ISSN: 2093-5374 / E-ISSN: 2093-5382

Comparison of Collective Diverse Arabic Sign Language Dataset

Zaid Saad Bilal[1*], Amir Gargouri[2], Hanaa F Mahmood[3], and Hassene Mnif[4]

[1*]National School of Electronics and Telecommunications, University of Sfax, Tunisia LETI Laboratory-ENIS, Tunisia. zaidsaadd1988@gmail.com, https://orcid.org/0009-0002-0602-6220

[2]Laboratory of Signals Systems Artificial Intelligence and networks "SM@RTS" Sfax University Sfax, Tunisia. amir.gargouri@enetcom.usf.tn, https://orcid.org/0009-0003-0377-4511

[3]College of Education for Pure Science, Department of Computer Science, University of Mosul, Iraq. dr.hanah@uomosul.edu.iq, https://orcid.org/0000-0001-5322-441X

[4]National School of Electronics and Telecommunications, University of Sfax, Tunisia LETI Laboratory-ENIS, Tunisia. hassene.mnif@enetcom.usf.tn, https://orcid.org/0000-0002-5912-752X

Received: July 23, 2024; Revised: September 02, 2024; Accepted: October 01, 2024; Published: December 30, 2024

Abstract

Machine learning researchers from all around the world continue to work out the best ways to collect data efficiently. Data collecting has recently emerged as a key concern for two primary reasons. Despite the fact that machine learning is making significant progress, there may not be enough labelled data for some new applications. Furthermore, deep learning methods have the benefit of automatically creating features, which is not the case with traditional machine learning methods. With this, model design becomes more affordable, although more labelled data may be required. Particularly, the collection of data research has been on the increase in recent years including data management, computer vision, machine learning, and natural language processing. The main reason for this is that it's necessary to handle and process enormous quantities of data successfully. This study primarily aims to present a publicly available dataset comparison that includes large samples of Arabic sign language images for the goal of sign language classification. This dataset collection has a lot of different videos and images that show different moves. The primary objective of this comparison is to show that there are different types of sign language datasets such as words based sign and alphabet based sign furthermore, the comparison include the Background of the Datasets, the Size of the Datasets, the Number of Samples, the Number of Training, Testing, and Validation Samples, the dataset types, and RGB or Binary Images. The main goals of future study will be to improve the method and test the model using AASL-annotated data.

Keyword: Arabic Alphabet Sign Language, Annotated Dataset, Activity Recognition, Dataset Collection, Image and Video.

1 Introduction

There has been a dramatic rise in the production and consumption of video material (Zenith Media, 2019). Video categorization (Karpathy et al., 2014), object identification (Liu et al., 2016; Ren et al.,

Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), volume: 15, number: 4 (December), pp. 133-150. DOI: *10.58346/JOWUA.2024.I4.009*

*Corresponding author: National School of Electronics and Telecommunications, University of Sfax, Tunisia LETI Laboratory-ENIS, Tunisia.

2016), tracking of objects (Wang et al., 2019), recognition of actions (Carreira & Zisserman, 2017; Madhan & Shanmugapriya, 2024; Kumar et al., 2023), visual query answering (Fukui et al., 2016), and video encoding (Lu et al., 2019) are some of the video data tasks that have been heavily focused on utilizing deep learning (Verkholyak et al., 2021). A number of these tasks are useful in the real world, and a few of them also have significant implications for accessibility.

Specifically, SLR (sign language recognition) is best understood as an activity recognition problem. The majority of the world's deaf population uses sign languages, which are visual forms of communication. Sign language is most often used to communicate spoken meaning via the use of hand, the arm, and face gestures. The syntax and vocabulary of sign languages are completely distinct from those of spoken languages. Even while certain languages share features, languages often vary from one area to another and are not always mutually intelligible (Ashish Doval, 2013). Their vocabulary and articulation rate are also different from regional languages (Gibson & Salamonson, 2023).

According to the World Health Organization (2015), around 1 billion people, or 14% of the global population, reside in Arab nations and utilize Arabic alphabets. Persian, Malay (Jawi), Uyghur, Pashto, Punjabi, Balti, Sindhi, Lurish, Balochi, Rohingya, Kashmiri, Urdu, Somali, Mandinka, and many more languages and dialects throughout Asia and Africa make use of the Arabic script (Mirdehghan, 2010). This clearly demonstrates the importance of Arabic alphabets, as they are utilized by about 25% of the global population.

The rule of multiple proportions states that among populations using Arabic alphabetic script, 5% have hearing loss. A large number like this is taken seriously. The significance of the Arabic Sign Language Alphabets (ArSLA) is therefore established. A representation of the Arabic alphabet in the form of sign language forms is known as the Arabic Sign Language Alphabets (ArSLA). Therefore, the hearing-impaired population use ArSLA as a means of overcoming the challenge of interacting with conventional Arabic language. Because of this, they are able to take part in the conventional pedagogical and educational practices (El-Bendary et al., 2010).

There have been a number of initiatives to create machine translation systems for sign language. Although the majority of these efforts focus on American or French languages rather than Arabic, a small number of them are using 3D human avatars to teach Arabic (Halawani, 2008; Vijay et al., 2022). The use of an avatar rather than a video is preferable since, with videos, various people may be animating different signals, which may lead to several participants in a single translated phrase. In situations where a human translator is not available, a machine translation system can be useful. Current studies have mainly concentrated on processing video images, but there has been substantial progress in examining methods of sign recognizing using spatial information from a single frame (Madhan & Shanmugapriya, 2024). This highlights the need to establish a set of a standard image datasets.

1.1. Research Gap

One of the main elements that change the accuracy of a system is the change of viewing situations. However, there are no benchmark sign databases, which makes it hard to compare current techniques and assess the impact of these approaches. This is also highlighted in (Camgözlü & Kutlu, 2023). The purpose of this study is to compare the latest Arabian Sign Language Collection datasets that have been made available to the public. This study is very helpful for researchers who want to create a recognition model for sign language in the future.

1.2. Contribution of this Research

Follow these steps to summarize the key findings of the present research:

- Giving details about two new Arabic sign language datasets that are available to the public, named Arabic Alphabet Sign Language (AASL) and Arabic Sign Language (ASL).
- Comparing between words based sign language and alphabet based sign language.
- Comparing the dataset's Background, Size, Number of Samples, Number of Training, Testing, and Validation Samples, as well as their RGB or Binary Images and Number of Samples.

1.3. Organization of this Research

This is the remaining structure of the paper. By collecting datasets and a sign language recognition system, Section 2 provides an outline of the relevant literature. In Section 3, we cover the actual need for the AASL dataset. Section 4 provides an example of a fundamental assessment of essential body signs and Arabic sign language. Section 5 provides a more in-depth explanation of the Arabic Alphabet Sign Language (AASL) dataset which is alphabet based sign language, including its collection and characteristics. The description of the Arabic Sign Language (ASL) dataset which is words based sign language and a number of examples of images are in Section 6. In Section 7, we compare the ASL and AASL datasets in further detail. The article and the paper are summed up in Section 9, which also compares the ASL and AASL datasets.

