advancement in Automatic Sign Language Recognition (ASLR), communication capabilities and social integration for speech and hearing impaired are improved [1][2]. Real-time ASLR is a multidisciplinary research area that involves image segmentation, pattern recognition, Computer Vision (CV) and Natural Language Processing (NLP). ASLR is a global issue due to the complication of the shapes of control. It requires an understanding of hands position, shape, motion and orientation [3][4][5]. Functional ASLR systems can be utilized to create speech and text, making the deaf and hearing-impaired citizens more autonomous. The most difficult part of any ASLR system is to detect, analyze and recognize the simplest and easiest hand gestures that must be recognized in the picture [6].

SL is a visible gesture language, which incorporates face, hands, and arms to pass on considerations and implications. It is produced in the deaf community to integrate deaf and hearing-impaired citizens with their peers, friends, relatives, and society at large. They utilize their hands, faces, arms, and body for communication. It is important to note that there is no international SL [7][8][9]. In fact, every country or region has its own SL, which will be discussed further in this paper. These SLs can be different from one country to another in terms of syntax and grammar. The SL that is used in Australia, for instance, is known as Australia Sign Language (Auslan), whereas the SL used in Japan is known as Japanese Sign Language (JSL), and the SL used in the Arab world is known as Arabic Sign Language (ArSL) etc. [10] [11] [12]. More details on various world standard SLs are shown in Table 3.

The SL is used among the deaf and hearing impaired citizens, including friends and relatives who have the same impairment, as well as the interpreters. In addition, the SLs are normally not familiar out of the scope of these communities; hence, communication constraints emerge between deaf and hearing impaired citizens with those outside the scope of their SLs.

Research efforts on ArSLR can be divided into two approaches such as Vision-Based Recognition (VBR) and Sensor-Based Recognition (SBR), which form our primary research taxonomy. These two approaches are generally used for the Arabic alphabet, isolated words, and continuous sign language recognition tasks. This serves as our secondary research taxonomy. It has been discovered that the VBR approach is most widely used in ArSLR compare to SBR approach. In addition, there have been many research attempts on the Arabic alphabet and isolated words sign language recognition compare to Arabic continuous sign language recognition [13].

This review is purposefully aiming at applying the prior knowledge and idea acquired about the automatic Arabic sign language in order to discuss difficulties in ArSLR, the taxonomy of ArSL approaches and algorithms. In addition, the taxonomy that comprises the challenges, the promising growth of this technology and the direction of further investigation is also highlighted. Furthermore, some related literature on the automatic Arabic sign language field, which consists of conference articles, certified journals, and laboratory reports, has also been widely investigated and analyzed.

This research has been conducted to provide a comprehensive review for Arabic sign language recognition including its taxonomies, difficulties, and approaches. However, the study carried out an extensive literature review in order to bridge the research gap, focusing on the related literature published within the period of 2001 to 2017. In addition, it also provides a comprehensive discussion on the fundamentals of sign language recognition, including its architecture, challenges, world standard SLs, and their available databases.

Taxonomy of research challenges and opportunities for Arabic sign language recognition alongside the potential research directions are highlighted and summarized in this article. The main objectives of this research include:

1. To identify the fundamentals and background of Arabic language and Arabic sign language recognition, including its difficulties, architecture, and world standard SLs and their available databases. Interpreting existing research conducted within this domain.
2. To interpret the current studies carried out in this field.
3. To classify the Arabic sign language recognition approaches and techniques.
4. To identify the core research gap should be further investigated by researchers in the domain.
5. To identify the potential areas and a roadmap that requires future research considerations.

This review is split into six sections, Section 2.0, provides sufficient research on the fundamentals of sign language recognition, including its architecture, taxonomies, and world standard SLs and their available databases. Section 3.0 presents the detail literature review and background on Arabic sign language recognition including its taxonomies, difficulties, approaches, and algorithms. A brief description of a unified Arabic sign language dictionary is also provided in Section. Section 4.0 provides the taxonomy of the research challenges and opportunities for Arabic sign language recognition, whereas Section 5.0 highlights and summarizes the future research roadmap to Arabic sign language recognition. The summary and conclusion of the study are finally presented in Section 6.0.

2.0 FUNDAMENTALS OF SIGN LANGUAGE RECOGNITION

The attentions of most researchers has been shifted to ASLR research, as it is now applied in many domains like machine control in the industrial domain [14][15], communication system for deaf and hearing-impaired people [16] [17], Human-Computer Interaction (HCI) [18], Virtual Reality (VR) [19][20], and many others. ASLR research is divided into two main groups, namely; 1) Static sign language. 2) Dynamic sign language [21] [22], as shown in Fig 1.

The position of the hand and its orientation are usually used to correct the static gestures (which consists of hand poses and postures) and dynamic gestures (which includes hand movements with a certain type like waving) in a given space and time without making any kind of movement. In addition, static gestures could involve a single hand orientation without performing any movement [1].

Static gesture takes single frames of signs as input, while the dynamic gesture takes continuous frames of signs in a video as input. Moreover, the static gesture is usually dependent on the angular direction and the shape of the finger while the position of the hand remains stationary within the duration of the gesture. In dynamic gesture. On the order hand, the position of the hand in dynamic hand gesture is continuously in motion with regards to the time and its message is a content of the stroke phase sequence. The dynamic gesture is subdivided into three phases of the motion such as stroke phase, retraction, and preparation [23].

According to the study carried out in [22], static gestures are characterized with orientations, shape, finger's flex angles, relative position to body, and context environment, while dynamic gestures are characterized with orientations, shape, finger's flex angles, hand trajectories and orientations, motion speed and direction, and scale.

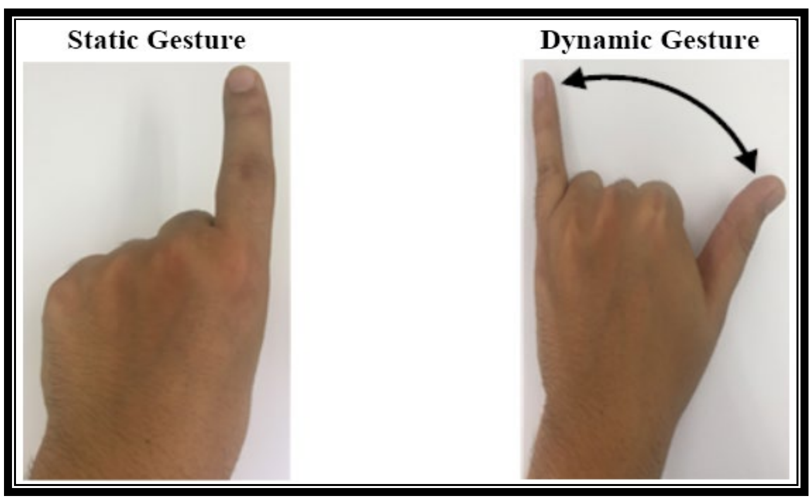

Fig. 1: Example for static and dynamic gesture

2.1 Architecture of Automatic Sign Language Recognition

ASLR comprises of four major phases in order to identify the correct gestures, namely; 1) data acquisition, 2) pre-processing and segmentation, 3) feature extraction and finally 4) classification as shown in Fig. 2. Another example of phases of processing model or cycle model of SLR and ArSLR is given in [134]. Once the hand image is captured using a suitable input device, the image is then segmented in order to find the location of the hand position within the parts of the body starting from the background. Next, the processing of the location image is then carried out in order to eliminate noises, identify contours and produce a suitable model. Again, after the pre-processing of the images and gestures, there comes a feature extraction process in which the shape position, orientation, movements and the location of the hands are been extracted for the classification purpose. Lastly, those captured images are now identified as suitable gestures on the bases of analysis and modeling [24][25]. Generally, there are three levels of SLR, which are the alphabets, isolated-word, and continuous sentences [134]. The detailed description of the ASLR Architectural phases is shown in Fig. 2 and sub-section below.

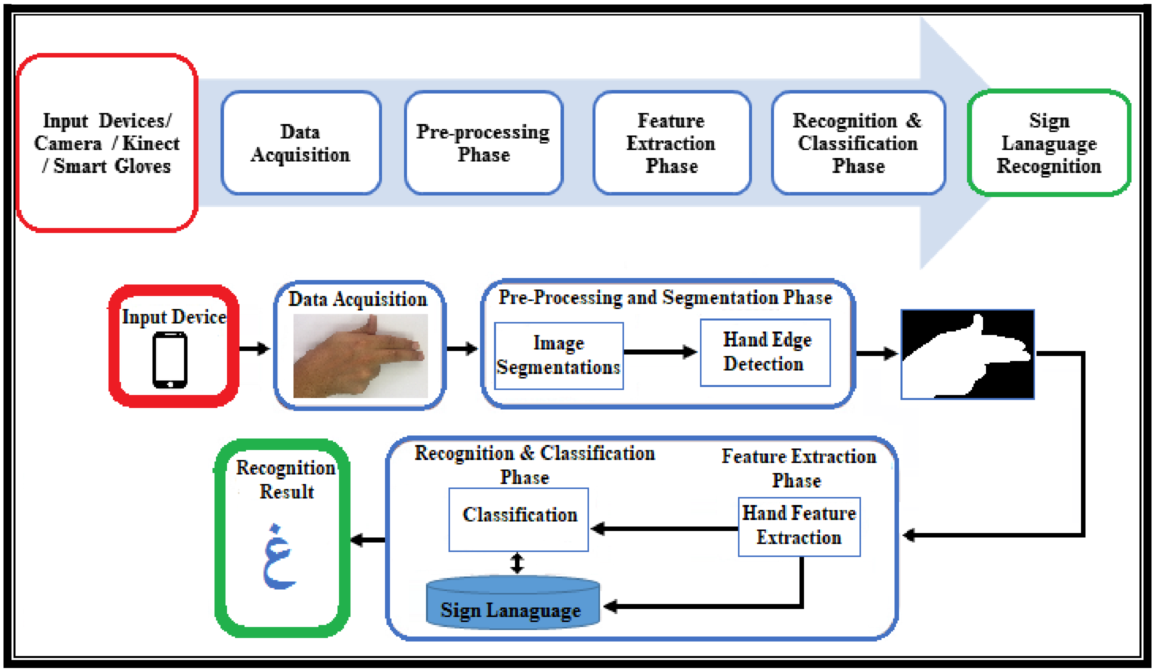

Fig. 2: Architectural design of the Automatic Sign Language Recognition

2.1.1 Data Acquisition Phase

The perfect the data acquisition is, the better and more efficient hand gesture recognition can be [26][27]. The frame of images is normally acquired as input, which is collected using recognized cameras which can be informed of stereo, webcam, thermal or video. In addition, state-of-the-art equipment such as leap monitor and Kinect can also be used. It is important to note that 3D cameras such as stereo cameras, Kinect and LMC can collect depth information [28]. This can be done by selecting a suitable input device to be used in data acquisition. Many input devices are available for use in data acquisition; including but not limited to the hand images (acquired from drawings), Kinect 3D sensor, marker, data gloves, stereo camera and webcam [25][29][30]. Arabic sign language recognition can acquire data by using two different approaches, namely; 1) vision-based approach and 2) Sensor-based approach.

A. Vision-based approach:

In this approach, the hand gesture images can be acquired by employing a video camera. In gesture recognition, this approach is made up of appearance and 3D hand model approach. The main gestures capturing technology in vision-based approach found in [28] study are:

1. Invasive techniques depend on body markers including colored gloves, wristbands, and LED lights.
2. Active techniques which use Kinect and LMC (Leap motion controller) for light projection.
3. A single camera such as smartphone camera, webcam, and video camera.
4. Stereo camera, which provides extensive information by employing various monocular cameras.

B. Sensor-based approach

In this approach, the location, hand motion, and velocity can be captured by the help of sensors and instruments. The main gestures capturing technology in sensor-based approach according to the study in[28] are:

1. Inertial measurement unit: this form of measurement uses accelerometer and gyroscope to measure the location, the intensity of recognition, and the acceleration of the fingers. The user's hands information orientation and motion can be obtained accurately by these sensors at a high frame rate (e.g., the Xsens MTw IMU has a frame rate of 50Hz) [31].

2. Wi-Fi and Radar that senses the changes in the strength of the signals in the air by using by using electromagnetic signs.

3. Electromyography (EMG) that detects the finger motion by taking the measurement of the electrical pulse in human muscle and reducing the bio-signal.

4. Others that utilize haptic technologies, mechanical, electromagnetic, ultrasonic, and flex sensors.

2.1.2 Pre-processing and Segmentation phase (Gesture Modeling Phase)

This phase is a vital part of conducting success gesture recognition. The various data collected from such devices should be modeled in a proper way through the appropriate application type. The respective modeling shall be carried out via different steps, namely; 1) pre-processing and segmentation, 2) noise/filter removal, 3) contour/edge detection, and 4) normalization [32].

Image pre-processing phase: this is applied on image or video inputs for system performance improvement and noise reduction using median and Gaussian filters followed by morphological operations in the acquired images or videos. In most studies, the pixel representation acquired is reduced into lesser sizes before moving into the next processing stages. In this method, the pixels of the input signals is lowered which is capable of enhancing the computing performance. The variation of the input images acquired under diverse atmosphere to unite the resolution and brightness of the image can be enhanced by employing histogram equalization [28].

Segmentation phase: The purpose of the segmentation phase is to partition the image into multiple distinct forms in such a way that the ROI (i.e. region of interest) is separated from the extra images. Segmentation phase can be in form of two states, such as contextual segmentation and non-contextual segmentation. The contextual segmentation considers the geometrical relationships within the features; for example, the identification of the edge technique. In the order hand, the non-contextual form groups the pixels based on the global attributes [28]. In Vision-based ASLR, structure, hand segmentation is the most important and challenging step in comparison with gesture recognition. Fig. 3 presents the main classification of image segmentation whereas Table 1 presents a comparison of various segmentation techniques.

