IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 10 July 2023, accepted 21 August 2023, date of publication 29 August 2023, date of current version 7 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3309642

RESEARCH ARTICLE

Challenges in Rendering Arabic Text to English Using Machine Translation: A Systematic Literature Review

SHAHAB AHMAD ALMAAYTAH[1] **AND SOLEMAN AWAD ALZOBIDY**[2]
[1]Department of English Language and Humanity, Applied College, King Faisal University, Hofuf 31982, Saudi Arabia
[2]Department of English Language and Translation Studies, College of Sciences and Theoretical Studies, Saudi Electronic University, Ryiadh 11673, Saudi Arabia

Corresponding author: Shahab Ahmad Almaaytah (Salmaaytah@kfu.edu.sa)

The authors extend their appreciation to the Deanship of Scientific Research at King Faisal University for funding this research work through the project number GRANT2, 503.

ABSTRACT The Arabic text can be translated into English using a variety of machine translation techniques. The translation of Arabic text into English still poses a number of challenges in contemporary Arabic. To identify these challenges that encounter while translating Arabic text into English using machine translation, a systematic literature review (SLR) approach is used. The SLR steps—protocol creation, first and final selection, quality assessment, data extraction and synthesis—are used. Nineteen challenges are reported during the SLR process based on fifty-six research papers. The four most important problems are carefully examined, and the possible solutions of other researchers are discussed. Word sense disambiguation, Arabic named entity, rich and complex morphology and low resource are the four critical challenges during rendering Arabic text to English text. Other challenges are also reported in this article.

INDEX TERMS Natural language processing, machine translation, Arabic, systematic literature review, challenges.

I. INTRODUCTION

Machine translation (MT) has advanced for practically all languages in recent years and has become quite important in many applications [1]. As a result, current MT advancements have greatly improved translation quality [2]. Machine translations that are correct and precise are becoming more and more in demand. Finding an adequate and ideal translation, however, is a difficult task in any linguistic context [3], [4]. Different machine translation systems already exist including Al-Mutarjim™ Al-Arabey 3.0[1], Sakhr[2], SYSTRAN[3], Shaheen[4], Bing Translator[5], Babylon[6], and Google Translate[7]. There are several challenges highlighted in various study works [3], [5], [6], [7], such as linguistic mistakes, that indicate the quality of the translation with reference to the Arabic language needs to be further improved. The challenges faced by machine translation can be broadly divided into two groups: technical challenges and linguistic challenges.

A major technical challenge associated with AMT is the lack of datasets and lexical resources that can be utilised as common benchmarks for conducting unified tests. As a matter of truth, academics frequently only collect data relevant to their own fields of study, ignoring a wide range of other fields in the process. They then used these data to try to fix the linguistic problems with Arabic. MT is made more difficult by additional technological difficulties such out-of-vocabulary (OOV), extremely long sentences, and out-of-domain test data [8]. Examples of effective solutions include BPE [9], character-level BPE variation [10], and hybrid approaches [11].

A main linguistic challenge is the nature of the Arabic language as great degree of ambiguity, linguistic complexity, and variety when compared to other languages. Other Arabic features like word order freedom, several diacritization schemas, a wide variety of dialectal variants along

The associate editor coordinating the review of this manuscript and approving it for publication was Agostino Forestiero.

[1]https://al-mutarjim-al-arabey.software.informer.com/3.0/
[2]http://www.sakhr.com/index.php/en/
[3]https://www.systransoft.com/
[4]https://mt.qcri.org/api
[5]https://www.bing.com/translator/
[6]https://translation.babylon-software.com/
[7]https://translate.google.com

social and geographic dimensions present serious linguistic problems to MT [12]. For instance, it has been demonstrated that improving the performance of AMT [13], [14], [15], [16], [17] through pre-processing the Arabic source by morphological segmentation [14], [15], syntactical reordering [16], and hybridization.