2 Related Work on Sign Language

The majority of speakers of sign languages, such as American Sign Language (ASL), are Deaf individuals, who themselves constitute a cultural minority. The use of languages and modalities that are not widely used creates several communication obstacles such as being excluded from voice-controlled agents, which are becoming more common. There are inadequate literacy rates among deaf children (Al-Dawoodi, 2015) and Deaf Education institutions fail to adequately promote first language learning (Kurbanazarova et al., 2024; Hall, 2017; Hall et al., 2017), thus even written expression of the majority of language might be inaccessible. Community remains small, therefore Deaf people typically know one another, despite modern technology making it easier for Deaf individuals from around the globe to connect (Hasan, 2023; Miyata et al., 2009). As members of the community can more readily identify individuals, privacy issues over video sharing may intensify. The term "audism" refers to the exclusion of the Deaf community because of their hearing impairment (Palmer et al., 2012). Various kinds of audism exist, such as state-mandated sterilization (Bauman, 2004) or the rejection of Deaf patients or workers by physicians or businesses. Some technology, including voice-controlled devices and airport PA announcements, unintentionally exclude Deaf persons due to the presence of audism. Decisions that will have a significant impact on the lives of the deaf community are often left out. An iconic case in point is the Deaf President Now demonstration when students at the only Deaf-specific university—a place that has always been governed by hearing people—demanded Deaf leadership (Lombardo, 2022; El et al., 2014). Like spoken words, signs are composed of phonological elements, which were once believed to include handshape, position, and movement (Rosa et al., 2024). The presence of audism may amplify privacy issues over video sharing by increasing the danger of personal identification and revelation of Deafness. A more exact and complete list of features is provided by current hypotheses (Stokoe, 1960). Eye contact, movement of the eyebrows or lips, or changes in head or body position may all contribute to the production of grammatical information without the need of hands (Brentari, 1996; Brentari, 2018) for reviews (Gümüş et al., 2022). In contrast to speech, which usually consists of a single sound, many characteristics (such as the handshape and its position) occur concurrently. In a world where data is already limited, the complexity of the language makes the amount of data required for successful modeling much more challenging.

One machine translation system that uses the Rule-based Interlingua technique is TEAM (Sanchez-Ancajima et al., 2022), which stands for Translation from the English language to American Sign Language, or ASL by Machine. There are two parts to TEAM. Firstly, it takes an English phrase as input and uses morphological information to create glosses, which are intermediate presentations. Secondly, it uses these glosses to control the human model and make ASL signs. Using a voice recognizer to convert spoken words into text sentences and then into signs is what makes up a speech-to-sign language translation system for Spanish (Zhao et al., 2000). It employs two methods: statistical methods and rule-based procedures. An eSIGN 3D avatar brings the translated words to life.

For Arabic sign language, San-Segundo et al., (2008) created a translating system. There is a knowledge method that the translation system is built on. To get around the issue of terms not being in the database, the technique incorporates finger spelling translation.

3 Importance of Dataset Datasets on ML Technology

Rapidly creating, evaluating, and refining models is the process that yields the best machine learning product. A crucial component of the success of models is the datasets themselves. For this reason, similar to modelling, data collection, preparation, and labelling should be viewed as an iterative procedure (Gibson & Salamonson, 2023).

Using data in this iterative fashion could seem complicated at first. Standard datasets are often used to report performance in ML research (Jonnerby et al., 2023; Christiansen & Barnartt, 2003). These datasets are considered benchmarks by the community and cannot be changed. Since deterministic rules are often written for programs in conventional software engineering, data is typically seen as something to be received, processed, and stored.

When it comes to creating things, ML technology is all about combining design with ML (Fiorino et al., 2019). So, our dataset is really one more tool in our toolbox for product development. Most of the time in ML development goes into selecting an initial dataset, maintaining and improving it on a regular basis, and adding to it. Figure 1 shows how the research and industrial workflows differ on dataset creating.

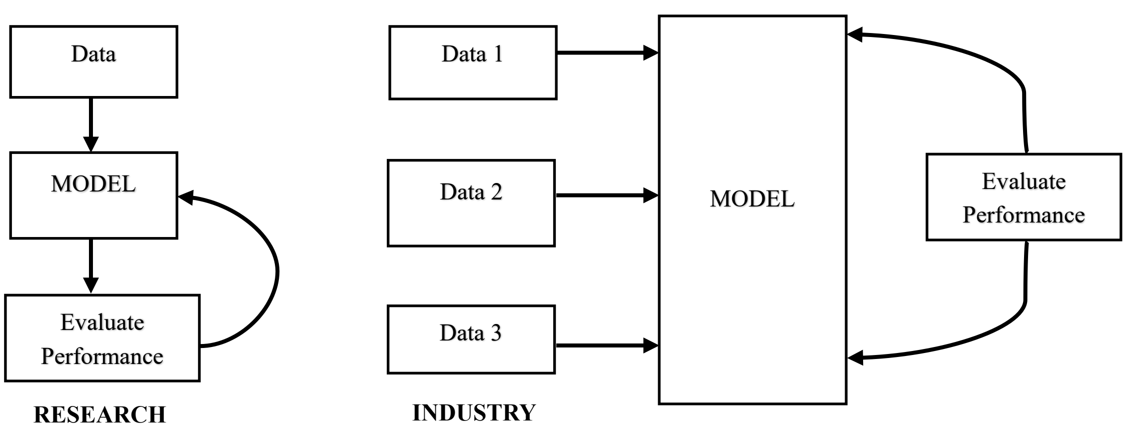


Figure 1: Dataset Analysis on Research and Industry Model

3.1. Strategies for Enhancing Communication for Individuals with Hearing Loss

The deaf community often uses sign languages as their primary means of communication. Deaf languages develop organically among deaf populations, much like any spoken language. These groups lose touch with the regional language wherever they settle and grow. Different populations in these areas utilize different sign languages, such as Arabic, American Sign Language, and French Sign Language.

Sign language and spoken language are quite different. The construction and reception of the communication components, however, constitute the primary distinction between them (Nirmala, 2023). Fingerspelling and cued speech are two additional forms of communication used by the deaf community, in addition to sign languages.

Similar to how the Arabic alphabet stands for the individual phonemes that comprise a word when spoken, the visual acoustical (optical phonemes) of fingerspelling form the indicating symbols of the sign language alphabet.

Signing out the letters of the alphabet with either one or both hands is called finger spelling. Arabic sign language includes finger spelling, which serves several functions. People who are deaf, as well as instructors and interpreters, may use it to spell people's names, confirm or explain something, or represent words without a corresponding sign. The deaf youngster may learn to read and write Arabic using it as a conceptual framework in educational activities. Another option is to employ a pronunciation pattern or a combination of them to indicate a guiding word when instructing people in sign language (Yan, 1993; Ahmed et al., 2017).