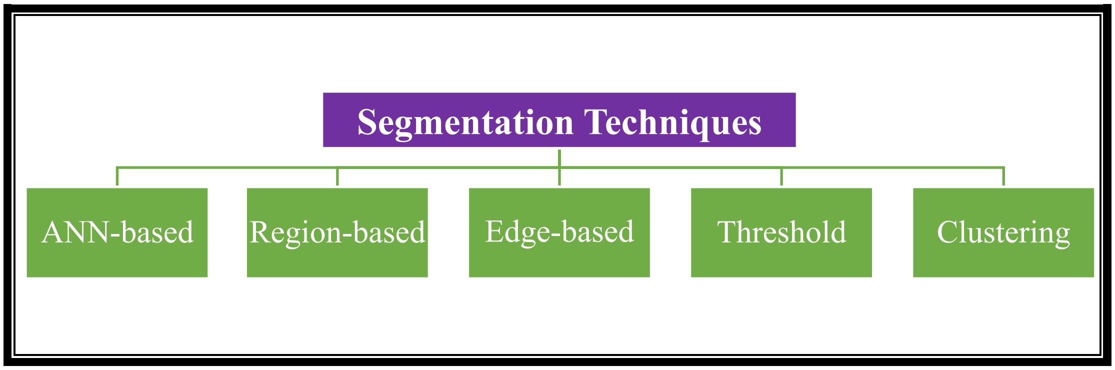

Fig. 3: General image segmentation techniques

Table 1: Comparison of various segmentation techniques

Segmentation Method	Description	Types	Advantages	Disadvantages
Artificial Neural Network Based (ANN-based)	-Based on the simulation of learning process for decision making -Mostly used for the segmentation of medical images.	-This method including two steps, Extracting features and segmentation by neural network.	-No need to write complex programs.	-More wastage of time in training.
Region Based	-Based on partitioning image into homogeneous regions.	-Region growing methods. -Region splitting and merging methods.	-More immune to noise, useful when it is easy to define similarity criteria.	-Expensive method in terms of time and memory.
Edge Based	-Based on discontinuity detection.	-Roberts edge, -Sobel edge, -Prewitt edge, -LoG edge -Canny edge.	-Good for images having better contrast between objects.	-Not suitable for wrong detected or too many edges.
Thresholding	-Based on the histogram peaks of the image to find particular threshold values.	- Global Threshold -Variable Threshold (Local and Adaptive) -Multiple Threshold	-No need of previous information, -simplest method in image segmentation. -Reduces the complexity	-Highly dependent on peaks, spatial details are not considered.
Clustering	-Based on division into homogeneous clusters.	-Hard Clustering -Soft Clustering	-Fuzzy uses partial membership therefore more useful for real problems.	-Determining membership function is not easy.

2.1.3 Feature Extraction Phase

Perfect segmentation process leads to perfect features extraction process [33][34]. The features are considered as the basic elements in order to establish hand gesture recognition. The feature extraction in gesture recognition context is expected to consist of suitable information obtained from the input to the hand gestures. However, this feature is depicted in a concise version as the gesture identity that is preserved for classification purpose outside the gestures obtained from the other parts of the body.

We can make use of a huge number of features extraction such as shaping, distance, motion, contour, textures, Centre of gravity, orientation, velocity and etc. For hand gesture recognition, we can identify hand gesture by using geometric features such as finger detections, hand contour, and fingertips. However, these features are neither available at all times nor reliable due to illuminations and occlusions. On the other hand, there are non-geometric features that can be used (such as Texture, silhouette, and color) that are also available for recognition [25] [35].

The important feature extraction techniques used in sign language are Shift-invariant feature transform (SIFT), Principal component analysis (PCA), Speeded up robust features (SURF), Linear discriminant analysis (LDA), Convexity defects and K-curvature [28] [36] [37].

1. **SIFT:** It the feature extraction method called rotation invariant and scale introduced by Lowe [38], which uses multiple scale technique for detection. In addition, the image is described by its interest point and Gaussian function is used to re-scaled and soften the image at each interval of the pyramid [28].

2. **PCA:** This is a numerical operation that obtains a set of value of imbalance variable (also known as a fundamental component) through the transformation of values of a balanced variable using orthogonal revolution [39].

3. **LDA:** This technique is used to find the definite combined features that best separate the object classes through augmentation of the adaptable class [28] [40]. In addition, the LDA technique is normally used as definite classifiers and in the dimension reduction[28]. It should be noted that PCA has nothing to do with the class differences but rather concentrate on determining the order of the highest variance that exists between the features [41]. However, both PCA and LDA techniques can be applied in determining the definite feature combination that gives the detail explanation of data.

4. **SURF:** This technique is established upon shift-invariant feature transformation that builds multiple scale pyramid (by using a difference-of-Gaussian operator to rotate the lower and upper scale of the image) and searches the scale space that contains the local extreme. In addition, SURF reduces the size of the image using a filtering approach rather than the iterative method. The scale-space can be obtained in SIFT by using difference-of-Gaussian (DoG) to approximate the LoG (Laplcian of Gaussian) [28].

5. **K-curvature and Convexity defects:** This technique is used to extract features in the palm center, convex defect and hull, the center of palm and the fingertips angle. Several types of researchers identified the gesture features using the universal features together with the convexity defects [28].

6. **Features extraction in frequency domain:** In this feature extraction technique, the frequency domain such as Fourier, Cosine, and Wavelet Transform are obtained by transforming the input value of the time domain [28].

2.1.4 Classification Phase

The last phase of the recognition system is represented by the classification of hand gestures. This phase has to be considered in order to make recognition technique and an effective classification algorithm available, which are useful in many gesture recognition research. This phase moves alongside with pattern recognition domain and machine learning. Hand gestures can be classified using rule-based and machine learning-based approaches [25] [35], which are further discussed in the following sections.

A. **Rule-based Approach:** In this approach, a number of manually encoded relationships also called rules between feature inputs are developed. Therefore, features are extracted from the input gesture and compared with the encoded rules. The matched rule is finally taken as a gesture. This approach lacks the human ability for encoding the rules and limits the success of the recognition process [35].

B. **Machine Learning-based Approach:** Due to the shortcoming of rule-based approach in gesture recognition as stated in the previous section, many researchers have turned into machine learning approach in order to find the mappings that exists in-between the gestures and high-dimensional feature sets. Machine learning-based approach considers gesture as the result of stochastic processes.

   Classification algorithm such as Conditional random fields (CRF) [1][42], K-means [1] [43], K-nearest neighbor (K-NN) [1] [44], Mean-shift clustering model [1] [45], support vector machine (SVM) [1] [46], Hidden Markov Models (HMMs) [1] [47], Dynamic Time Warping (DTW) [1] [48], Time-delay neural network (TDNN) [1] [49], and Finite-State Machine (FSM) [1] [50], Artificial Neural Networks (ANNs) [1] are used in sign language and gesture recognition. Other researchers have also employed Gaussian mixture distribution to achieve gestures classification and Euclidian distance measure [35] [51] [52]. Details of each algorithm and further discussions and comparisons can be found in our Systematic Literature Review [1]

2.2 Corpus of Main World Standard Sign Languages

Based on the World Health Organization (WHO), it is reported that over 5% of the world population suffers from deaf-mute and hard hearing disabilities [53]. Such people utilize hand, head, and/or body gestures to communicate their emotions and ideas [54]. Therefore, every country could have its own SL, which can differ from one country to another.

Before starting the design and development of any ASLR system, it is essential to acquire large volume and comprehensive corpus of signs for a particular standard language. Without sufficient volumes of prepared and ready-to-use training and testing data, it is impractical to develop any ASLR system. This section reports some of the remarkable databases from all around the globe in alphabetical order. We attempt to cover as many standard SL corpus as possible from all over the world. These corpus are particularly generated for the advancement of Machine assisted ASLR systems. Table 2 shows the main standard SL corpuses in different countries.

Table 2: Main standard sign language databases in different countries

Sign language	Abbreviations/ SIL code	Library database name	Description	Ref.
***A**merican Sign Language*	ASL	The Purdue RVL-SLLL Database	*RVL-SLLL* Database is a comprehensive database of *ASL* gestures, movements, words and sentences. It was performed by fourteen signers. This database comprises of 2576 videos of 39 motion primitives, 62 hand shapes, and sentences.	[55]
		RWTH-BOSTON-400	*RWTH-BOSTON-400* is gathered for the development of isolated *ASL* Recognition. It comprises of 843 sentence. It was performed by four signers	[56]
***A**ustralian Sign Language*	Auslan	Auslan Sign bank	*Auslan-Sign-bank* is comprises 7415 words in *Auslan*, the corpus consists more the 1000 separate video clips. It was performed by one hundred signers	[57]
***A**rabic Sign Language*	ArSL	Arabic Sign Language Database	Arabic Sign Language Database comprises of forty sentence. Each sentence was repeated nineteen times. This corpus is fully segmented and labeled database for continuous ArSL. It was performed by eighty signers.	[58]
		Signs World Atlas, a benchmark Arabic Sign Language	The Signs World Arabic Sign Language Database is contain of a picture and video clips. The authors have been improved and developed the database to assess their algorithms and methods for posture recognition and real-time Arabic Sign Language. Database contains about 500 static gestures (manual signs) include “finger spelling, hand motions”, and dynamic gestures (non-manual signs) NMS elements include: lip reading, body language, and facial expressions.	[59]
***B**ritish Sign Language*	BSL	British Sign Language Corpus	*BSL* Corpus is consists of videos presenting conversations of two hundred forty nine participants. The corpus consists annotations of 6330 gestures from the signers in the conversational dataset.	[60]
***B**razilian Sign Language*	Libras	LIBRAS-HC-RGBDS	*LIBRAS-HC-RGBDS* corpus consists of sixty one hand configurations of the Libras. The data were acquired using the Kinect sensor. There are 610 video clips of five signers in the DB.	[61]
***G**erman Sign Language*	DGS	The SIGNUM Database	The SIGNUM Database is Contain of vocabulary size four hundred fifty basic gestures in DGS. Based on this DB, seven hundred eighty sentence. It was performed by twenty five signers	[62][63]
***G**reek Sign Language*	GSL	Greek Sign Language Corpus	Corpus is contain of sentences level gesture samples and respective annotations. Video clips recording of the DB have been performed by four signers.	[64]
Indian Sign Language	IPSL	Indian Sign Language Database	ISL is performed by 9 signers, each signer repeated the gesture 20 times under various conditions with a total number of 1440 gesture video clips for a total of eighty gestures.	[65]
***I**rish Sign Language*	ISL	The ATIS Sign Language	The ATIS corpus is contain of five hundred ninety five gesture and English expressions and sentences	[66]
***I**talian Sign Language.*	LIS	The A3LIS-147	A3LIS-147 is Contain of one hundred forty seven distinct gestures from LIS. DB is organized in 6 groups, depend on various daily life scenarios. This DB have been performed by ten signers.	[67]
***K**orean Sign Language*	KSL	Korean Sign Language corpus	KSL corpus is contain of less than 25 fundamental gestures. However, there are six thousand vocabulary words in KSL database.	[68]
***M**alaysian Sign Language*	MSL	The MSL Database	The MSL Database is contains of isolated gestures and continuous gestures in MSL, Each sign is repeated 20 times.	[69]
***P**akistani Sign Language*	PSL	Pakistani Sign Language Database	A fairly small database. It contain of thirty seven signs. This gestures is generated based on Urdu. Signs in the corpus match to fingerspelling alphabet.	[70]
***P**ersian Sign Language*	PSC	Persian Sign Language Database	***P**ersian Sign Language is contain of six hundred forty image of thirty two gestures that match to the Persian alphabet. The images are taken in controlled environment, and the background is black.	[71]
***S**panish Sign Language*	LSE	Spanish Sign Language Corpus	***S**panish Sign Language Corpus is contain of four thousand eighty Spanish sentences translated into the sign language. The Corpus also have the gestures for all the alphabets, numbers from zero – hundred.	[72]
***T**urkish Sign Language*	TSL/TSM	The Buhmap	The BUHMAP is contains of one hundred thirty two video clips of eight dynamic gesture. This DB are performed by eleven signers.	[73]

3.0 BACKGROUND ON ARABIC LANGUAGE AND ARABIC SIGN LANGUAGE

This section provides detailed background about the difficulties in the Arabic language as well as the difficulties in Arabic sign language. A taxonomy of Arabic sign language recognition approaches is also provided.

3.1 Difficulties in Arabic Language

Being a Semitic, Syntactically and morphologically rich language, vocabularies in the Arabic language are expected to be very huge and larger than in any other languages [74] [75]. In Arabic language construction, words are joined together by using a preposition, conjunctions, pronouns, and articles owing to the fact that the language is so dynamic that differs in agglutinative and sentence structure. However, the concatenation of words is done by filling in the suffixes and prefixes to the word stem, which raises the out-of-vocabulary rate and forms a large list of potential word form for applications such as Automatic Speech Recognition (ASR) [74] [76] [77] [78].

The Arabic language is highly demanded by the NLP community as a result of its challenges and socio-political importance, which are presented by complex morphology, its dialect differences, non-transparent orthography and diglossia [79]. Thus, the Arabic language is one of the most complex natural languages especially when processed using machine learning. The complexity of the Arabic language is due to the following reasons:

1. The existence of distinct characteristics among the three different forms of Arabic language (i.e. in the form of modern standard, classical or Dialectal Arabic) [77] [78] [80].

2. The omission of diacritical marks is modern standard Arabic which makes the readers speculate it from the context [80]. However, mistakes in identifying the right diacritics may result in a different meaning of the same word.

3. The possibility of changes in the letter sequence depending on letter position as a result of the unavailability of the letters for short vowel representation. [81].

4. Every Arabic letter can be written in various ways that can result in two or more forms according to its position in a word, which could be at the beginning of a word, middle, end, or stand-alone [82].

5. The absence of capitalization, the inconsistent and irregular use of punctuation marks causes serious challenges to many NLP tasks which include POS tagging, NER (named entity recognition), parsing, tokenization, and many others [79] [81].

6. The difficulty in the adaptation of the Arabic language with other languages initially developed with the NLP tools as a result of peculiar features inherent in the Arabic language [81].

7. Typographically, the Arabic character set is different from the Latin character set. In order to view Arabic fonts and process Arabic scripts correctly, computers need to be Arabic enabled [83].

8. The creation of Arabic word comprise of joining of suffixes and prefixes to the word stem from various linguistic rule. This causes limitation in structure analysis and root recovery especially in computerization and linguistic theory [79].

3.2 Difficulties in Arabic Sign Language Recognition

Sign language may vary based on the country or even region because it is not considered as a universal language. In the Arabic nations such as Gulf States, Jordan, Tunisia, Algeria, Iraq, Palestine, Syria, Sudan, and Djibouti, there have been many efforts to establish and standardize the sign language, in order to expand it among the impaired hearing people within those nations. Consequently, similar sign letters are used to produce various sign languages in those Arabic-speaking nations [84] [85]. Fig. 4 shows the gestures used in ArSL alphabets.