A survey on Arabic machine translation was conducted to explore the techniques that employ machine translation available in literature and to encourage researchers to study these techniques. This survey focused on the summarization of major techniques used in machine translation from Arabic into English, and discusses their strengths and weaknesses [4].

Various surveys [4], [18], [19] are conducted in which the topic of Arabic machine translation to other languages was thoroughly examined. All of these earlier analyses and studies came to the conclusion that it is difficult to design a good MT system that satisfies human criteria [4]. However, none of the mentioned survey papers performed a systematic literature review to identify the existing challenges of Arabic Machine Translation (AMT). This research paper adopt the systematic literature review for the identification of various challenges and their possible solutions exists in the literature. Further classification of these challenges is performed. In order to accomplish this, we intend to address the following research questions:

Research Question 1: What are the challenges, as identified in the literature, of rendering Arabic text to English using machine translation?

Research Question 2: What are the proposed solutions and its limitations, as identified in the literature of rendering Arabic text to English using machine translation?

II. ARABIC MACHINE TRANSLATION MECHANISMS

Rule-based, statistical, and neural machine translation are the three basic mechanisms for machine translation [4]. These three approaches are also used in Arabic machine translation.

A. RULE-BASED MACHINE TRANSLATION

In Rule-Based machine translation, a set of linguistic rules are used to translate the source text to the target text [4], [20], [21], [22]. A language specialist usually develops the rules. The use of bilingual or multilingual lexicons, including those for Arabic and other languages, is another component of this strategy. Keep in mind that the lexicons and rule collection were constructed manually.

The main Arabic MT system, known as UniArab, was created by Salem et al. [23] as a global MT system based on a linguistic model. This method's strength lies in its ability to thoroughly examine both the syntax and semantic levels, as indicated in [4], and the fact that it still works for language pairings with little available parallel data, such as low-resource language pairs [22]. However, it is hard to create laws that apply to all languages because doing so would need extensive linguistic expertise and a top-notch dictionary. The latter is more expensive to construct and might not include all the terms. In addition, linguistic specialists are required to develop thorough norms.

B. STATISTICAL MACHINE TRANSLATION

Statistical machine translation [4], [20], [21], [22] uses statistical models from a group of datasets made up of parallel corpora that have been sentence-aligned. For the majority of languages, phrase-based models to SMT provide the most cutting-edge performance [24]. In this approach, initially, the translation model is trained on the bilingual corpus to estimate the probability of the source sentence being a translated version of the target sentence. Then the language model is trained on monolingual corpora which is used to improve the fluency of the output translation. At the end, the maximum probability of product of both the language model and the translation model is computed which gives the most probable sentence in the target language. Phrase-based, syntax-based, and hierarchical phrase-based models are the three types of SMT models [22].

Statistical MT can handle ambiguity by recording phrase-based translations with their frequency of occurrence on a phrase table [4], [20]. The translation result generates through this approach is more fluent and natural. In addition this mechanism is language independent, easy, cheap and fast to build.

C. NEURAL MACHINE TRANSLATION

Neural MT models have been proposed and have outperformed than other mechanisms though these models need a huge amount of parallel data to be trained. Convolutional neural networks (CNN) are used to encode a source text into a continuous vector, and recurrent neural networks (RNN) are used as the decoder to predict the word in the destination language. The concept of the attention mechanism was developed by Bahdanau et al. [2], where the decoder pays attention to input or to any element of the input text. A vector with the same size as the input sequences is produced by calculating attention using each encoder output and the current hidden state.

A neural MT model was created between Arabic text and English by Almahairi et al. [25]. In some studies, neural characteristics for Arabic text is investigated [25], and [26]. The primary difference between neural and statistical MT is that the former has a specific language model while the latter has seen success in a variety of domains [25] in terms of fluency and accuracy. The fundamental issue with NMT, however, is that it necessitates the use of a large parallel corpus [25], which increases the complexity of the training model.