4 Sign Language and Diglossia

Some have proposed that ASL, BSL, and DSL (Danish Sign Language) have diglossia characteristics (Deuchar, 1977; Stokoe, 1969; Lawson, 1981). When it comes to Arabic, everything is different. ARSLs are not diglossia, but Arabic is. The presence of more than one sign language in Arabic was unexpected. Similar to the way there are several spoken varieties of Arabic rather than a single standard form, experts and members of the general public outside of Deaf communities find it hard to comprehend the concept of additional sign language vernaculars. Unfortunately, this has led to the current efforts to standardize ARSL, which so far have failed. Diglossia symptoms are beginning to show up among BSL or ASL interlocutors in non-diagnostic speech communities, which is somewhat different from the situation in the United Kingdom and the United States.

Possible explanations include the "primitive" character of ARSLs and the reliability and complicated nature of BSL and ASL. Furthermore, most Arab nations do not have sign language variants that have been developed as a result of the structured formal education of the deaf in countries like the United Kingdom and the United States. The necessity for a specific variant (H) in addition to the other (L) has grown due to the superior/inferior viewpoints held about sign languages and their spoken equivalents. For example, there are claims that sign languages lack syntax, are incorrect, do not exist, etc. Sign language vernacular learners who are hearing have a hard time understanding Arabic since each word is signed as it would be pronounced. There is a significant distinction between the Arabic letters H and L, but when translating from or communicating in sign language, both are considered as one. Despite this, there is no problem distinguishing between the colloquial and conventional forms of an assertion. Due to the lack of a written form, members of the Deaf community may not be fully aware that L variants exist.

4.1. Arabic Sign Language's Evolution

Worldwide sign languages have been around for a long time. They were as real as any language that people could hear. Nobody may claim credit for their innovation. As with other vocal languages, they evolved organically (Bellis, 2004). Likewise, ARSLs have been evolving in a completely natural manner. Within their "natural environment,"" ARSLs evolved into distinct modes of communication. These aren't actually interpretations of spoken Arabic or standard Arabic. ARSLs exhibit both shared and distinct features. Following all, this is the case for every language; in fact, there are commonalities throughout the world's sign languages that suggest a degree of universality. While certain ARSLs have profited from the pioneering work of others, generally speaking, ARSLs evolved separately.

The practice of finger spelling is very recent and has evolved from a mix of mimicry and originality. For words without a sign equivalent or proper nouns, it is employed to spell them out. On the other hand, reading or communicating the traditional form of Arabic does not need finger spelling. As a result, there isn't currently a "manual Arabic." However, if the deaf are to be instructed with sign language, and the demand for a standard signed Arabic language ever develops, then possibly a type of signed standard Arabic could possibly develop. Additionally, as far as we aware, no one has tried to record ARSLs (sign writing) thus far. Even though American Sign Language (ASL) has its own writing systems, they haven't been utilized much for recording ASL literature. Nevertheless, you can find a lot of ASL literature on DVDs, videotapes, and compact CDs (Wilcox & Peyton, 1999). In contrast, Arabic contains a large amount of signed literature, mostly from films, TV shows, and news broadcasts; nevertheless, this material has neither been documented nor used to create Arabic sign vernaculars.

4.2. Arabic Sign Language (ArSL)

There are five universal factors or features shared by all sign languages (Fukui et al., 2016). The initial one is a sign-making hand form. The second variable is the sign-making palm orientation. Position of the symbol relative to the body is the third. The motion of the sign-making handshape is the fourth parameter. The final one is the meaning-making body language and facial emotions. To create a sign in a discussion, you need to meet these conditions:

- Keep your hand facing forward so that the palm is facing you.
- The palm of the hand should be directly on the shoulders, level with the chin, and not too far away from the body.
- Finger clarity (shape and number);
- No unnecessary finger motions that might tamper with the meaning of the communication;
- When presenting the letters, keep your hand as steady as possible;
- During the presentation, there will be no acceleration.

In (Samrine & Benali), they found three grammatical forms of sign sentences in their extensive research on Arabic sign language: verb + subject + object, subject + verb + object, and object + verb + subject. The ArSL, however, makes extensive use of the first order. A noun, pronoun, verb, and particle are the four main word classes that make up an Arabic sign word.

Nouns

When a term is feminine in gender, the deaf community uses the sign for feminine to communicate this. Deaf people use finger spelling to sign proper nouns including names of people, places, and nations.

Additionally, there are three forms for nouns: single, dual, and plural. There is an ArSL sign for every noun in its single form, but the sign for a dual noun is the combination of the signs for the noun in its singular form plus the number two. There are two alternative indications for nouns in plural. One is the singular version of the noun and the other is the term كثير meaning non-intelligent beings. The second signifies an intelligent entity by means of its single word that is repeated three times.

Pronouns

- The orientation of the hand is mostly used by the deaf to sign personal pronouns such as "you," "they," "he," so on.
- Pronouns that pose questions: (do you) هل, (how) كيف...: the deaf may employ facial expressions and tools that are often present at the initial part of a question phrase to form an interrogative phrase. The sign of reply particles, such as "yes" (نعم) or "no" (لا), is used by the deaf to answer questions.
- The expressive movements may be used to show relative and conditional pronouns, which do not have a sign and are used for males (الذي) and females (التي).
- The use of other Arabic pronouns is unnecessary in ArSL.

Particles

People who are deaf do not place a significance on particles. Their sentencing could be reduced.

Verbs

Time is indicated via tenses. By referencing the location of the body relative to the timeline, the ArSL differentiate between the overarching concepts of the present, the past, and the future. In Figure 2 we can see Arabic sign language in action.

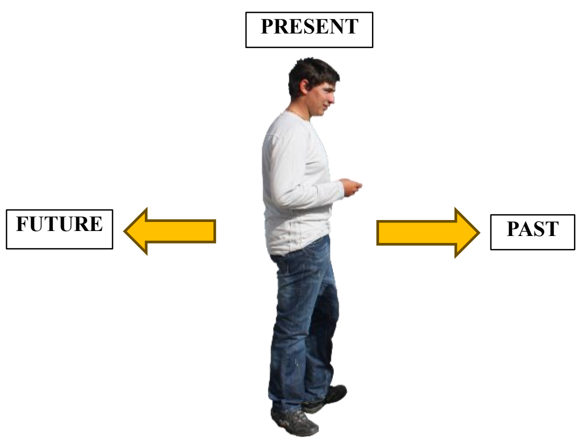


Figure 2: Sign Language for Verbe Tense

5 AASL Dataset

We are required to provide the dataset for the Arabian Alphabet sign language in this research paper which are publicly available in (Shaik, 2020). In this case, the dataset has been compiled from a variety of sources and situations. See Figure 3 for a schematic of this data collecting process. According to this flow diagram, sign language has been gathered from a variety of sources. In the same way, the AASL dataset values are enhanced by collecting the sign alphabet from diverse individuals.