ArSL contains more than nine thousand signs and gestures, and also uses twenty-six (26) and five (5) for both static and dynamic hand gestures representation of Arabic language letters [86] [84] [87].

In Arab countries, assisting deaf students to identify suitable environments is the main aim of establishing hard-of-hearing schools. One gesture (sign) is produced by combining specific and clear handshapes, palm orientations and movements. However, sign language is considered as a universal native language in hearing for impaired and deaf. ArSL is considered as the official language for deaf people in Arab countries.

Stated below are examples of the various problems that are normally faced by ArSL researchers when translating into Arabic sign language [88] [89]:

1. Lack of syntax, linguistic and morphology studies on Arabic sign language.
2. The huge size of the translation corpus while generating and creating an Arabic sign language translation technique.
3. The complication in representing Arabic sign language translation output.
4. Difficulty in finding a method that evaluates Arabic sign language translation output.

Arabic sign language is similar to other sign languages in that they are gestural and spatial languages. However, Arabic sign language differs in terms of syntax, morphology and language structure from Arabic spoken language. Consequently, it becomes hard to compare the sign language with the corresponding spoken form similar to other languages due to the fact that many concepts in spoken languages are described inadequately for sign languages [90].

In Arabic sign language, signed and gestured sentences do not have a plural, dual, and singular agreements as shown in Table 3. There are many countable nouns in the Arabic language, but they do not exist in the Arabic sign language. For example, the Arabic word /**فُنْجَانَيْنْ**/ /***fundʒaa:najn***/ which means ***two cups*** in the English language is represented in the Arabic sign language by two consecutive words as follows: the sign/gesture for the word /**فُنْجَانْ**/ / ***fundʒaa:n*** / which means ***cup*** in English language, followed by the sign/gesture of the Arabic number /**إِثْنَيْنْ**/ /***ʔiθnajn***/ which means ***two*** in English language, in order to indicate the dual.

Table 3: Syntax in Arabic sign language

No	Arabic Language Syntax	Arabic Sign Language Syntax
1	Singular	Singular
2	Dual	Singular + Two
3	Plural	Singular + Corresponding Number

On the other hand, the Arabic sign language does not benefit from grammatical tenses similar to those that exist in written and spoken forms. Tenses in Arabic sign language are easily and practically applicable. Tenses, which consists of present, past, and future are pointed out in the beginning (starting points) of a discussion and a conversation, and as well switches when there is a need to indicate another tense. In addition, interrogatives and negatives can be expressed in more than one form. Facial expressions, longer signing time, dramatization and repetition are used for emphasis, (for instance, the relationship that exists among different hands is used for manual explanation of adverbs. Furthermore, sentence boundaries, conditional expressions, and turn-taking can be achieved by using nominal features obtained from the facial expression and there is no existence of passivation and declension features. [88]. Table 4 presents the syntax variation between the Arabic language and the Arabic sign language.

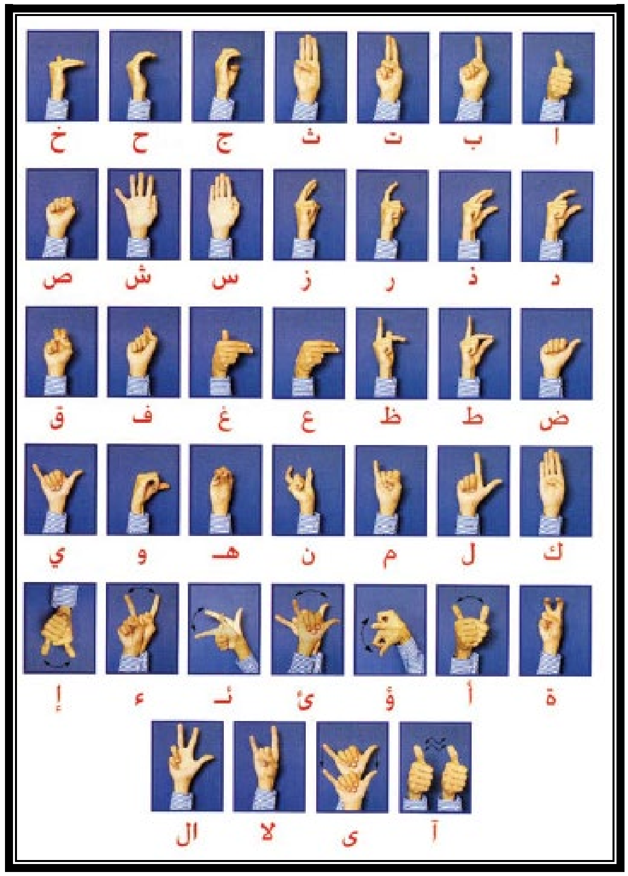

Fig. 4 ArSL alphabet [91]

Table 4: Syntax variation between Arabic language and Arabic sign language [92]

No	Arabic Language Syntax	Arabic Sign Language Syntax
1	Subject + Verb	Subject + Verb
2	Verb + Subject	Subject + Verb
3	Subject + Predicate	Subject + Predicate
4	Subject + Verb + Object	Subject + Object + Verb
5	Subject + Verb + Object (Adjective, Adverb)	Subject + Object + Verb (Adjective, Adverb)
6	Subject + Predicate + (Adjective, Adverb)	Subject + Predicate + (Adjective , Adverb)
7	Subject + Verb + Pronoun	Subject + Verb
8	Verb + Object	Object + Verb

Syntax in sign language is synchronous with a parallel spatial and temporal arrangement but the syntax of sentences in an oral language (spoken language) is linear; as word follows each other. The syntax of Arabic Language sentences is partitioned into two categories of sentences. The first category includes the Subject, Verb, Predicate, and Object. Table 4 presents a syntax variation between the Arabic language and Arabic Sign Language. The second category sometimes starts with subject and sometimes with Verb [88]. Therefore, it is better to start with subject in any Arabic sign language sentence. The example of such sentence with its representation in the Arabic sign language is shown in Table 5.

Table 5: Syntax comparison between Arabic language and Arabic sign language

Criteria	Arabic Language Representation	Arabic Sign Language Representation
Sentence	قَامَ الطَّبِيْبُ بِإجرَاءِ عَمَلِيةٍ فِيْ القَلْبُ	الطَّبِيْبُ عَمَلِيةِ القَلْبُ
English Translation	The Doctor conducted a surgery in the heart	Doctor surgery heart
IPA Representation	qaa:ma a:ltˤabijbu biʔdʒraa:ʔi ʕamalijti: fij a:lqalb	/a:ltˤabijbu ʕamalijti: a:lqalb/

There are differences in the negative sentence ordering in Arabic sign language compared with the Arabic and spoken languages. Furthermore, the adjectives can be translated from the Arabic language to Arabic sign language in two ways. The first way is by using the adjective sign immediately and directly, whereas the second way is by using negation of an equivalent negative verb with the adjective. For example, the word /**حُبْ**/ /***ħub***/ which means ***love*** in the English language does not necessarily be the word /كُرْه/ /***kurh***/ which means ***hatred*** in the English language as shown in Table 6.

Table 6: Differences between negative syntax in Arabic language and Arabic sign language [92]

No	Arabic Language Syntax	Arabic Sign Language Syntax
1	Negative + Verb	Verb + Negative
2	Subject + Negative + Verb + Object	Object + Subject + Verb + Negative
3	Subject + Negative + Verb	Subject + Verb + Negative
4	Negative + (Adjective, Adverb)	(Adjective , Adverb) + Negative
5	Adjective	Verb + Negative

Based on the above illustration, it is clear that the Arabic sign language consists of many varieties among different Arab countries. The documentary history of ArSL is found in Egypt dictionary as far as in the year 1927 after investigation and was written by the association of impaired hearing and deaf in Egypt [93]. Consequently, other Arab countries made several attempts to standardize and spread the local Arabic sign language among their deaf community. As a result, standard sign languages has now be recognized in various Arabic nations such as Qatar (2010), Sudan (2009), Yemen (2009), Kuwait (2007), Tunisia (2007), Jordan (1993), Libya (1992), Morocco (1987), Iraq, UAE, Palestine, and Saudi Arabia [94].

There are gestural repertoire and cultural values similarities that exist within the Arabic nations that should be noted because it has created a belief that ArSLs are almost identical for all Arab countries. It is crucial to producing a standardized dictionary for Arabic sign language of all Arabic countries. Therefore, 18 Arab countries launched a standard dictionary called UASLD (i.e. the Unified Arabic Sign Language Dictionary) for impaired hearing people in 2007, which was an initiative in 1999 from the League of Arab States (LAS) and the Arab League Educational, Cultural and Scientific Organization (ALECSO). This UASLD is divided into two parts. The first part, which consists of over 1600 words, was published in Tunisia in 2001. The words was further divided into various categories such as food, family, home, etc., whereas the second part contained additional signs that are well organized and was published with a strong collaboration among ALECSO, LAS, the Arab Union for Deaf (AUD), and the Supreme Council for Family Affairs in Qatar funded in 2006 by the Arab deaf. The two video DVD version of the gesture was published in 2007 [94].

3.3 Taxonomy of Arabic Sign Language Recognition Approaches

The research efforts on ArSLR have witnessed a significant increase in the last few years. The researchers used various methods, frameworks, and techniques in this domain. Based on our literature investigation, the taxonomy of research efforts made on the Arabic sign language recognition approaches are divided into two main categories;

1) Vision-Based Recognition (VBR)
2) Sensor-Based Recognition (SBR)

The illustration is shown in Fig 5 and Table 7 provides a comparison of the vision-based and sensor-based approaches based on the design concepts.

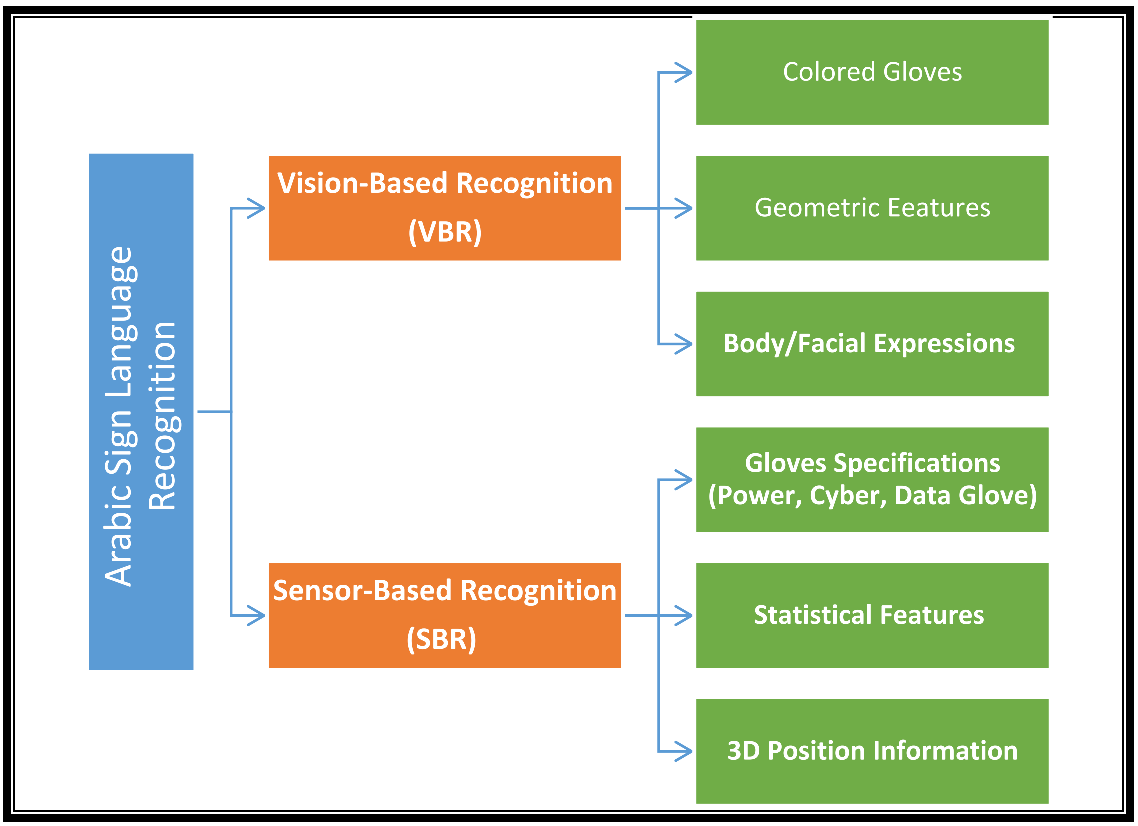

Fig. 5: Taxonomy of Arabic sign language recognition approaches

Table 7: Comparison of the vision-based and sensor-based approaches based on the design concepts [1]

Concept	Sensor-based approach	Vision-based approach
World wide availability	Low likely	High likely
Feature extraction	Relatively easier	Challenging
User experience	Inconvenient	Good
Cost	High	Low
User dependency	Less prone	Highly prone
Calibration	Required but stable	Environment Dependent

Vision-based recognition approach depends on providing a group of static and dynamic images (such as video). Generally, the signers are requested to purse shortly between the signs so that the produced signs are of good quality and to ease the signs segmentation process. The major benefit of vision-based approach in ArSL is that the users are not required to wear the ponderous DataGlove. On the other hand, there are various challenges posed by using the vision-based approach in ArSL such as images background, hand segmentation, face segmentation, and lighting conditions. In addition, the segmentation of lips, facial expressions, and hands gesture are computationally expensive. Nowadays, the techniques and algorithms have the ability to implement and perform segmentation in real time. However, vision-based recognition approach is still limited and needs more to be developed.

On the other hand, sensor-based recognition approach processes data obtained from Smart Gloves, which depends on sensors. The Power-Glove, Data-Glove, and Cyber-Glove have been generally used for ArSLR as shown in Fig. 6. Using the data obtained from the smart gloves, a huge number of features can be extracted such as finger bending, the orientation of the hand, movement, rotation, and position. In addition, the classification algorithm uses those features in detecting and recognizing the optimal sign.