Baniata et al. [27] introduced Transformer-based neural machine translation model for Arabic text. This system used subword units and shared vocabulary within the Arabic dialect to enhance the behavior of the multi-head attention sublayers for the encoder. Experiments are carried out to validate that the proposed mechanism adequately addresses the unknown word issue and boosts the quality of Arabic translation. Self-attention-based Transformer [28] is a stack of layers in a sequence-to-sequence model. To create non-linearity, each layer first uses self-attention to extract information from the entire sentence. This is followed by a point-wise feed-forward network. To enhance the outcomes on the Arabic-English, a Deep Learning architecture based on Convolutional Neural Networks (CNNs) and the transformer model was developed [29]. Experiments on the UN Arabic-English datasets achieved that transformer based model performs better than the most advanced Arabic MT systems.

III. RESEARCH METHODOLOGY

A Systematic Literature Review (SLR) process [30] is used for data collection, which is a structured and defined procedure for finding, evaluating, and analysing published primary studies in order to answer a particular research question. Systematic literature reviews are different from ordinary literature survey because they are explicitly planned and methodically executed. A systematic review may offer a higher level of validity in its conclusions by finding, analysing, and summarising all available information on a particular research subject than may be achievable in any one of the papers it has examined.

To plan the review's strategy, a systematic review protocol is created. These are the key steps in this methodology:

A. SEARCH STRATEGY

The following search strategy steps are used for the construction of search terms.

Step 1: Use the Research Questions for the derivation of major terms, by identifying population, intervention and outcome

Step 2: For these major terms, find the alternative spellings and synonyms

Step 3: Verify the key words in any relevant paper;

Step 4: Use Boolean Operators for conjunction if the database allows, in such a way, to use 'OR' operator for the concatenation of alternative spellings and synonyms whereas 'AND' for the concatenation of major terms.

Process of **step 1** on research questions are as:

RQ1: Machine translation, Challenges, Rendering advertisement Arabic text to English

RQ2: Machine translation, Mechanisms and its limitations, rendering advertisement Arabic text to English

Process of **step 2** on research questions are as:

RQ1:

Machine translation: (MT, Computer translation, Automatic translation, Automatic text conversion)

Challenges: (Challenges, Difficulties, Threats, Complaints, Hardships, Hardness, Problems, Complications, Obstacles)

Rendering advertisement Arabic text: (Arabic translating, Arabic text translation, Arabic broadcasting text)

RQ2:

Machine translation: (MT, Computer translation, Automatic translation, Automatic text conversion)

Solution and limitation: (Mechanisms, Techniques, Methods, Strategies, issues, restrictions)

Rendering advertisement Arabic text: (Arabic translating, Arabic text translation, Arabic broadcasting text)

Process of **step 3** on research questions are as:

Machine translation, Arabic machine translation, Machine translation challenges, advertisement text translation,

Process of **step 4** on research questions are as:

RQ1:

"Machine translation" OR MT OR "Computer translation" OR "Automatic translation" OR "Automatic text conversion" AND Challenges OR Difficulties OR Threats OR Complaints OR Hardships OR Hardness OR Problems OR Complications OR Obstacles AND "Rendering advertisement Arabic text" OR "Arabic translating" OR "Arabic text translation" OR "Arabic broadcasting text" OR "Arabic language"

RQ2:

"Machine translation" OR MT OR "Computer translation" OR "Automatic translation" OR "Automatic text conversion" AND "Solution and limitation" OR Mechanisms OR Techniques OR Methods OR Strategies OR issues OR restrictions AND "Rendering advertisement Arabic text" OR "Arabic translating" OR "Arabic text translation" OR "Arabic broadcasting text" OR "Arabic language"

B. RESOURCES TO BE SEARCHED

The following digital libraries and databases will be searched.

- IEEEXplore (ieeexplore.ieee.org)
- ACM Digital Library (www.acm.org)
- Google Scholar (scholar.google.com)
- ScienceDirect (sciencedirect.com)
- SpringerLink (springerlink.com)

C. SEARCH CONSTRAINTS AND VALIDATION

We are searching for all published relevant literature to our search terms (strings) using the aforementioned resources. We are searching for all relevant literature and hence do not put any date boundaries. A prior trial search was conducted on ScienceDirect (sciencedirect.com) and IEEE Explore (ieeexplore.ieee.org) digital libraries using a set of major terms ("machine translation") AND (challenges OR solutions OR "proposed solutions" OR limitations) AND ("Arabic text" OR "Arabic advertisement text").