There should be two characteristics in the AASL dataset, and they are:

- Every letter of the Arabic alphabet has to be in the dataset.
- Static images should make up the dataset.

Through a form, participants were requested to submit images they had taken. The Arabic sign language alphabets are categorized into five primary groups, with many alphabets in each group. The participants will be shown the Arabic sign language alphabets and will need to imitate them. All uploaded images are reviewed by hand to ensure they are of high quality and acceptability. The Collected dataset specification in given in Table 1.

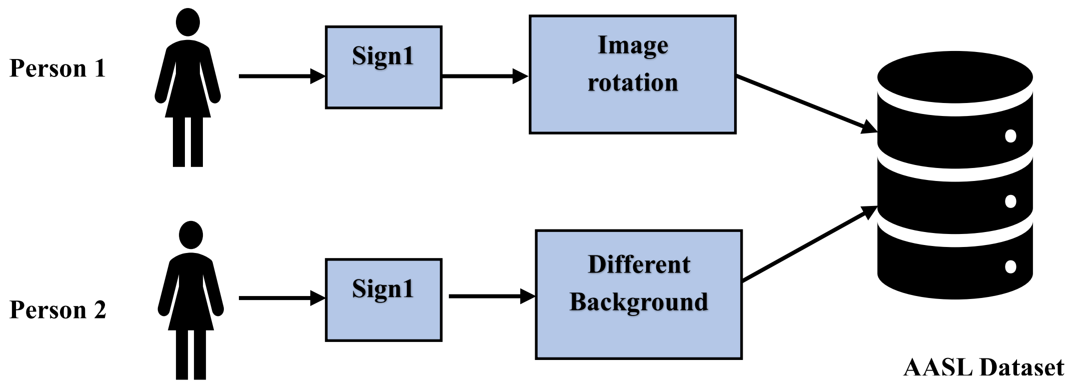


Figure 3: AASL Dataset Collection Procedures

Table 1: AASL Dataset Specification

Dataset	RGB-Image Based Arabic Sign Language Dataset
Type of data	Images
Hardware device	Images in this dataset were acquired using different types of cameras (webcam, digital camera. and camera phone).
Data format	Labelled filtered RGB images with different extensions ('jpg': 6545, 'jpeg': 1211, JPG: 80, JJPEG': 21)
Data source location	Jordan.

5.1. AASL Dataset Collection Procedure

Here, they are going over some information gathered using the laptop's built-in webcam and a mobile phone's smart camera, which lets snap and save many photographs in rapid succession. A technique for recognizing sign language required to create AASL images. The following Table 2 provides a description of the AASL datasets that were developed and utilized as groups in training and testing the system after processing. The dataset will allow researchers and those interested in this field to investigate and create computerized systems for the deaf and hard of hearing employing machine learning techniques, computer vision, deep learning algorithms, human interaction with computers, and these image datasets. All academics and the general public will have access to these comprehensive datasets that begin with alphabetic letters; this will aid the deaf in recognising the Quranic alphabet.

The current research mainly provides the AASL Language Dataset (Shaik, 2020). When they were designing AASL, they had two main concepts in mind:

- Each video should depict a single letter of the alphabet,
- The dataset should have several videos for each of the sign. The signs in the dataset should be varied.

The following method of data collection was devised in accordance with these concepts. Images taken by volunteers of different ages at Jordon make up the AASL dataset. To get these images, they used a smart camera that was mounted on a steady surface. Researchers had the subjects stand around one meter from the camera. Images were altered by changing the backdrop, time, lighting, and viewpoints. Although precisely the number of images for each letter differs, a total of 7857 images were collected for the entire set. The images needed to be pre-processed so they could be used for identification and classification since they were captured in RGB format and had varying size.

5.2. Data Collection

For the purpose for collecting images, an online form was created along with a set of instructions. Individuals might choose to upload photos of themselves doing the alphabets if they felt comfortable doing so. Because of this, the maximum number of images that may be submitted was not specified. Many social media sites have the online form's URL publicized. People of all ages and genders took part in their study, which they conducted at various locations. The participants took images in a variety of settings, including with various cameras, backdrops, lighting, and image sizes. All participants' identities were concealed. The five-month data gathering period began in March 2022.

More than 200 people who shared at least one alphabet worked together to create the RGB Arabic Alphabet Sign Language (AASL) dataset. Webcams, digital cameras, and mobile phone cameras captured the vast majority of the images. There are 7,857 Arabic sign language tagged images in the AASL collection. The images were supervised, verified, and filtered by a team of Arabic sign language specialists to guarantee a high-quality dataset. A total of 31 folders, one for each letter, contain the dataset. The number of images in each folder is highlighted in Table 2, and Figure 4 shows an instance of images for various alphabets from (Shaik, 2020).