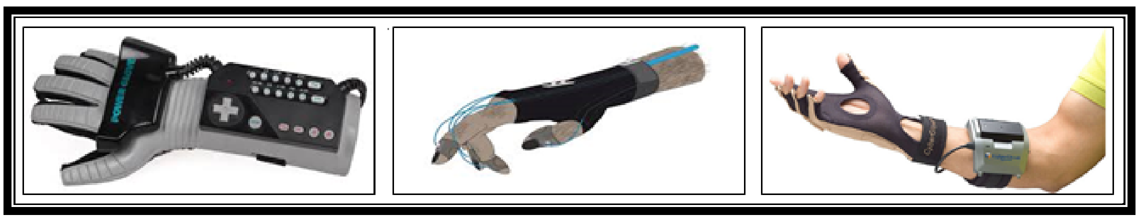

Fig. 6: Several types of smart gloves: Power Glove (left), DT-Data Glove (Center), Cyber Glove (right)

The secondary research taxonomy for ArSLR research efforts using vision-based and sensor-based approaches ranges from developing systems that can recognize small forms and segments such as alphabets, moving to larger yet small forms such as isolated words, and finally to the largest and most difficult forms such as full sentences. The ASLR task becomes more difficult when the unit to be recognized increases. For simplicity of the secondary research taxonomy, this work recommends the classification of Arabic ASLR research efforts into three broad categories, which include 1) alphabet sign language recognition, 2) isolated words sign language recognition, and 3) continuous sign language recognition. The following sections provide a literature investigation and comparison between various research efforts for each category.

3.3.1 Alphabet Sign Language Recognition

Based on our literature investigation, there various research attempts and efforts for Arabic alphabet sign language recognition by using vision-based recognition approach as shown in Table 8, which depends on performing each alphabet independently. Generally, alphabets are performed by a fixed position and the vocabularies database is restricted in size. Although there are 28 alphabets for the Arabic language, thus, ArSL utilizes 39 signs as previously mentioned in Fig. 4 [91]. Each Arabic alphabet is represented by one sign and the 11 remainders are used for other orthographical representation of the Arabic alphabets including **(أ/إ/ئ/ؤ/ء/ـئ/آ/ى/ة/لا/ال)**. In addition, the case of (**لا** and **ال**), represents 2 (**ل** and **ا**) Arabic alphabets but they have special cases when used in sign language recognition in which they are represented as one sign too.

To the best of our knowledge, we were unable to find any research publication for Arabic alphabet sign language recognition using sensor-based recognition approach.

3.3.2 Isolated Words Sign Language Recognition

Based on our literature investigation, there were various research attempts and efforts for Arabic isolated words sign language recognition using both vision-based recognition approach as shown in Table 9 and sensor-based recognition approach as shown in Table 10. The isolated words sign language recognition depends on the performance and analysis of each isolated word separately by analyzing a set of images to perform one sign only.

3.3.3 Continuous Sign Language Recognition

Based on our literature investigation, there were very limited research attempts and efforts made on Arabic continuous sign language recognition using vision-based recognition approach as shown in Table 11 and sensor-based recognition approach as shown in Table 12. We were able to find only two research attempts using sensor-based recognition approach, which were recently published in 2015. This shows that there is a research gap in attempts and efforts for Arabic continuous sign language recognition using both sensor and vision-based recognition approaches, which opens room for further research.

The continuous sign language recognition depends on performing and analyzing full sentences, which is more challenging than the two previous types (alphabet and isolated words). Continuous sign language recognition is considered more efficient and appropriate in real life. Continuous ASLR systems are more suitable for impaired hearing and Deaf Community. An ideal continuous ArSLR should be reliable and available in real-time with a high recognition rate. Detecting the additional movements from the transition between a set of signs, recognition, and modeling are considered as major challenges in continuous sign language recognition.

Table 8: Performance Comparison of Alphabet Sign Language Recognition for ArSL Research Efforts Using Vision-Based Approach

Source	**Year**	**Main Task**	**Techniques**		**Tools**	**Sign language Data**		**Signer Dependency**	**Recognition Rate (%)**
			Features Extraction	**Classifier used**		**Training**	**Testing**		
[95]	2014	Arabic Sign Language Recognition using the Leap Motion Controller (LMC) (28 alphabet)	Hus moments	Nave Bayes Classifier (NBC), Multilayer Perceptron (MLP)	MATLAB	2800	2800	N/A	98% with the Nave Bayes, 99% using the MLP
[96][97]	2013	Arabic alphabet sign language recognition (30 alphabet)	N/A	Pulse coupled neural network (PCNN)	N/A	60	240	Signer-Dependent	90%
[98]	2012	Arabic alphabet sign language recognition (28 alphabet)	Finding histograms	K-Nearest Neighbor (KNN)	N/A	N/A	N/A	N/A	White gloves: 80%, a black gloves: 65%, a red gloves: 75%, Naked hand: 50%.
[99]	2011	Arabic sign language Alphabets Recognition (converts signs into voice)	YCbCr space, PCA, Prewitt edge	K-Nearest Neighbor (KNN)	N/A	150	N/A	N/A	97%
[84]	2010	Arabic alphabet sign language recognition (30 alphabet)	Edge detection stage, feature vector creation stage	Nearest-Neighbor Technique (NNT)	N/A	N/A	N/A	N/A	91.3%
[100]	2008	Arabic alphabet sign language recognition using one hand (28 alphabet)	Color layers expanded	Fully recurrent network, Elman network algorithms	Matlab, C	900	900	2 Signer-Dependent	Recurrent Networks: 95.11%, Elman Network: 89.66%
[101]	2007	Arabic alphabet sign language recognition (30 gesture)	Boundary information and region information	Adaptive Neuro-Fuzzy Inference System (ANFIS)	N/A	1200	600	N/A	97.5%
[102]	2005	Arabic alphabet sign language recognition (30 alphabet)	Segmented color regions , geometric measures	Polynomial classifier	N/A	1625	698	Signer-Dependent	93.41%
[103]	2001	Arabic alphabet sign language recognition (28 alphabet)	Hand region	Adaptive Neuro-Fuzzy Inference System (ANFIS)	N/A	N/A	N/A	N/A	95.5%
[104]	2001	Arabic alphabet sign language recognition (30 gesture)	Border Information	Neuro-Fuzzy Inference System (ANFIS)	Matlab	1200	1800	N/A	93.55%

Table 9: Performance Comparison of Isolated Words Sign Language Recognition for ArSL Research Efforts Using Vision-Based Approach

Source	Year	Main Task	Techniques		Tools	Sign language Data		Signer Dependency	Recognition Rate (%)
			Features Extraction	Classifier used		Training	Testing		
[105]	2016	Arabic isolated word sign language recognition (20 isolated gesture)	Fisher linear discriminant analysis (LDA)	Fisher linear discriminant analysis (LDA)	Microsoft Kinect	N/A	N/A	N/A	N/A
[106]	2016	Arabic isolated word sign language recognition (40 isolated gesture)	Go-Stop Detector	Hidden Markov Model (HMM)	Microsoft Kinect	2400	1200	Signer-Dependent, Signer-Independent	95.125% for signer-Dependent, (92.5% for signer-independent
[107]	2014	Arabic isolated word sign language recognition (23 isolated words)	Appearance-based , Local Binary Patterns (LBP) and Principal Component Analysis (PCA)	Hidden Markov Model (HMM)	N/A	2415	1035	3 Signer-Dependent	99.97%
[108]	2013	Arabic isolated word sign language recognition (50 isolated words)	Fitness Function	Hybrid PCNN, graph matching approach	N/A	N/A	N/A	Signer-Independent	96% for Pose-Invariant constraints with indulgence of up to 90º.
[109]	2013	Arabic isolated word sign language recognition (40 isolated words)	Gesture Mixture Model (GMM)	Time delay neural network (TDNN)	Matlab	N/A	N/A	Signer-Dependent	77.43%
[110]	2012	Arabic isolated word sign language recognition (300 isolated words)	Gaussian skin color, simple region growing, Two colored gloves	Hidden Markov model (HMM)	Matlab	1000	500	Signer-Independent	50 signs was 98%, 300 signs was 95%
[111]	2012	Arabic isolated word sign language recognition (180 isolated words)	Fourier descriptors (FDs)	K Nearest Neighbor (KNN)	Visual c#, OpenCV, LifeCam VX-5500	N/A	N/A	N/A	90.55 %.
[10]	2011	Arabic isolated word sign language recognition (23 isolated gesture)	Zonal coding, Discrete cosine transform (DCT)	K Nearest Neighbor (KNN)	N/A	2300	1150	3 Signer-Independent	87%
[112]	2011	Arabic isolated word sign language recognition (20 isolated gesture)	N/A	Hidden Markov model (HMM)	N/A	900	810	Signer-Independent	82.22%

[113]	2010	Arabic isolated word sign language recognition (40 isolated gesture)	spatial domain analysis	Hidden Markov model (HMM)	N/A	1920	1280	Signer-Independent, Signer-Dependent	70.5% user independent mode, 92.5% user-dependent mode.
[114]	2009	Arabic isolated word sign language recognition (30 isolated gesture)	discrete cosine transform (DCT)	Hidden Markov models (HMMs)	camera	2730	1315	Signer-Independent, Signer-Dependent	94.2% user independent offline mode, 90.6% user independent online mode. 97.4% user dependent offline mode, 93.8% user dependent online mode.
[115]	2007	Arabic isolated word sign language recognition (23 isolated gesture)	Zonal coding, Discrete cosine transform (DCT)	Hidden Markov model (HMM)	N/A	805	345	Signer-Independent	95%

Table 10: Performance Comparison of Isolated Words Sign Language Recognition for ArSL Research Efforts Using Sensor-Based Approach

Source	Year	Main Task	Techniques		Tools	Sign language Data		Signer Dependency	Recognition Rate (%)
			Features Extraction	Classifier used		Training	Testing		
[116]	2013	Arabic isolated word sign language recognition (20 isolated gesture)	Linear Discriminant Analysis (LDA)	Minimum Distant (MD), Dempster-Shafer theory of evidence (DS)	Two CyberGlove (GLV), two Flock-of-Birds (FOB)	1000	1000	N/A	Hand tracking: 84.7%, Cyber Glove: 91.3%. Feature based fusion: 96.2%. Fusion at the decision level: 98.1%
[117]	2013	Arabic isolated word sign language recognition (100 isolated gesture)	Principal component analysis (PCA)	Support Vector Machines (SVMs)	Two CyberGlove	1500	500	2 Signer-Independent	99.6%
[58]	2012	Arabic isolated word sign language recognition (10 isolated gesture)	accumulated differences (ADs), a regression technique	K-Nearest-Neighbor KNN classifier with Manhattan distance measure.	DG5-VHand data glove	500	500	Signer-Independent, Signer-Dependent	92.5% in the signer independent mode, 95.3% in the signer dependent.
[118]	2012	Developed Intelligent Computer-Based System for learning ArSL (32 dynamic gestures, 33 static gestures)	N/A	N/A	DataGlove, Bend sensors, push button	N/A	N/A	4 Signer-Independent	93%
[119]	2004	Arabic isolated word sign language recognition (42 isolated gesture)	Statistical Features	Support Vector Machines (SVMs)	PowerGlove	720	120	N/A	90%

Table 11: Performance Comparison of Continuous Sign Language Recognition for ArSL Research Efforts Using Vision-Based Approach

Source	Year	Main Task	Techniques		Tools	Sign language Data		Signer Dependency	Recognition Rate (%)
			Features Extraction	Classifier used		Training	Testing		
[120]	2013	Arabic continuous sign language recognition (3 to 4 words) (30 sentences)	Multi-Layer Perceptron (MLP)	Pulse coupled neural network (PCNN), Graph Matching.	N/A	N/A	N/A	N/A	70%
[121]	2010	Arabic continuous sign language recognition (40 sentences)	Discrete cosine transform (DCT), Spatio-temporal	Hidden Markov model (HMM)	Georgia Tech Gesture Recognition Toolkit (GT2K), Hidden Markov Model Toolkit (HTK).	532	228	Signer-Dependent	73.3%
[122]	2008	Arabic continuous sign language recognition (40 sentences)	Discrete Cosine Transform (DCT),	Hidden Markov model (HMM)	Georgia Tech Gesture Recognition toolbox (GT2K), HTK Toolkits, digital video camera.	N/A	N/A	Signer-Dependent	75%

Table 12: Performance Comparison of Continuous Sign Language Recognition for ArSL Research Efforts Using Sensor-Based Approach

Source	Year	Main Task	Techniques		Tools	Sign language Data		Signer Dependency	Recognition Rate (%)
			Features Extraction	Classifier used		Training	Testing		
[123]	2015	Arabic continuous sign language recognition using Sensor-Based (40 sentences)	Window-based approach	Modified the polynomial classifier to suitable with sequential data	Two DG5-VHand data gloves	280	120	N/A	85%
[13]	2015	Arabic continuous sign language recognition using Sensor-Based (40 sentences)	Window-based approach	Modified k-Nearest Neighbor (MKNN)	Two DG5-VHand data gloves	N/A	N/A	Signer-Dependent	98.9%

4.0 TAXONOMY OF OPEN ISSUES AND CHALLENGES FOR ARABIC SIGN LANGUAGE RECOGNITION

A taxonomy of this of this review shown in Fig. 7 is developed for open issues and challenges inherent in ArSLR research. The issues and the challenges according to the taxonomy are classified based on Language, System, Environment, and Gesture which are further described in the subsection below.