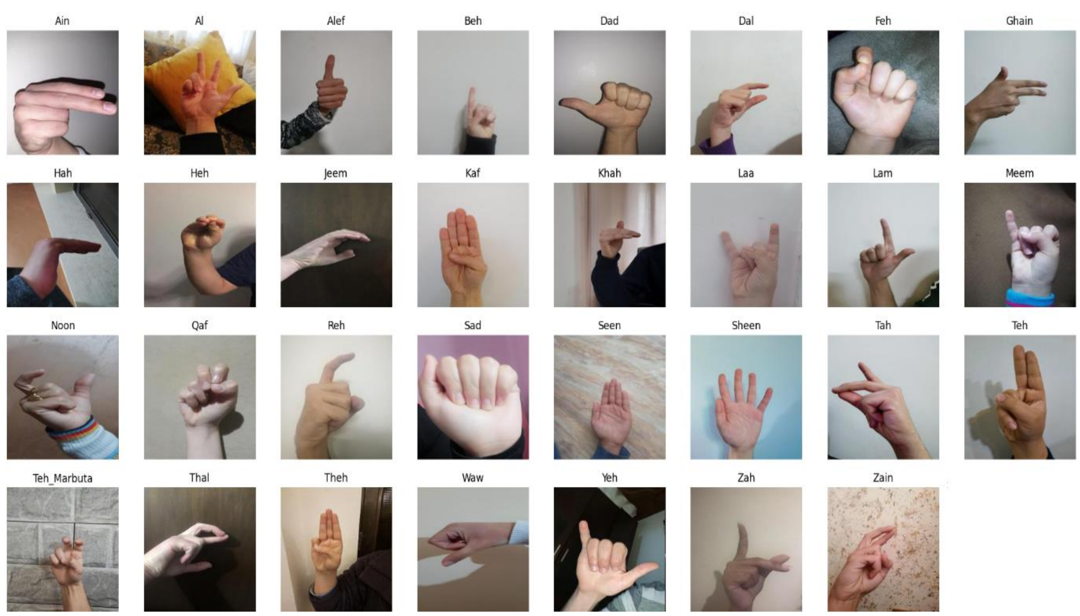


Figure 4: Sample Dataset Images

Table 2: Dataset Description

#	Letter name in English Script	Letter name in Arabic Script	# of Images
1	ALEF	ألف	287
2	BEH	باء	307
3	TEH	تاء	226
4	THEH	ثاء	305
5	JEEM	جيم	210
6	HAH	حاء	246
7	KHAH	خاء	250
8	DAL	دال	235
9	THAL	ذال	202
10	REH	راء	227
11	ZAIN	زاي	201
12	SEEN	سين	266
13	SHEEN	شين	278
14	SAD	صاد	270
15	DAD	ضاد	266
16	TAH	طاء	227
17	ZAH	ظا	232
18	AIN	عين	244
19	GHAIN	غين	231
20	FEH	فاء	255
21	QAF	قاف	219
22	KAF	كاف	264
23	LAM	لام	260
24	MEEM	ميم	253
25	NOON	نون	237
26	HEH	هاء	253
27	WAW	واو	249
28	YEH	ياء	272
29	TEH MARBUTA	تاء مربوطة	257
30	AL	ال	276
31	LAA	لا	268

5.3. Various Circumstance

Bright, luminous, and textural backdrops are the main points of this collection (Shaik, 2020), which aims to demonstrate the transformation of many types of backgrounds. Figure 5 shows a typical example of the dataset collection, which includes images of the same sign taken against several backdrops. In contrast, the degree to which a sign may be distorted in a recorded image is seen in Figure 6. Images saved in the JPEG format have a resolution of 640 X 480 pixels. A number of steps are done to acquire the proper settings of the camera's characteristics in order to guarantee that the captured image visually matches the viewing circumstances.

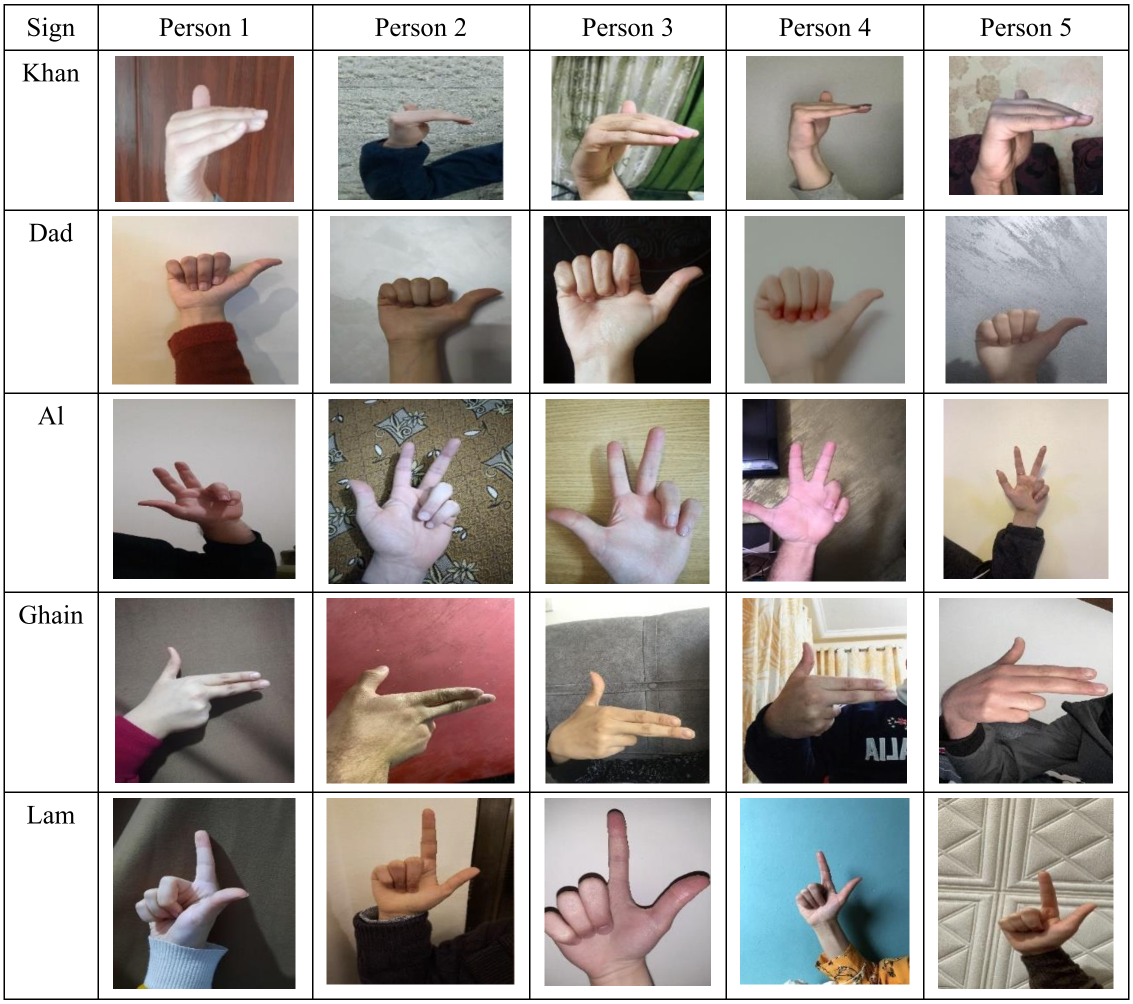


Figure 5: Sample AASL Dataset from Various Orientation

Figure 6: Sample Image for Various Orientation in AASL Dataset

5.4. AASL Dataset Quality

They searched for and validated ground-truth images of static AASL alphabets from specialists in the area of AASL interpretation so that we might make a contribution to the classification of Arabic sign language. In addition, the professionals offered advice on how to correctly execute each letter of the alphabet.

(a) Ground Truth Image (b) Correct Image (c) Wrong Image

Figure 7: AASL Image Correction

Two members of their study team were assigned the responsibility of manually assessing each and every image that was submitted (Shaik, 2020). Their primary function was to verify that submitted images matched the ground-truth images of a given alphabet and that image labels were accurate. The ground-truth depiction of an alphabet is on the left side of Figure 7, whereas the middle and right side display an example of an incorrectly produced alphabet. One member of their research team then performed a last round of review on the whole dataset, making sure that all submitted images were accurate. By the end of the review procedure, the number of right images in the dataset had dropped from 8,042 to 7,857. Lastly, a simple script was executed to automatically label the whole dataset. The labels read "Aphabet-Name ID" on every image. The IDs began at 0 and continued up to the entire number of images in a certain folder that correspond to a particular alphabet, The sample labelled image from AASL dataset is shown in Figure 8 (Shaik, 2020). The value of the AASL dataset is given below.