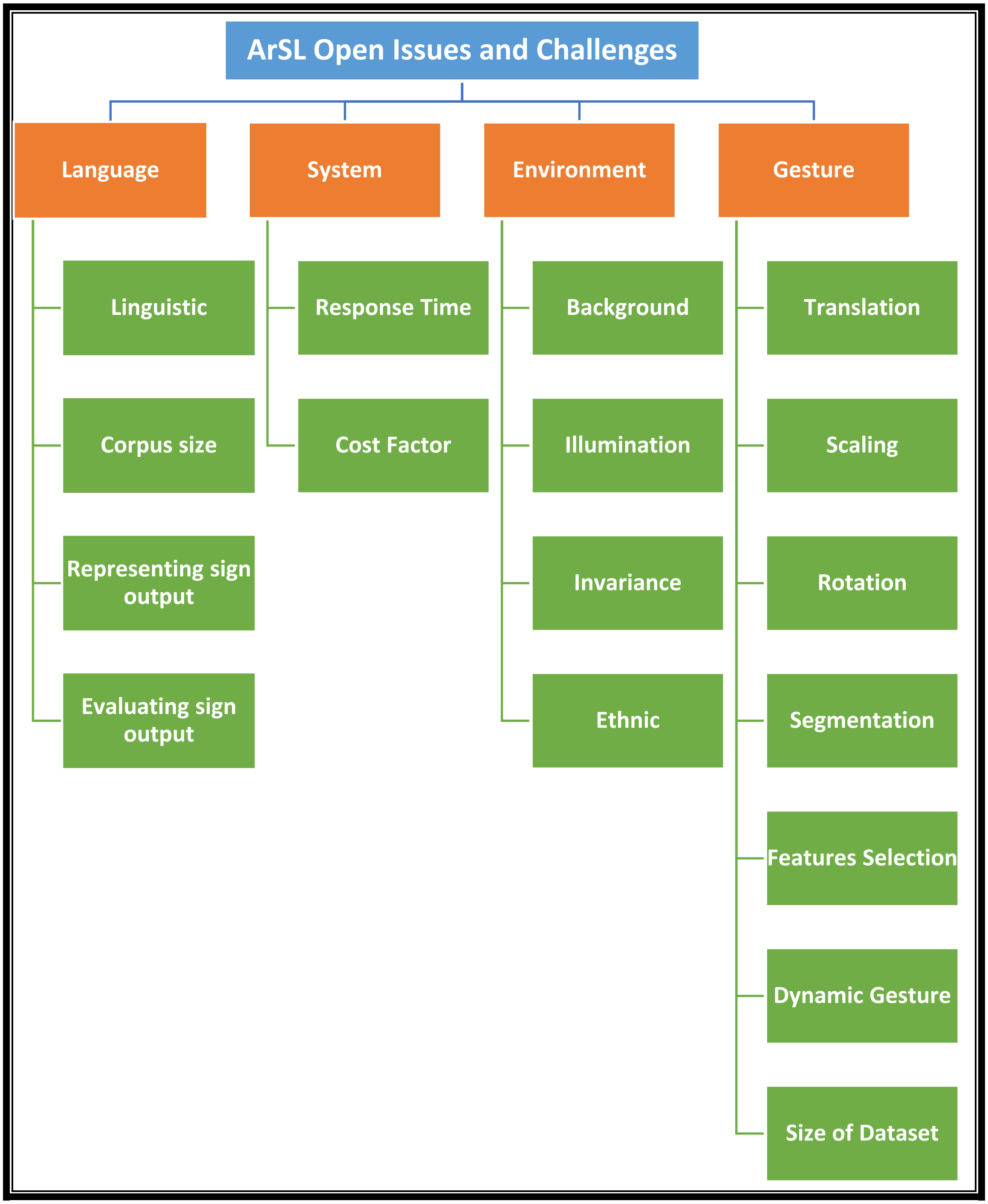

Fig. 7: Taxonomy of ArSLR Open Issues and Challenges

A. Language Perspective:

1. Linguistic:

Linguistic studies that tackle ArSL are scarce, especially those related to grammar and structure. Consequently, different misunderstanding of natural language might occur and the concerned researchers could find it difficult to design and develop usable ArSL translation systems [89]. One of the misunderstandings assumes that SL is a universal language that can be communicated by deaf individuals anywhere worldwide but, there are different forms of SLs in the world such as British SL, Irish SL, ArSL, and many more. Secondly, some people assume that ArSL is dependent on Arabic language but in reality, it does not but has its special grammar, structure, and idioms [84].

2. Corpus size:

The size of the translation corpus is another important challenge inherent in ArSL research. This is because the linguistic studies of ArSL's grammar and structure are scarce. Normally, the data-driven approach depends heavily on the corpus used and its size, which consequently affects the translation accuracy. Furthermore, a written document does not exist for ArSL, which indicates that ArSL documents are not inexistent for designing and developing a translation corpus. Instead, it is essential that the translation corpus is in the visual form (albeit with annotation). Therefore, the translation corpus for ArSL has to be designed and developed from scratch and limited in size in order to accurately translate the signed sentences [89].

3. Representing sign output:

Representation of the output sign sentences is another important challenge. Spoken languages rely on sounds to produce utterances, whereas 3D space is employed in SL for representing signs. Signs are continuous that can be gathered using avatar or video clips' concatenation [88] [89].

4. Evaluating sign output:

Finding a way for evaluating SL output is another essential challenge because SL normally uses multi-channel representations. Translation systems using machine learning can also be affected [88] [89].

B. System Perspective:

1. Response Time:

The required time for gesture algorithm's execution and computer response should be as fast as possible and should be generally acceptable in a specific real-time application for their widespread [124]. Therefore, response time should be investigated more by the researchers on the field of ArSLR, in order to solve the problem of execution.

2. Cost Factor:

Costly specialized equipment and devices including camera, sensor, data glove, and many more are required for ArSLR research and development and need sufficient sponsorship and budget [124]. Hence, the cost factor is another essential challenge for ArSLR study which is open for further investigation in the field as a result of the high cost on the specialized equipment and the scarcity of the sponsorship and budget. However, in such circumstances, a collaboration between various academic institutions, industrial companies, and industrial research and development institutions is advisable. Besides, discussing these challenges with sponsoring bodies and parties for possible funding opportunities is required.

C. Environment Perspective:

1. Background Challenge:

Environmental conditions such as background and lightning may have positive or negative impacts on the extracted features during ArSLR design and development, especially when using simple and low-quality equipment such as camera and data gloves [124]. Therefore, the background feature is another important open issue and challenge that must be tackled especially in vision-based gesture recognition.

2. Illumination Challenge:

Light conditions may change due to factors related to normal or artificial outdoor or indoor, which consequently result in illumination. Therefore, illumination regarded as a problem that requires more investigation for ArSLR, and it is crucial to resolving it by adopting a segmentation approach or suitable color space models like (Value (HSV), Saturation, Hue, and (Y) Green, (Cb) Blue, (Cr) Red which occur in pre-processing segment and is referred to as (YCbCr) video color space. The combination of both the segmentation approach and color space conversion can also be used for resolving the illumination challenge [23] [125].

3. Invariance Challenge:

Another problem inherent in ArSLR research which is open for further investigation is invariance. This problem might affect features especially those resultants from the feature selection stage. Invariance is normally handled in rotation, scaling, translation, and illumination [126].

4. Ethnic Challenge:

It is important to adhere to variations in ethnic groups such as skin colors, gender, origin, and many more in order to conduct ArSLR research and development. Thus, ArSLR research can only be reliable once it considers the ethnic variations between its users [124].

D. Gesture Perspective:

1. Translation Challenge:

This refers to a state of positioning and location of hand-object within an image which might be changed through adjusting of hand location or camera in the capturing process [124]. Therefore, translation remains a problem in an ArSLR research and development from gesture point of view that needs more investigation.

2. Scaling Challenge:

The distance between the hand and the camera lens is referred to as scaling. The measurement of the scaling can be very high when the position of the hand is not close to the camera lens. This high distance may have a negative impact on the overall recognition in the ArSLR task. Therefore, researchers use various techniques such as histogram and scale matching in order to overcome this challenge [127].

3. Rotation Challenge:

The accuracy of ArSLR systems can be influenced by the rotation of a gesture or a hand object. This challenge can be addressed and resolved by applying soft computing techniques [127].

4. Segmentation Challenge:

Hand segmentation is considered an essential part of ASLR task that is needed to be performed before useful features can be extracted from the gesture image. In addition, it is very important in ArSLR research and development as it makes the segmentation from the image sequence very fast and accurate [128]. There are various factors that may affect the hand segment such as the complex background, varied hand size, and unforeseen physical factors. In addition, other processes that can directly affect the hand segment are feature extraction, gesture tracking and the classification of the sign language. These processes have a consequent influence on the overall performance of the ArSLR r. thus, the recognition of the foreground region from its background is very important [126]. Our investigation revealed that hand detection can be carried out by using either geometric features (such as the orientation of the finger, fingertips, and contour [129] or non-geometric (like motion, strip, and color) [130][131].

5. Features Selection Challenge:

In SLR, feature selection is a prior process to classification. In this process, the most important sets of features that will contribute to an effective classification result are selected and minimized. Features selection acts as a crucial open issue and challenge for ArSLR research and development [126].

6. Dynamic Gesture Challenge:

The dynamic gesture has served as a bedrock of the research work on ArSL. It can also be referred to as a gesture in motion. However, such research efforts are limited in terms of applicability. In ArSLR research and development, dynamic gestures have open issues and challenges such as the identification of the motion, isolation of sign recognition and the sequence of the gesture identification from the start to the end of recognition [126].

7. Size of Dataset Challenge:

The size of the dataset is another challenge faced with Arabic sign language recognition. The current research works rely on datasets that are limited and small in terms of size in which the alphabets of languages that are static one-handed signs are used and 8 to 10 phrases that represent the target are used for continuous ArSLR. However, such research work will suffer usability problem provided that huge amount of dataset is utilized due to the limited size of the dataset [126]. In addition, the preference on the type of the application to be chosen for obtaining gesture either in continuous or in an isolated form relies so much on the size of the dataset. More importantly, the dataset size irrespective of the kind of application plays a vital role in the realization of an effective result of the experiment [18]. Thus, further research on ArSLR should focus more on the application that utilizes the large datasets and accurate data sets and designing of adequate databases.

5.0 FUTURE RESEARCH TRENDS AND DIRECTIONS FOR ARABIC SIGN LANGUAGE RECOGNITION

Based on our literature investigation and analysis of more than 100 articles, various research trends, research directions, and potential research topics are drawn for ArSLR research and development as summarized in Fig. 8.

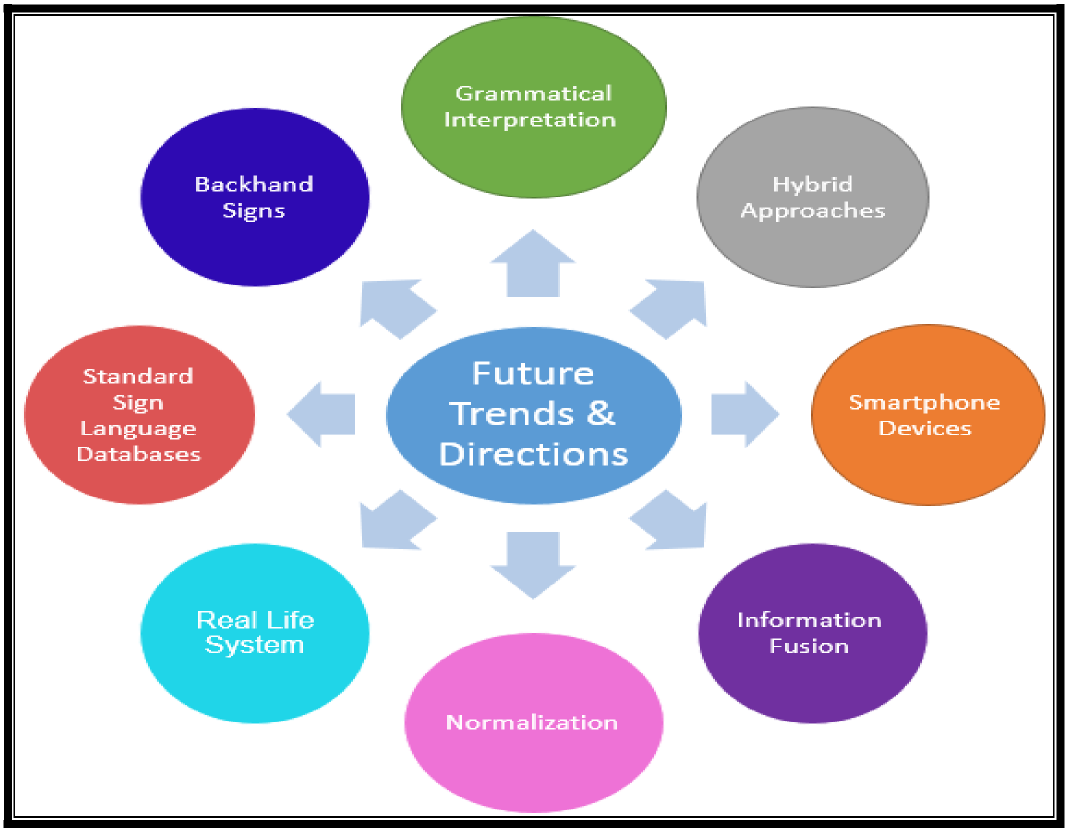

Fig. 8: Potential Trends and Directions for ArSLR Research and Development

Firstly, grammatical rules in SLs and their corresponding spoken languages are not sufficiently studied and thus, considered limited. It is important to investigate and include grammatical rules of ArSL into its spoken form and have more related research initiatives in this aspect, which consequently can be useful for future ArSLR research and development.

Secondly, developing hybrid approaches is a common trend in many fields recently. This trend should be investigated for ArSLR research and development. Data fusion is the most important aspect of this approach. This can be achieved by applying VBR and SBR techniques to obtain data from the camera and smart glove and merging it together. However, the fused data acquired can be in three levels, such as data level, feature level, and decision level. The effectiveness of this fusion produces a better accuracy of ArSLR systems in performance in comparison with ArSLR systems that employ VBR or SBR approaches individually. Thus, the implementation of the hybrid technique by fusing both techniques (SBR and VBR) acts as a potential future direction for ArSLR research and development.

Thirdly, embedding ArSLR systems into smartphone devices is predicted due to their wide use and ease of reach. This indicates that more efforts should be done to stabilize the performance of other related technologies embedded on smartphone devices such as Automatic Speech Recognition (ASR) systems, where human speech and text can be recognized and translated into gestures as performed on smartphone devices by human or avatar signers. Contributions to this approach can enhance communication in both directions through SL.

Fourthly, satisfactory results are achieved in ArSLR research using simple environmental setup and scenarios. Thus, better performance and result accuracy still remains a challenge in producing stable ArSLR systems. Therefore, moving towards hybrid approaches is advisable, which involve the integration of various techniques that use data obtained from a different sources like a smart glove, sensors, digital cameras, and Kinect. This acts as another future direction in ArSLR research and development that worth investigation.

Fifthly, the normalization between the ways different singers perform the same gesture is not well studied. Each signer performs a particular gesture differently due to serious differences in terms of style, body size, and timing. Hence, the signer differences can be tackled using enhanced or new techniques and can be another future research direction for ArSLR research and development.

Sixthly, In order to bridge the communication gap and making it easier between both the deaf and hearing citizens, the current state-of-the-art technology (like a system that can translate text or speech) must be employed in the development and integration of real life ArSLR systems. These systems are indispensable and highly demanded, which act as important research and industrial future direction.