- The data is adaptable since it is gathered using a variety of parameters, including illumination, backdrop, picture orientation, picture size, and resolution.
- This dataset is ideal for training ML systems to identify Arabic sign language.
- Subject-matter specialists check and double-check the dataset.
- This dataset is the first publicly accessible RGB high-resolution dataset for Arabic sign language, as far as we are aware.

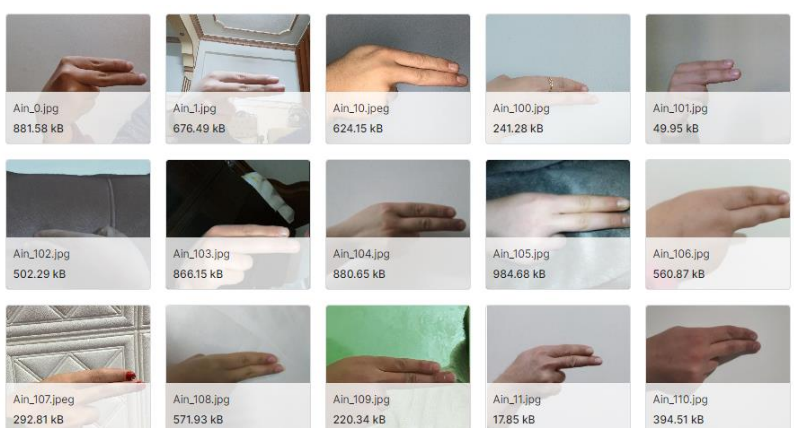


Figure 8: Sample Labelled Image for “Ain” Sign in AASL Dataset

6 ASL Dataset

There are a number of computer vision applications that make use of the Arabic Sign Language dataset (Al-Barham et al., 2023). The three classes that make it up are the following: Angry, Happy and Sorry (Arabic Sign Language Recognition Computer Vision Project, 2022). The collection has a total of 2,889 images. You may see an example of an image from each category in Figure 9. Table 3 states that, the classes of the ASL dataset with the corresponding number of images in each class.

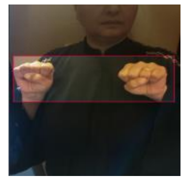
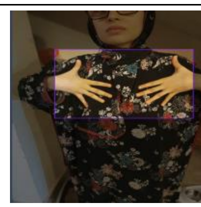
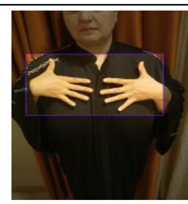
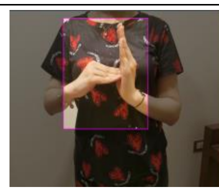
Classes	Sample Image 1	Sample Image 2
Angry		
Happy		
Sorry		

Figure 9: Sample Image of ASL Dataset for Various Classes

Table 3: ASL Dataset Classes Details

Classes	Number of Images
Angry	1000
Happy	889
Sorry	1000

7 Comparison of AASL Vs ASL Dataset

Both the AASL and AASL datasets' performance is detailed here. Table 4 displays the datasets with a detailed explanation.

Table 4: Comparison of AASL and ASL Datasets

Description	AASL	ASL
Size of Datasets	7857	2889
Image or Video or sensor Dataset	Image	Image
Number of Samples	200	Nearly10
Number of Training, Testing, and Validation Samples	Training = 70% Testing = 20% Validation = 10%	Training = 1394 Testing = 917 Validation = 578
Alphabet or Word Datasets	Alphabet	Word of 3 class
Color or Black and White Images	RGB	RGB
Year of Publication	2023	2022

In accordance with the information shown in Table 3, the size of the AASL dataset is three times more than that of the ASL dataset. The number of images in AASL is 7857, but the number of images in ASL is just 2889. While JORDAN is the place for the image collecting information of AASL, the data collection about ASL is unclear. Both sets of data provide examples in the form of images. These are RGB-based images that are used by systems that allow for the identification of sign language (Bellugi & Fischer, 1972). Additionally, in the AASL, a total of 200 samples were used for the purpose of data gathering, while in the ASL, only seven to ten samples are utilized for the same purpose. Sign language alphabet construction is the primary focus of the Arabic Sign Language (AASL) dataset. For the purpose of sign recognition, the ASL paid particular attention to the word of three class images.

8 Discussion

The first RGB sign language dataset developed as AASL dataset is described in this research. This collection of data includes images of signs with the phonemes that go with them, written out in the ArSL alphabet. The author created this dataset because them believe there needs to be more thorough information available to help move the growing field of sign language research and application further.

To build a comprehensive dataset that captures every aspect of sign language expression, we have described in full the steps required to record phonemes and improve their quality in this paper. The method we used for manually reviewing quality images was an essential component of building their collection.

Earlier test in data labelling production have also provided evidence for the relevance of dataset. In order to generate a sign language detection system, they researched the feasibility of using their dataset to train a transformer-based model. The research can demonstrate the innovative and useful nature of their dataset in future sign language identification model.

9 Conclusion

An analysis of the similarities and differences between the features of the Arabic Word Sign Language (ASL) and Arabic Alphabet Sign Language (AASL) datasets is presented in this article. The Arabic Alphabet Sign Language (AASL) dataset is suitable for the AASL Sign Language Recognition model since it is both a big and high-quality dataset. Detailed explanations have been provided for each and every stage of the process. It is made up of 7,857 images, all of which are fully labelled representations of the Arabic alphabet used in sign language. A broad variety of background settings, including but not limited to lighting, background, graphic orientation, image size, and quality, were used in order to capture the Arabic Alphabet Sign Language (AASL) statements of more than two hundred persons. In contrast to the ASL dataset, the AASL dataset is well suited for the development of future model recognition models. Our ultimate goal is to make use of their Arabic Alphabet Sign Language (AASL) dataset in order to construct Arabic sign recognition models that are capable of correctly generalizing to all messages.

References

[1] Ahmed, A. M., Alez, R. A., Tharwat, G., Taha, M., Ghribi, W., Badawy, A. S., ... & Bose, J. S. C. (2017, November). Towards the design of automatic translation system from Arabic Sign Language to Arabic text. In *2017 International Conference on Inventive Computing and Informatics (ICICI)* (pp. 325-330). IEEE. https://doi.org/10.1109/ICICI.2017.8365365

[2] Al-Barham, M., Alsharkawi, A., Al-Yaman, M., Al-Fetyani, M., Elnagar, A., SaAleek, A. A., & Al-Odat, M. (2023). RGB Arabic alphabets sign language dataset. https://doi.org/10.48550/arXiv.2301.11932

[3] Al-Dawoodi, A. G. M. (2015). An improved Bees algorithm local search mechanism for numerical dataset. *Universiti Utara Malaysia*.

[4] Arabic Sign Language Recognition Computer Vision Project, (2022). https://universe.roboflow.com/arabic-sign-language-recognition/arabic-sign-language-recognition

[5] Ashish Doval. (2013). The People's Linguistic Survey of India SIGN LANGUAGE.