Seventhly, designing, preparing and creating standard sign language databases for ArSL is indeed amongst the most important research directions for ArSLR research. The ArSLR research community is mainly in need of sign language databases for ArSL for continuous sign language recognition tasks due to the limited number of publically available databases. Advancements in large vocabulary continuous sign language databases will certainly improve ArSLR systems in terms of accuracy and vocabulary coverage.

Finally, the previous sections reviewed studies and research works pertaining to sign language recognition systems for ArSL using forehand images and continuous gestures. However, it is found that there is no study or research that uses backhand images and continuous gestures to recognize ArSL signs. Therefore, we believe that this aspect is not fulfilled and can be investigated further in the future.

Based on this discussion, Fig. 9 provides seminal works and a roadmap of the research initiatives for ArSLR since 2001 and the potential future research trends and directions until 2020.

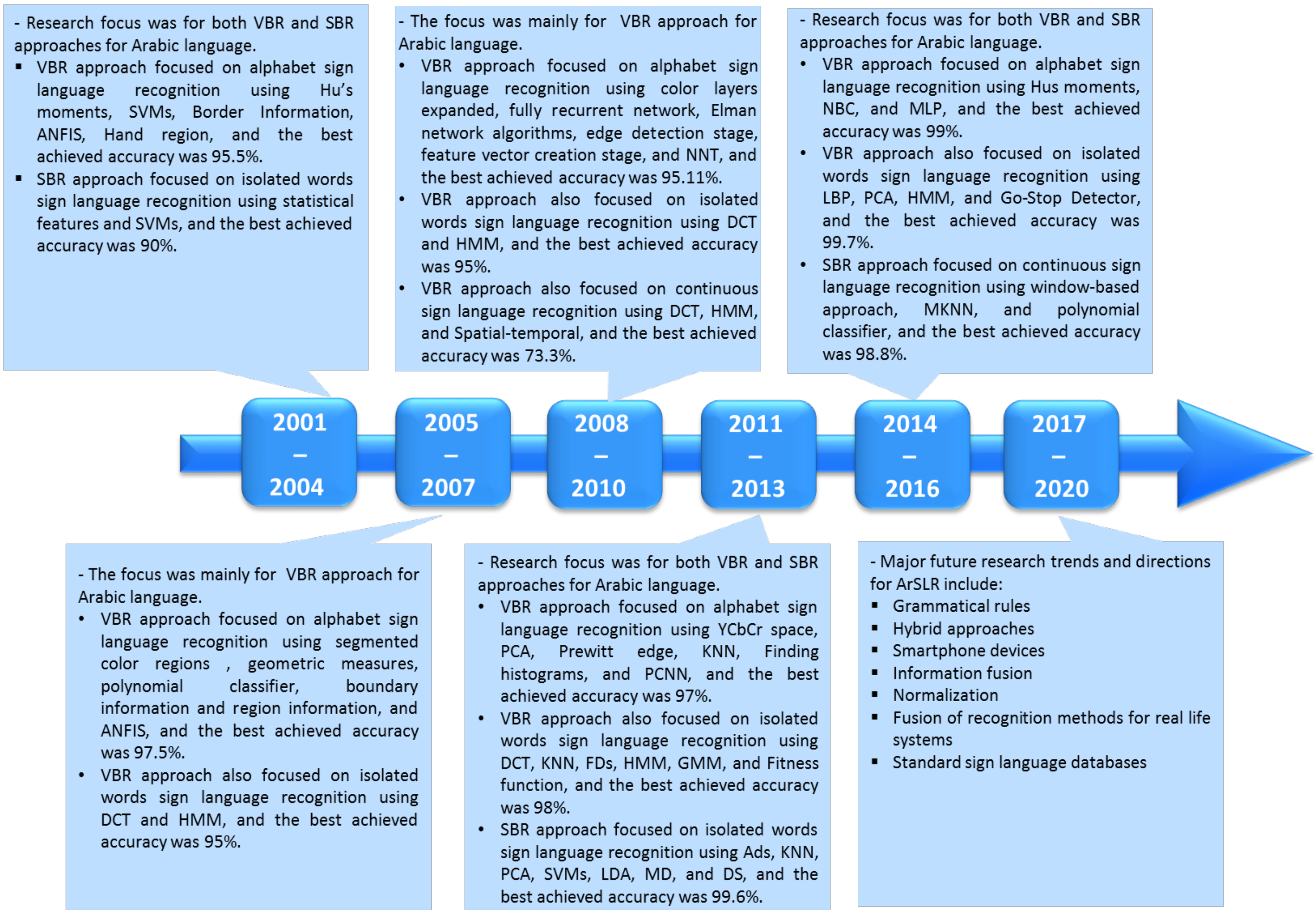

Fig. 9: Seminal works and roadmap of the research initiatives for ArSLR since 2001 and the future research trends and directions until 2020

6.0 SUMMARY AND CONCLUSION

Although research attempts and efforts in ASLR started several years ago, yet, it is still limited in many aspects. This area can still be perceived in its infancy. Based on our knowledge, no comprehensive systems have been developed which cover the present extensive scope. Research in this domain will certainly influence other domains that encompass human-computer interaction (HCI) elements. In this paper, we summarized two main approaches that are used in translating SL into text, namely: 1) Vision-Based Recognition (VBR) approach, and 2) Sensor-Based Recognition (SBR) approach. These approaches were discussed in this paper with special emphasis on ArSL. The main target of any approach in communicating between normal people and deaf is to simplify the communication in a real-life environment without any constraint and limitation such as wearing of colored gloves or cumbersome devices. To ease and reduce these constraints, the researchers put huge efforts to develop these approaches [132] [133]. For example, Microsoft Kinect is used as an interface for ASLR. However, it is not used widely for Arabic sign language recognition.

Based on our literature investigation, Arabic alphabet sign language recognition using vision-based recognition approach has achieved a recognition rate of 99% [52] as shown in Fig. 10, which indicates that research attempts were able to achieve excellent performance in this category. It is important to note that we are unable to find any previous study on sign language recognition for the Arabic alphabet that used sensor-based recognition approach.

In addition, based on our literature investigation, Arabic isolated words sign language recognition using vision-based recognition approach has achieved a recognition rate of 99.97% [107]. On the other hand, Arabic isolated words sign language recognition using sensor-based recognition approach has achieved a recognition rate of 99.60% [117] as shown in Fig. 10. This indicates that research attempts using both vision-based and sensor-based approaches were able to achieve excellent performance in this category.

Based on our literature investigation for the third category namely the continuous sign language recognition, research works using vision-based recognition approach has achieved a recognition rate of 75% [122], whereas research attempts using sensor-based recognition approach has achieved a recognition rate of 98.90% for signer dependent testing [13] as shown in Fig. 10. This shows that this category requires more research attempts using both vision-based and sensor-based approaches, especially for signer independent applications in order to achieve better performance. Research attempts in this category are still inadequate and the performance is not yet satisfactory compared to the Arabic alphabet and isolated words sign language recognition categories.

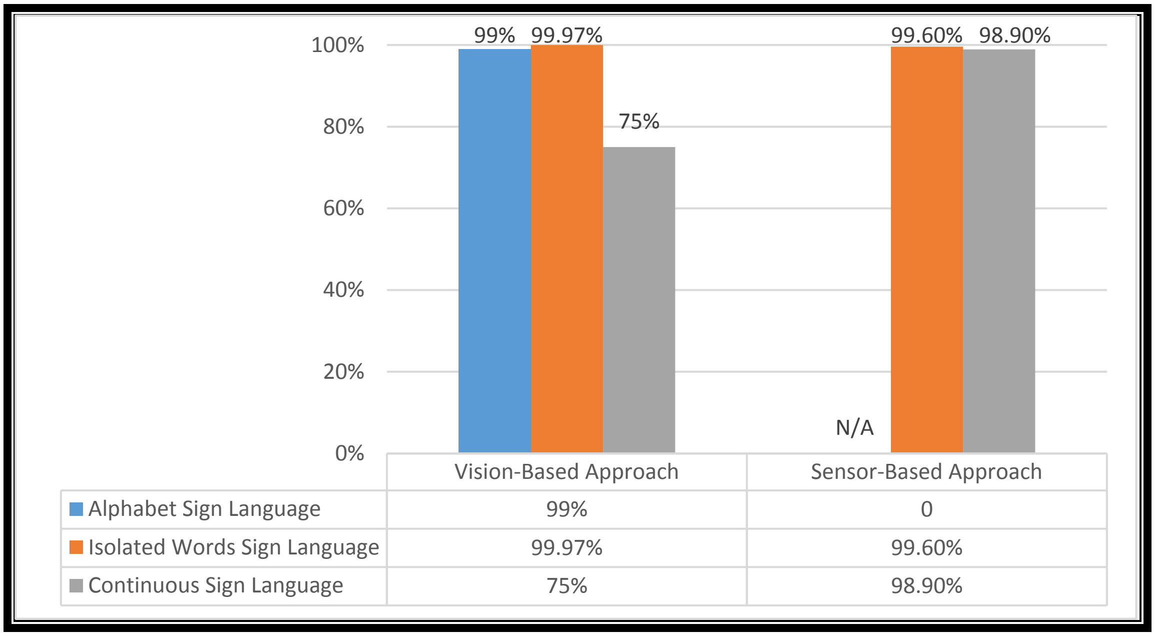

Fig. 10 Summary of Performance Accuracy

Based on our review, it is found that the SBR approach can be a better option compared to VBR approach for ArSLR tasks. Research efforts on SBR approaches have concentrated on two major directions:

1) Selection of the suitable glove: In this research direction, there are three major problem is that needed to be resolved for ArSLR system.

   - The location and the number of sensors that has a direct influence on the length of the resultant dictionary.
   - No detailed analysis of the various aspects of the system, which include the electronics, sensors, and support.
   - Measurement calibration, owing to the fact that people have hands and finger that consists of different sizes and thickness

Consequently, the sensors of the glove may not be in line with the locations of the finger joints. However, there is a need for calibration of gloves for a specific user in order to reduce inaccuracy of the result. This brings us to the notice about the question of whether the SBR approach is more or less appropriate for the signers-related settings.

2) The robust signal processing: This research direction focuses on the stage of signal processing in ArSLR.

Traditionally, the sign language recognition techniques have been viewed as an ideal pattern recognition techniques. These techniques consist of three main processes, namely, preprocessing, feature extraction, and feature classification. One of the main challenges is the extraction of features from noisy data taking into consideration that patterns are represented with dynamic gestures that produces non-stationary signals. On the other hand, the classification process can also be a substantial challenge when considering vocabularies that comprise one-hand and both hands signs.

For the functional deployment of the translation techniques of the SL, previous studies found that the SBR systems are less attractive to users than the VBR systems. The SBR approach demands that the signers wear data gloves that are tied with specialized Digital Signal Processing (DSP) boards, while the VBR approach does not. In addition, the VBR approach for ArSL can take advantage of the additional information, which is obtained from facial expressions, and lip and/or head motions. Thus, this acquired VBR information is not presently employed in the ArSL recognition systems globally. Furthermore, the VBR approach still demands special settings that include the lighting, the background, and signer's camera(s) in order to acquire the signs, which have substantial effects on the overall efficiency and performance of the VBR system.

RESEARCH TEAM

A team of four researchers co-authored this article and developed this research. Ahmad Sami Al-Shamayleh, Rodina Ahmad, and Nazeean Jumhari are affiliated with the University of Malaya, Malaysia, whereas Mohammad A. M. Abushariah is affiliated with The University of Jordan, Jordan. Ahmad Sami Al-Shamayleh and Rodina Ahmad are the corresponding authors for this research article.

The Authors focus on the design and evaluation of interactive applications that help people with special needs, including autism, deaf and syndrome down to gain spiritual knowledge. In addition, the authors have research experience in different topics including pattern recognition, speech recognition and gesture recognition. With reference to ASLR, the authors created the first corpus for Arabic sign language based on systematic criteria and strict rules. This corpus has all possible types of signs (alphabet and numbers, isolated words and continuous sentences), which can be used for ArSLR research and development. In addition, the research team has many research publications related to the Arabic language patterns and assistive technology, and recently conducted a research for Arabic sign language recognition based on backhand and independent user. This research is among the earliest initiatives for Arabic sign language using backhand sign recognition. Recently, the research team published a journal article that provides a systematic literature review (SLR) on vision based hand gesture recognition, which is the first SLR within its scope. At present, we are working on proposing a novel multimodal framework for Arabic Sign Language Recognition using sensor devices, which will be the initial article on Arabic sign language using this technology that intend to overcome many of the present challenges in this area.

REFERENCES

[1] A. S. Al-Shamayleh, R. Ahmad, M. A. M. Abushariah, K. A. Alam, and N. Jomhari, “A systematic literature review on vision based gesture recognition techniques,” *Multimed. Tools Appl.*, vol. 77, no. 21, pp. 1–64, Apr. 2018.

[2] M. K. Bhuyan, D. Ghoah, and P. K. Bora, “A Framework for Hand Gesture Recognition with Applications to Sign Language,” in *2006 Annual IEEE India Conference*, 2006, pp. 1–6.

[3] A. Kuznetsova, L. Leal-Taixe, and B. Rosenhahn, “Real-Time Sign Language Recognition Using a Consumer Depth Camera,” in *2013 IEEE International Conference on Computer Vision Workshops*, 2013, pp. 83–90.

[4] J. Molina, M. Escudero-Viñolo, A. Signoriello, M. Pardàs, C. Ferrán, J. Bescós, F. Marqués, and J. M. Martínez, “Real-time user independent hand gesture recognition from time-of-flight camera video using static and dynamic models,” *Mach. Vis. Appl.*, vol. 24, no. 1, pp. 187–204, 2013.

[5] Z. Lu, X. Chen, Q. Li, X. Zhang, and P. Zhou, “A hand gesture recognition framework and wearable gesture-based interaction prototype for mobile devices,” *IEEE Trans. human-machine Syst.*, vol. 44, no. 2, pp. 293–299, 2014.

[6] S. Kang, K. Chung, K.-W. Rim, and J.-H. Lee, “Development of Real-Time Gesture Recognition System Using Visual Interaction,” in *Proceedings of the International Conference on IT Convergence and Security 2011*, 2012, pp. 295–306.

[7] M. Abhishek and H. Divakar, “A Video Based Sign Language Recognition System,” *Int. J. Eng. Sci. Comput.*, vol. 7, no. 9, pp. 1–3, 2017.

[8] M. Raees and S. Ullah, “Continuous Number Signs Recognition,” in *2014 12th International Conference on Frontiers of Information Technology*, 2014, pp. 274–279.

[9] M. Avraam, “Static gesture recognition combining graph and appearance features,” *Int. J. Adv. Res. Artif. Intell.*, vol. 3, no. 2, pp. 1–4, 2014.

[10] T. Shanableh and K. Assaleh, “User-independent recognition of Arabic sign language for facilitating communication with the deaf community,” *Digit. Signal Process. A Rev. J.*, vol. 21, no. 4, pp. 535–542, 2011.

[11] S. M. Halawani, D. Daman, S. Kari, and A. R. Ahmad, “An Avatar Based Translation System from Arabic Speech to Arabic Sign Language for Deaf People,” *IJCSNS Int. J. Comput. Sci. Netw. Secur.*, vol. 13, no. 12, pp. 43–52, 2013.

[12] S. Wei, X. Chen, X. Yang, S. Cao, and X. Zhang, “A component-based vocabulary-extensible sign language gesture recognition framework,” *Sensors*, vol. 16, no. 4, pp. 1–16, 2016.

[13] N. Tubaiz, T. Shanableh, and K. Assaleh, “Glove-Based Continuous Arabic Sign Language Recognition in User-Dependent Mode,” *IEEE Trans. Human-Machine Syst.*, vol. 45, no. 4, pp. 526–533, Aug. 2015.

[14] R. S. Choraś, “Hand shape and hand gesture recognition,” in *IEEE symposium on industrial electronics and applications (ISIEA 2009)*, 2009, pp. 145–149.

[15] C. Tran and M. M. Trivedi, “3-D Posture and Gesture Recognition for Interactivity in Smart Spaces,” *IEEE Trans. Ind. Informatics*, vol. 8, no. 1, pp. 178–187, 2012.

[16] N. Tubaiz, T. Shanableh, K. Assaleh, and S. Member, “Glove-Based Continuous Arabic Sign Language Recognition in User-Dependent Mode,” vol. 45, no. 4, pp. 1–8, 2015.

[17] T. Chouhan, A. Panse, A. K. Voona, and S. M. Sameer, “Smart glove with gesture recognition ability for the hearing and speech impaired,” in *2014 IEEE Global Humanitarian Technology Conference - South Asia Satellite (GHTC-SAS)*, 2014, pp. 105–110.

[18] S. Bilal, R. Akmeliawati, A. A. Shafie, and M. J. E. Salami, “Hidden Markov model for human to computer interaction: a study on human hand gesture recognition,” *Artif. Intell. Rev.*, vol. 40, no. 4, pp. 495–516, 2011.

[19] J. Cheng, C. Xie, W. Bian, and D. Tao, “Feature fusion for 3D hand gesture recognition by learning a shared hidden space,” *Pattern Recognit. Lett.*, vol. 33, no. 4, pp. 476–484, 2012.

[20] M. Vafadar and A. Behrad, “A vision based system for communicating in virtual reality environments by recognizing human hand gestures,” *Multimed. Tools Appl.*, vol. 74, no. 18, pp. 7515–7535, 2015.

[21] S. S. Rautaray, A. Agrawal, S. S. Rautaray, and A. Agrawal, “Vision based hand gesture recognition for human computer interaction: a survey,” *Artif. Intell. Rev.*, vol. 43, pp. 1–54, 2012.

[22] S. Kausar and M. Y. Javed, “A Survey on Sign Language Recognition,” in *2011 Frontiers of Information Technology*, 2011, pp. 95–98.

[23] P. K. Pisharady and M. Saerbeck, “Recent methods and databases in vision-based hand gesture recognition: A review,” *Comput. Vis. Image Underst.*, vol. 141, pp. 152–165, 2015.

[24] T. Aujeszky and M. Eid, “A gesture recogintion architecture for Arabic sign language communication system,” *Multimed. Tools Appl.*, vol. 75, no. 14, pp. 8493–8511, 2015.

[25] A. RaySarkar, G. Sanyal, and S. Majumder, “Hand Gesture Recognition Systems: A Survey,” *Int. J. Comput. Appl.*, vol. 71, no. 15, pp. 25–37, 2013.

[26] S. Pranjali and U. V.S, “Hand Gesture Recognition Systems : A Survey,” *Int. J. Inven. Eng. Sci.*, vol. 3, no. 3, pp. 1–5, 2015.

[27] H. Cooper, B. Holt, and R. Bowden, “Sign Language Recognition,” in *Hilton A., Krüger V., Sigal L. (eds) Visual Analysis of Humans*, Springer, London, 2011, pp. 539–562.

[28] M. J. Cheok, Z. Omar, and M. H. Jaward, “A review of hand gesture and sign language recognition techniques,” *Int. J. Mach. Learn. Cybern.*, vol. 10, no. 1, pp. 131–153, 2019.

[29] M. M. Hasan and P. K. Mishra, “Hand Gesture Modeling and Recognition using Geometric Features : A Review,” *Can. J. Image Process. Comput. Vis.*, vol. 3, no. 1, pp. 12–26, 2012.

[30] D. Chen, G. Li, Y. Sun, J. Kong, G. Jiang, H. Tang, Z. Ju, H. Yu, and H. Liu, “An interactive image segmentation method in hand gesture recognition,” *Sensors*, vol. 17, no. 2, p. 253, 2017.

[31] S. Ruffieux, D. Lalanne, and E. Mugellini, “ChAirGest: a challenge for multimodal mid-air gesture recognition for close HCI,” in *Proceedings of the 15th ACM on International conference on multimodal interaction*, 2013, pp. 483–488.

[32] Y. Ying Wu and T. S. Huang, “Hand modeling, analysis and recognition,” *IEEE Signal Process. Mag.*, vol. 18, no. 3, pp. 51–60, May 2001.

[33] A. G. Bairagi and Y. . Kapse, “Survey on Sign language to Speech Conversion,” *Int. J. Innov. Res. Comput. Commun. Eng.*, vol. 6, no. 1, pp. 267–274, 2018.

[34] R. Zaman Khan, “Hand Gesture Recognition: A Literature Review,” *Int. J. Artif. Intell. Appl.*, vol. 3, no. 4, pp. 161–174, 2012.

[35] G. Murthy and R. Jadon, “A review of vision based hand gestures recognition,” *Int. J. Inf. Technol. Knowl. Manag.*, vol. 2, no. 2, pp. 405–410, 2009.

[36] H. Bay, T. Tuytelaars, and L. Van Gool, “Surf: Speeded up robust features,” in *European conference on computer vision*, 2006, pp. 404–417.

[37] J. Rekha, J. Bhattacharya, and S. Majumder, “Hand gesture recognition for sign language: A new hybrid approach,” in *The 2011 International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV)*, 2011, pp. 80–86.

[38] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[39] G. Kumar and P. K. Bhatia, “A detailed review of feature extraction in image processing systems,” in *Advanced Computing & Communication Technologies (ACCT), 2014 Fourth International Conference on*, 2014, pp. 5–12.

[40] K. Delac, M. Grgic, and S. Grgic, “Independent comparative study of PCA, ICA, and LDA on the FERET data set,” *Int. J. Imaging Syst. Technol.*, vol. 15, no. 5, pp. 252–260, 2005.

[41] M. Suriya, N. Sathyapriya, M. Srinithi, and V. Yesodha, “Survey on real time sign language recognition system: an LDA approach,” in *International conference on exploration and innovations in engineering and technology, ICEIET*, 2016, pp. 219–225.

[42] J. Lafferty, A. McCallum, and F. C. N. Pereira, “Conditional random fields: Probabilistic models for segmenting and labeling sequence data,” 2001.

[43] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, “An efficient k-means clustering algorithm: analysis and implementation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 881–892, Jul. 2002.

[44] S. Thirumuruganathan, “A detailed introduction to K-nearest neighbor (KNN) algorithm,” 2010.

[45] K. G. Derpanis, “Mean Shift Clustering,” pp. 1–3, 2005.

[46] C. J. C. Burges, “A Tutorial on Support Vector Machines for Pattern Recognition,” *Data Min. Knowl. Discov.*, vol. 2, no. 2, pp. 121–167, 1998.

[47] D. Ramage, “Hidden Markov models fundamentals,” 2007.

[48] P. Senin, “Dynamic Time Warping Algorithm Review,” in *Information and Computer Science Department University of Hawaii at Manoa Honolulu*, USA 885, 2008, pp. 1–23.

[49] C. Wohler and J. K. Anlauf, “An adaptable time-delay neural-network algorithm for image sequence analysis.,” *IEEE Trans. Neural Netw.*, vol. 10, no. 6, pp. 1531–6, Jan. 1999.

[50] G. J. Holzmann, *Design and validation of computer protocols*. New Jersey: Prentice-Hall, Inc., 1990.

[51] A. Chaudhary, J. L. Raheja, K. Das, and S. Raheja, “A survey on hand gesture recognition in context of soft computing,” in *Communications in Computer and Information Science*, 2011, vol. 133 CCIS, no. PART 3, pp. 46–55.

[52] H.-R. Choi and T. Kim, “Combined Dynamic Time Warping with Multiple Sensors for 3D Gesture

Recognition," *Sensors*, vol. 17, no. 8, pp. 1–15, 2017.

[53] World Health Organization (WHO), "Deafness and hearing loss," 2017. [Online]. Available: http://www.who.int/en/news-room/fact-sheets/detail/deafness-and-hearing-loss.

[54] K. K. R. Ashok K Sahoo, Gouri Sankar Mishra, "Sign Language Recognition: State Of The Art," *ARPN J. Eng. Appl. Sci.*, vol. 9, no. 2, pp. 116–134, 2014.

[55] A. M. Martinez, R. B. Wilbur, R. Shay, and A. C. Kak, "Purdue RVL-SLLL ASL database for automatic recognition of American Sign Language," in *Proceedings. Fourth IEEE International Conference on Multimodal Interfaces*, 2002, pp. 167–172.

[56] P. Dreuw, C. Neidle, V. Athitsos, S. Sclaroff, and H. Ney, "Benchmark Databases for Video-Based Automatic Sign Language Recognition.," in *Proceedings of the Sixth International Conference on Language Resources and Evaluation (Lrec)*, 2008, pp. 1–6.

[57] Johnston, T., A. Schembri, R. Adam, J. Napier, and D. Thornton, "Auslan SignBank: the Auslan lexical database." [Online]. Available: http://www.auslan.org.au/.

[58] K. Assaleh, T. Shanableh, and M. Zourob, "Low Complexity Classification System for Glove-Based Arabic Sign Language Recognition," in *Proc. 19th International Conference. Neural Information Processing*, vol. 7665, Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 262–268.

[59] S. Shohieb, H. Elminir, and A. Riad, "SignsWorld Atlas; a benchmark Arabic Sign Language database," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 27, no. 1, pp. 68–76, 2015.

[60] A. Schembri, J. Fenlon, R. Rentelis, S. Reynolds, and K. Cormier, "Building the British Sign Language Corpus," *Lang. Doc. Conserv.*, vol. 7, pp. 136–154, 2013.

[61] A. J. Porfirio, K. L. Wiggers, L. E. S. Oliveira, and D. Weingaertner, "LIBRAS Sign Language Hand Configuration Recognition Based on 3D Meshes," in *2013 IEEE International Conference on Systems, Man, and Cybernetics*, 2013, pp. 1588–1593.

[62] U. Agris and K.-F. Kraiss, "The SIGNUM Database," *International Gesture Workshop on Gesture in Human-Computer Interaction and Simulation*, 2009. [Online]. Available: http://www.phonetik.uni-muenchen.de/forschung/Bas/SIGNUM/.

[63] U. Von Agris and K.-F. Kraiss, "Towards a Video Corpus for Signer-Independent Continuous Sign Language Recognition," *GW 2007 7th Int. Work. Gesture Human-Computer Interact. Simul.*, pp. 10–11, 2007.

[64] E. Efthimiou and S.-E. Fotinea, "GSLC: Creation and Annotation of a Greek Sign Language Corpus for HCI," in *Universal Acess in Human Computer Interaction. Coping with Diversity*, vol. 4554, Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 657–666.

[65] P. V. V. Kishore, P. R. Kumar, E. K. Kumar, and S. R. C. Kishore, "Video Audio Interface for Recognizing Gestures of Indian Sign Language," *Int. J. Image Process.*, vol. 5, no. 4, pp. 479–503, 2011.

[66] J. Bungeroth, D. Stein, P. Dreuw, H. Ney, S. Morrissey, A. Way, and L. van Zijl, "The ATIS sign language corpus," in *Proceedings of the International Conference on Language Resources and Evaluation (LREC)*, 2008, pp. 2943–2946.

[67] M. Fagiani, E. Principi, S. Squartini, and F. Piazza, "A New Italian Sign Language Database," in *In Proceeding of International Conference on Brain Inspired Cognitive Systems (BICS),* vol. 3139,

Shenyang, China: Springer Berlin Heidelberg, 2012, pp. 164–173.

[68] J. S. Kim, W. Jang, and Z. Bien, "A dynamic gesture recognition system for the Korean sign language (KSL).," *IEEE Trans. Syst. Man, Cybern. Part B*, vol. 26, no. 2, pp. 354–359, 1996.

[69] H. A. Q. Maarif, R. Akmeliawati, and S. Bilal, "Malaysian Sign Language database for research," in *2012 International Conference on Computer and Communication Engineering (ICCCE)*, 2012, pp. 798–801.

[70] S. Kausar, M. Javed, and S. Sohail, "Recognition of gestures in Pakistani sign language using fuzzy classifier," *Proc. 8th Int. Conf. Signal Process. Comput. Geom. Artif. Vis.*, pp. 101–105, 2008.

[71] A. Karami, B. Zanj, and A. K. Sarkaleh, "Persian sign language (PSL) recognition using wavelet transform and neural networks," *Expert Syst. Appl.*, vol. 38, no. 3, pp. 2661–2667, 2011.

[72] R. San-Segundo, J. M. Pardo, J. Ferreiros, V. Sama, R. Barra-Chicote, J. M. Lucas, D. Sánchez, and A. García, "Spoken Spanish generation from sign language," *Interact. Comput.*, vol. 22, no. 2, pp. 123–139, Mar. 2010.

[73] O. Aran, I. Ari, A. Guvensan, H. Haberdar, Z. Kurt, I. Turkmen, A. Uyar, and L. Akarun, "A Database of Non-Manual Signs in Turkish Sign Language," in *Signal Processing and Communications Applications. IEEE 15th*, 2007, pp. 1–4.