[6] Bauman, H. D. L. (2004). Audism: Exploring the metaphysics of oppression. *Journal of deaf studies and deaf education*, *9*(2), 239-246. https://doi.org/10.1093/deafed/enh025

[7] Bellis, M. (2004). Innovations for the hearing impaired. http://inventors.about.com/library/inventors/bltty.htm

[8] Bellugi, U., & Fischer, S. (1972). A comparison of sign language and spoken language. *Cognition*, *1*(2-3), 173-200. https://doi.org/10.1016/0010-0277(72)90018-2

[9] Brentari, D. (1996). Trilled movement: Phonetic realization and formal representation. *Lingua*, *98*(1-3), 43-71. https://doi.org/10.1016/0024-3841(95)00032-1

[10] Brentari, D. (2018). Representing Handshapes in Sign Languages Using Morphological Templates 1. *Gebärdensprachen: Struktur, Erwerb, Verwendung*, *13*, 145.

[11] Camgözlü, Y., & Kutlu, Y. (2023). Leaf Image Classification Based on Pre-trained Convolutional Neural Network Models. *Natural and Engineering Sciences*, *8*(3), 214-232. https://doi.org/10.28978/nesciences.1405175

[12] Carreira, J., & Zisserman, A. (2017). Quo vadis, action recognition? a new model and the kinetics dataset. In *proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 6299-6308). http: //arxiv.org/abs/1705.07750 arXiv: 1705.07750

[13] Christiansen, J. B., & Barnartt, S. N. (2003). *Deaf president now!: The 1988 revolution at Gallaudet University*. Gallaudet University Press.

[14] Deuchar, M. (1977). Sign language diglossia in a British deaf community. *Sign Language Studies*, *17*(1), 347-356. https://doi.org/10.1353/sls.1977.0017

[15] El, A. E. E., El, M. M. R., & El Atawy, S. M. (2014). Intelligent Arabic text to arabic sign language translation for easy deaf communication. *International Journal of Computer Applications*, *92*(8).

[16] El-Bendary, N., Zawbaa, H. M., Daoud, M. S., Hassanien, A. E., & Nakamatsu, K. (2010, October). Arslat: Arabic sign language alphabets translator. In *2010 international conference on computer information systems and industrial management applications (CISIM)* (pp. 590-595). IEEE. https://doi.org/10.1109/CISIM.2010.5643519

[17] Fiorino, D., Collotta, M., & Ferrero, R. (2019). Usability evaluation of touch gestures for Mobile Augmented Reality applications. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 10*(2), 22-36. https://doi.org/10.22667/JOWUA.2019.06.30.022

[18] Fukui, A., Park, D. H., Yang, D., Rohrbach, A., Darrell, T., & Rohrbach, M. (2016). Multimodal compact bilinear pooling for visual question answering and visual grounding. https://doi.org/10.48550/arXiv.1606.01847

[19] Gibson, K., & Salamonson, Y. (2023). Image processing application: Overlapping of Images for faster video processing devices. *International Journal of Communication and Computer Technologies (IJCCTS), 11*(1), 10-18. https://doi.org/10.31838/ijccts/11.01.02

[20] Gümüş, A. E., Uyulan, Ç., & Guleken, Z. (2022). Detection of EEG Patterns for Induced Fear Emotion State via EMOTIV EEG Testbench. *Natural and Engineering Sciences, 7*(2), 148-168. https://doi.org/10.28978/nesciences.1159248

[21] Halawani, S. M. (2008). Arabic sign language translation system on mobile devices. *IJCSNS International Journal of Computer Science and Network Security*, *8*(1), 251-256.

[22] Hall, W. C. (2017). What you don't know can hurt you: The risk of language deprivation by impairing sign language development in deaf children. *Maternal and child health journal*, *21*(5), 961-965. https: //doi.org/10.1007/s10995-017-2287-y

[23] Hall, W. C., Levin, L. L., & Anderson, M. L. (2017). Language deprivation syndrome: A possible neurodevelopmental disorder with sociocultural origins. *Social psychiatry and psychiatric epidemiology*, *52*, 761-776. https://doi.org/10.1007/s00127-017-1351-7

[24] Hasan, Z. (2023). Deep Learning for Super Resolution and Applications. *Journal of Galoitica: Journal of Mathematical Structures and Applications*, *8*(2), 34-42. https://doi.org/10.54216/GJMSA.080204

[25] Jonnerby, J., Brezger, A., & Wang, H. (2023). Machine learning based novel architecture implementation for image processing mechanism. *International Journal of Communication and Computer Technologies (IJCCTS), 11*(1), 1-9. https://doi.org/10.31838/ijccts/11.01.01

[26] Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., & Fei-Fei, L. (2014). Large-scale video classification with convolutional neural networks. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition* (pp. 1725-1732). https://doi.org/10.1109/CVPR. 2014.22

[27] Kumar, A., Joshi, P., Bala, A., Sudhakar Patil, P., Jang Bahadur Saini, D. K., & Joshi, K. (2023). Smart Transaction through an ATM Machine using Face Recognition. *Indian Journal of Information Sources and Services*, *13*(2), 7–13. https://doi.org/10.51983/ijiss-2023.13.2.3752

[28] Kurbanazarova, N., Shavkidinova, D., Khaydarov, M., Mukhitdinova, N., Khudoymurodova, K., Toshniyozova, D., ... & Alimova, R. (2024). Development of Speech Recognition in Wireless Mobile Networks for An Intelligent Learning System in Language Education. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), 15*(3), 298-311. https://doi.org/10.58346/JOWUA.2024.I3.020

[29] Lawson, L. (1981). The role of sign in the structure of the deaf community. *Perspectives on British sign language and deafness*, 166-177.

[30] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). Ssd: Single shot multibox detector. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14* (pp. 21-37). Springer International Publishing. https://doi.org/10.1007/978-3-319-46448-0_2

[31] Lombardo, P. A. (2022). *Three generations, no imbeciles: Eugenics, the Supreme Court, and Buck v. Bell*. JHU Press.

[32] Lu, G., Ouyang, W., Xu, D., Zhang, X., Cai, C., & Gao, Z. (2019). DVC: An End-To-End Deep Video Compression Framework. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, 10998-11007. https://doi.org/10.1109/CVPR.2019.01126

[33] Madhan, K., & Shanmugapriya, N. (2024). Efficient Object Detection and Classification Approach Using an Enhanced Moving Object Detection Algorithm in Motion Videos. *Indian Journal of Information Sources and Services, 14*(1), 9–16. https://doi.org/10.51983/ijiss-2024.14.1.3895

[34] Mirdehghan, M. (2010). Persian, Urdu, and Pashto: A comparative orthographic analysis. *Writing Systems Research, 2*(1), 9-23. https://doi.org/10.1093/wsr/wsq005

[35] Miyata, S., Yanou, A., Nakamura, H., & Takehara, S. (2009). Road sign feature extraction and recognition using dynamic image processing. *International Journal of Innovative Computing, Information and Control*, *5*(11), 4105-4113.

[36] Nirmala, M. S. (2023). Behavioural Analysis of Deaf and Mute People Using Gesture Detection. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, *14*(3), 125-142. https://doi.org/10.58346/JOWUA.2023.I3.010

[37] Palmer, J. L., Reynolds, W., & Minor, R. (2012). "You Want What on Your Pizza!?" Videophone and Video-Relay Service as Potential Influences on the Lexical Standardization of American Sign Language. *Sign Language Studies*, *12*(3), 371-397.

[38] Ren, S., He, K., Girshick, R., & Sun, J. (2016). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, *39*(6), 1137-1149. https://doi.org/10.1109/TPAMI.2016.2577031

[39] Rosa, C., Wayky, A. L. N., Jesús, M. V., Carlos, M. A. S., Alcides, M. O., & César, A. F. T. (2024). Integrating Novel Machine Learning for Big Data Analytics and IoT Technology in Intelligent Database Management Systems. *Journal of Internet Services and Information Security*, *14*(1), 206-218. https://doi.org/10.58346/JISIS.2024.I1.014

[40] Samrine, S., & Benali, M. E. *Standard Arabic Sign Language Rules*. Qatar National Library ISBN: 99921-53-41-0.

[41] Sanchez-Ancajima, R. A., Peres, S. M., López-Céspedes, J. A., Saly-Rosas-solano, J. L., Hernández, R. M., & Saavedra-López, M. A. (2022). Gesture Phase Segmentation Dataset: An Extension for Development of Gesture Analysis Models. *Journal of Internet Services and Information Security*, *12*(4), 139-155. https://doi.org/10.58346/JISIS.2022.I4.010

[42] San-Segundo, R., Barra, R., Córdoba, R., D'Haro, L. F., Fernández, F., Ferreiros, J., ... & Pardo, J. M. (2008). Speech to sign language translation system for Spanish. *Speech Communication*, *50*(11-12), 1009-1020. https://doi.org/10.1016/j.specom.2008.02.001

[43] Shaik, S. (2020). A coplanar wave guide fed compact antenna for navigational applications. *National Journal of Antennas and Propagation*, *2*(1), 7-12. https://doi.org/10.31838/NJAP/02.01.02

[44] Stokoe, W. C. (1960). Sign language structure (Studies in Linguistics). Occasional paper, 8.

[45] Stokoe, W. C. (1969). Sign language diglossia. Studies in Linguistics, 21, 27–41

[46] Verkholyak, O., Dvoynikova, A., & Karpov, A. (2021). A Bimodal Approach for Speech Emotion Recognition using Audio and Text. *Journal of Internet Services and Information Security, 11*(1), 80-96. https://doi.org/10.22667/JISIS.2021.02.28.080

[47] Vijay, V., Chaitanya, K., Pittala, C. S., Susmitha, S. S., Tanusha, J., Venkateshwarlu, S. C., & Vallabhuni, R. R. (2022). Physically unclonable functions using two-level finite state machine. *Journal of VLSI circuits and systems*, *4*(01), 33-41. https://doi.org/10.31838/jvcs/04.01.06

[48] Wang, Q., Zhang, L., Bertinetto, L., Hu, W., & Torr, P. H. (2019). Fast online object tracking and segmentation: A unifying approach. In *Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition* (pp. 1328-1338).

[49] Wilcox, S., & Peyton, J. K. (1999). *American Sign Language as a foreign language*. ERIC Clearinghouse. https://eric.ed.gov/?id=ED429464

[50] Yan, K. L. (1993). A comparison of sign language and spoken language rate of articulation and primacy effect.

[51] Zenith Media. 2019. Online video viewing to reach 100 minutes a day in 2021. https://www.zenithmedia.com

[52] Zhao, L., Kipper, K., Schuler, W., Vogler, C., Badler, N., & Palmer, M. (2000). A machine translation system from English to American Sign Language. In *Envisioning Machine Translation in the Information Future: 4th Conference of the Association for Machine Translation in the Americas, AMTA 2000 Cuernavaca, Mexico, October 10–14, 2000 Proceedings 4* (pp. 54-67). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-39965-8_6

Authors Biography

Zaid Saad Bilal, is a seasoned researcher and educator with a Master's degree from Al-Hadithah University for Business and Science in Lebanon. His professional journey includes a tenure at the National School of Electronics and Telecommunications, University of Sfax in Tunisia, where he was associated with the LETI Laboratory at ENIS. With a strong focus on advanced topics in electronics and telecommunications, his interests extend to artificial intelligence. Currently, he imparts knowledge as a secondary school teacher.

Amir Gargouri, received his MSc in Electrical Engineering from the National Engineering School of Sfax, Tunisia, in 2009. Following this, he joined the Control and Energy Management Laboratory (CEM Lab) to pursue his thesis and earned his PhD in 2013. His research interests include neural network implementations, deep learning, machine learning, and Internet of Things (IoT) applications. Currently, he is a member of the Laboratory of Signals, Systems, Artificial Intelligence, and Networks (SM@RTS) and serves as an Associate Professor at the National School of Electronics and Communication, University of Sfax.

Hanaa F Mahmood, received the B.Sc. degree in computer science from the University of Mosul, Iraq, The M.Sc. degree in Artificial Intelligent form the university of Mosul, and the Ph.D. degree in Artificial Intelligent from Loughborough university, U.K. She is a lecturer at department of computer science, college of Education for Pure Sciences, University of Mosul, Iraq. Her research interests include machine learning, deep Learning, computer vision and intelligent systems.

Hassene Mnif, was born in Sfax, Tunisia, in 1975. He received the Engineer and Master Diplomas in electrical engineering from the University of Sfax (ENIS) in 1999 and 2000, respectively, the PhD. degree in electronics from the University of Bordeaux I, France, in 2004 and the HDR degree from the University of Sfax in 2011. He is currently full Professor and Vice-President of the University of Sfax responsible for scientific research. He was the Director of this school between 2014 and 2020, where he has multiple innovative engineering education initiatives. He is a member of the Electronic and Information Technology Laboratory. His research interests include Energy Harvesting, Design of Radio-Frequency Integrated Circuits, Characterization, and compact modeling of both high frequency devices and future emerging technologies like Carbon Nanotube Field Effect Transistor (CNTFET). He also participates in research for real time image and video text extraction and micro mobility systems. He has authored and co-authored about 100 journal publications and conference papers and has gathered significant scientific coordination experience within national and international collaborative research projects. He participated in the organization of several IEEE conferences and workshops. He served as the Tunisia Section treasurer between 2011 and 2013, he is currently the IEEE Tunisia Section Chair-Elect.