[74] K. Kirchhoff, J. Bilmes, S. Das, N. Duta, M. Egan, G. Ji, F. He, J. Henderson, D. Liu, and M. Noamany, "Novel approaches to Arabic speech recognition: report from the 2002 Johns-Hopkins summer workshop," in *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03)*, 2003, pp. 344–347.

[75] N. Y. Habash, *Introduction to Arabic Natural Language Processing*, vol. 3, no. 1. Morgan and Claypool Publishers, USA, 2010.

[76] M. Alghamdi, M. Elshafei, and H. Al-Muhtaseb, "Arabic broadcast news transcription system," *Int. J. Speech Technol.*, vol. 10, no. 4, pp. 183–195, 2007.

[77] M. A. M. Abushariah, "TAMEEM V1. 0: speakers and text independent Arabic automatic continuous speech recognizer," *Int. J. Speech Technol.*, vol. 20, no. 2, pp. 261–280, 2017.

[78] M. A. M. Abushariah, R. N. Ainon, R. Zainuddin, M. Elshafei, and O. O. Khalifa, "Phonetically rich and balanced text and speech corpora for Arabic language," *Lang. Resour. Eval.*, vol. 46, no. 4, pp. 601–634, 2012.

[79] A. Soudi, G. Neumann, and A. Van den Bosch, "Arabic computational morphology: knowledge-based and empirical methods," in *Arabic Computational Morphology*, Springer, 2007, pp. 3–14.

[80] M. Elmahdy, R. Gruhn, W. Minker, and S. Abdennadher, "Survey on common Arabic language forms from a speech recognition point of view," in *proceeding of International conference on Acoustics (NAG-DAGA)*, 2009, pp. 63–66.

[81] A. Farghaly and K. Shaalan, "Arabic natural language processing: Challenges and solutions," *ACM Trans. Asian Lang. Inf. Process.*, vol. 8, no. 4, p. 14, 2009.

[82] H. Althobaiti and C. Lu, "A survey on Arabic optical character recognition and an isolated handwritten Arabic character recognition algorithm using encoded freeman chain code," in *Information Sciences and Systems (CISS), 2017 51st Annual Conference on*, 2017, pp. 1–6.

[83] M. Attia and H. Somers, *Handling Arabic morphological and syntactic ambiguity within the LFG*

framework with a view to machine translation, vol. 279. University of Manchester Manchester, 2008.

[84] N. El-Bendary, H. M. Zawbaa, M. S. Daoud, A. E. Hassanien, and K. Nakamatsu, "ArSLAT: Arabic Sign Language Alphabets Translator," in *2010 International Conference on Computer Information Systems and Industrial Management Applications, CISIM 2010*, 2010, pp. 590–595.

[85] S. M. Halawani, "Arabic sign language translation system on mobile devices," *Int. J. Comput. Sci. Netw. Secur. (IJCSNS),* vol. 8, no. 1, pp. 251–256, 2008.

[86] T. Aujeszky and M. Eid, "A gesture recogintion architecture for Arabic sign language communication system," *Multimed. Tools Appl.*, vol. 75, no. 14, pp. 8493–8511, 2016.

[87] M. Tolba, A. Samir, and M. Abul-Ela, "A proposed graph matching technique for Arabic sign language continuous sentences recognition," *2012 8th Int Conf Informatics Syst*, pp. 14–20, 2012.

[88] M. A. Abdel-Fattah, "Arabic Sign Language: A Perspective," *J. Deaf Stud. Deaf Educ.*, vol. 10, no. 2, pp. 212–221, 2005.

[89] A. Almohimeed, M. Wald, and R. I. Damper, "Arabic Text to Arabic Sign Language Translation System for the Deaf and Hearing-Impaired Community," in *Proceedings of the Second Workshop on Speech and Language Processing for Assistive Technologies*, 2011, pp. 101–109.

[90] N. Jemina, M. Rachel, and D. Goswell, *Sign language interpreting : theory and practice in Australia and New Zealand*, 2nd ed. Annandale, N.S.W. : Federation Press, 2010.

[91] "The Arabic Dictionary of Gestures for the Deaf.," 2007. [Online]. Available: http://www.menasy.com/arab Dictionary for the deaf 2.pdf.

[92] O. Altun and S. Albayrak, "Turkish fingerspelling recognition system using Generalized Hough Transform, interest regions, and local descriptors," *Pattern Recognit. Lett.*, vol. 32, no. 13, pp. 1626–1632, 2011.

[93] S. Smreen, "Linguistic Culture of Deaf," in *In: International Workshop on "Deaf Communication", Tunisia (in Arabic)*, 2004.

[94] Y. O. Mohamed Elhadj, Z. Zemirli, and K. Ayyadi, "Development of a bilingual parallel corpus of Arabic and Saudi Sign Language: Part I," in *Advances in Intelligent Systems and Computing*, 2013, pp. 285–295.

[95] M. Mohandes, S. Aliyu, and M. Deriche, "Arabic sign language recognition using the leap motion controller," *2014 IEEE 23rd Int. Symp. Ind. Electron.*, pp. 960–965, 2014.

[96] A. Samir Elons, M. Abull-ela, and M. F. Tolba, "Neutralizing lighting non-homogeneity and background size in PCNN image signature for Arabic Sign Language recognition," *Neural Comput. Appl.*, vol. 22, no. SUPPL.1, pp. 47–53, 2013.

[97] A. SamirElons, M. Abull-Ela, and M. F. Tolba, "Pulse-coupled neural network feature generation model for arabic sign language recognition," *IET Image Process.*, vol. 7, no. 9, pp. 829–836, 2013.

[98] R. Naoum, H. Owaied, and S. Joudeh, "Development of a new Arabic sign language recognition using k-nearest neighbor algorithm," *J. Emerg. Trends Comput. Inf. Sci.*, vol. 3, no. 8, pp. 1173–1178, 2012.

[99] E. E. Hemayed and A. S. Hassanien, "Edge-based recognizer for Arabic sign language alphabet (ArS2V-Arabic sign to voice)," in *Computer Engineering Conference (ICENCO). IEEE*, 2011, pp.

121–127.

[100] M. Maraqa and R. Abu-Zaiter, “Recognition of Arabic Sign Language (ArSL) using recurrent neural networks,” *1st Int. Conf. Appl. Digit. Inf. Web Technol. ICADIWT 2008*, vol. 2012, no. February, pp. 478–481, 2008.

[101] O. Al-Jarrah and F. A. Al-Omari, “Improving Gesture Recognition in The Arabic Sign Language Using Texture Analysis,” *Appl. Artif. Intell.*, vol. 21, no. 1, pp. 11–33, 2007.

[102] K. Assaleh and M. Al-Rousan, “Recognition of Arabic Sign Language Alphabet Using Polynomial Classifiers,” *EURASIP J. Adv. Signal Process.*, vol. 2005, no. 13, pp. 2136–2145, 2005.

[103] M. Al-Rousan and M. Hussain, “Automatic recognition of Arabic sign language finger spelling,” *Int. J. Comput. Their Appl. - IJCA- Spec. issue Fuzzy Syst.*, vol. 8, no. 2, pp. 80–88, 2001.

[104] O. Al-Jarrah and A. Halawani, “Recognition of gestures in Arabic sign language using neuro-fuzzy systems,” *Artif. Intell.*, vol. 133, no. 1–2, pp. 117–138, 2001.

[105] S. Aliyu, M. Mohandes, M. Deriche, and S. Badran, “Arabie sign language recognition using the Microsoft Kinect,” in *2016 13th International Multi-Conference on Systems, Signals & Devices (SSD)*, 2016, pp. 301–306.

[106] O. Amin, H. Said, A. Samy, and H. K. Mohammed, “HMM based automatic Arabic sign language translator using Kinect,” in *2015 Tenth International Conference on Computer Engineering & Systems (ICCES)*, 2016, pp. 389–392.

[107] A. A. Ahmed and S. Aly, “Appearance-based Arabic Sign Language recognition using Hidden Markov Models,” in *2014 International Conference on Engineering and Technology (ICET)*, 2014, pp. 1–6.

[108] A. S. Elons, M. Abull-Ela, and M. F. Tolba, “A proposed PCNN features quality optimization technique for pose-invariant 3D Arabic sign language recognition,” *Appl. Soft Comput. J.*, vol. 13, no. 4, pp. 1646–1660, 2013.

[109] F. Fares, A. Mashagba, E. Fares, A. Mashagba, and M. O. Nassar, “Automatic isolated-word Arabic sign language recognition system based on time delay neural networks: New improvements,” *J. Theor. Appl. Inf. Technol.*, vol. 57, no. 11, pp. 42–47, 2013.

[110] M. Mohandes, M. Deriche, U. Johar, and S. Ilyas, “A signer-independent Arabic Sign Language recognition system using face detection, geometric features, and a Hidden Markov Model,” *Comput. Electr. Eng.*, vol. 38, no. 2, pp. 422–433, 2012.

[111] N. R. Albelwi and Y. M. Alginahi, “Real-Time Arabic Sign Language (ArSL) Recognition,” in *International Conference on Communications and Information Technology (ICCIT)*, 2012, pp. 497–501.

[112] A. A. . Youssif, A. E. Aboutabl, and H. H. Ali, “Arabic Sign Language (ArSL) Recognition System Using HMM,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 2, no. 11, pp. 45–51, 2011.

[113] M. Al-Rousan, O. Al-Jarrah, and M. Al-Hammouri, “Recognition of dynamic gestures in arabic sign language using two stages hierarchical scheme,” *Int. J. Knowledge-Based Intell. Eng. Syst.*, vol. 14, no. 3, pp. 139–152, 2010.

[114] M. AL-Rousan, K. Assaleh, and A. Tala’a, “Video-based signer-independent Arabic sign language recognition using hidden Markov models,” *Appl. Soft Comput.*, vol. 9, no. 3, pp. 990–999, 2009.

[115] T. Shanableh and K. Assaleh, "Video-based feature extraction techniques for isolated Arabic sign language recognition," in *9th International Symposium on Signal Processing and Its Applications (ISSPA)*, 2007, pp. 1–5.

[116] M. Mohandes and M. Deriche, "Arabic sign language recognition by decisions fusion using Dempster-Shafer theory of evidence," in *Computing, Communications and IT Applications Conference (ComComAp)*, 2013, pp. 90–94.

[117] M. A. Mohandes, "Recognition of Two-Handed Arabic Signs Using the CyberGlove," *Arab. J. Sci. Eng.*, vol. 38, no. 3, pp. 669–677, 2013.

[118] T. Ritchings, A. Khadragi, and M. Saeb, "An Intelligent Computer-Based System for Sign Language Tutoring," *Assist. Technol.*, vol. 24, no. 4, pp. 299–308, 2012.

[119] M. Mohandes and S. A-Buraiky, "Automation of the Arabic Sign Language Recognition using the PowerGlove," *Artif. Intell. Mach. Learn.*, vol. 7, no. 1, pp. 41–46, 2007.

[120] M. F. Tolba, A. Samir, and M. Aboul-Ela, "Arabic sign language continuous sentences recognition using PCNN and graph matching," *Neural Comput. Appl.*, vol. 23, no. 3, pp. 999–1010, 2013.

[121] K. Assaleh, T. Shanableh, M. Fanaswala, F. Amin, and H. Bajaj, "Continuous Arabic Sign Language Recognition in User Dependent Mode," *J. Intell. Learn. Syst. Appl.*, vol. 02, no. 01, pp. 19–27, 2010.

[122] K. Assaleh, T. Shanableh, M. Fanaswala, H. Bajaj, and F. Amin, "Vision-based system for continuous arabic sign language recognition in user dependent mode," *Proceeding 5th Int. Symp. Mechatronics its Appl. ISMA 2008*, pp. 25–29, 2008.

[123] M. Tuffaha, T. Shanableh, and K. Assaleh, "Novel Feature Extraction and Classification Technique for Sensor-Based Continuous Arabic Sign Language Recognition," in *Neural Information Processing: 22nd International Conference, ICONIP, Part IV*, 2015, pp. 290–299.

[124] E. Aghajari and D. Gharpure, "Real Time Vision-Based Hand Gesture Recognition for Robotic Application," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 4, no. 3, pp. 2277–128, 2014.

[125] E. Stergiopoulou and N. Papamarkos, "Hand gesture recognition using a neural network shape fitting technique," *Eng. Appl. Artif. Intell.*, vol. 22, no. 8, pp. 1141–1158, Dec. 2009.

[126] S. Kausar and M. Y. Javed, "A Survey on Sign Language Recognition," in *Frontiers of Information Technology (FIT)*, 2011, pp. 95–98.

[127] H. Hasan and S. Abdul-Kareem, "Human–computer interaction using vision-based hand gesture recognition systems: a survey," *Neural Comput. Appl.*, vol. 25, no. 2, pp. 251–261, 2014.

[128] Y. Zhou, G. Jiang, and Y. Lin, "A novel finger and hand pose estimation technique for real-time hand gesture recognition," *Pattern Recognit.*, vol. 49, pp. 102–114, 2016.

[129] K. Oka, Y. Sato, and H. Koike, "Real-time fingertip tracking and gesture recognition," *IEEE Comput. Graph. Appl.*, vol. 22, no. 6, pp. 64–71, 2002.

[130] J. Triesch and C. Von Der Malsburg, "A system for person-independent hand posture recognition against complex backgrounds," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 12, pp. 1449–1453, 2001.

[131] M. a Berbar, H. M. Kelash, and A. a Kandeel, "Faces and Facial Features Detection in Color Images," in *Geometric Modeling and Imaging – New Trends (GMAI06)*, 2006, pp. 209–214.

[132] Z. Zafrulla, H. Brashear, T. Starner, H. Hamilton, and P. Presti, "American sign language recognition with the kinect," in *Proceedings of the 13th international conference on multimodal interfaces*, 2011, pp. 279–286.

[133] A. Memis and S. Albayrak, "Turkish Sign Language recognition using spatio-temporal features on Kinect RGB video sequences and depth maps," in *Signal Processing and Communications Applications Conference (SIU)*, 2013, pp. 1–4.

[134] Ahmed, A. M., Alez, R. A., Taha, M., & Tharwat, G. (2016). Automatic translation of Arabic sign to Arabic text (ATASAT) system. Journal of Computer Science and Information Technology, 6, 109-122.