We got certain related research papers from our trail search. These related papers will be used for the validation of our search terms (strings).

D. PUBLICATION SELECTION

Publication selection procedure will be carried out by using publication inclusion criteria, publication exclusion criteria and selection of the primary sources. The main purpose of this publication selection procedure is to choose only those search results which are relevant to our research questions. We will only select those research papers/reports/books that are relating to Arabic machine translation. Others research papers/reports/books not related to Arabic machine translation will be will be ignored.

The inclusion criteria are listed as below:

- Research work that describe challenges in Arabic machine translation
- Research work that describes difficulties in translation advertisement text from Arabic to English
- Research work that identifies the Arabic machine translation system
- Research work that shows different solutions for translating Arabic text to English
- Research work that describes limitations during translating Arabic text

Exclusion criteria are used to decide which piece of literature (research papers/reports/books) found by the search term will not be selected for review.

The criteria are listed below:

- Research work that is not relevant to the research questions
- Research work that don't describe Arabic machine translation
- Research work other than machine translation

E. SELECTING PRIMARY SOURCES

Primary sources will be initially selected by analyzing the title, keywords and abstracts of searched literature. This review will exclude/ignore those searched literature which have no relevance to the research questions.

The primary sources chosen during this initial selection process will be checked against the above inclusion/exclusion criteria by reviewing through full text of the research papers.

If any uncertainty occurs regarding the inclusion/exclusion decision, the case will be sent to the secondary reviewer. The process will be checked by the third reviewer.

Inclusion/exclusion decision record regarding each primary source will be maintained properly. This will include the justification whether or not the primary source has been included in the final review.

F. PUBLICATION QUALITY ASSESSMENT

The publication quality assessment is carried out when the final selection of publications is completed. This assessment is performed parallel with data extraction process.

The quality assessment checklist contains the following questions, which will be marked as "Yes" or "No" or "partial" or "NA":

- Is it clearly identified challenges/difficulties during tracking object(s) in the augmented reality environment?
- Is it clearly identified fields/area in the augmented reality tracking environment?
- Is it clear how to solve challenges/difficulties during tracking object(s) in the augmented reality environment?

IV. RESULTS

The publication follow the SLR procedure as given in [31].

A. SEARCHING AND SELECTING SOURCES

Process of searching and selecting the primary research for both research questions are shown in Table 1.

The planned selection process had two parts: an initial selection from the search results of papers that could plausibly satisfy the selection criteria, based on a reading of the title and abstract of the papers; followed by a final selection from the initially selected list of papers that satisfy the selection criteria, based on a reading of the entire papers. In order to reduce the researcher's bias the inter-rater reliability test was performed where the secondary reviewer selected five publications randomly from the list of "Total Searched" and performed the initial and final selection processes. The results were compared with the results produced by the primary author and no disagreements were found. We have identified fifty six (56) research articles for research question 1 and thirty nine (39) research articles for research question 2 as shown in Table 1. We noticed that eighteen (18) research articles are common in both research questions. Data extraction is performed from seventy seven (77) research papers.

From the final selected research papers, the data related to our questions are extracted. This extraction contains the challenges reported in these research articles and their possible solutions.

B. DATA ANALYSIS

In data analysis, the first step to organize quantitative data to group scores or values into frequencies, because frequency analysis is helpful for the treatment of descriptive information. The number of occurrences and percentages of each challenge in Arabic MT are reported using these frequency tables.

In order to answer RQ1, Table 2 shows the list of challenges while translating Arabic advertisement text to English as identified through the SLR. "Word sense disambiguation/Ambiguity" and "Arabic named entity" are the most common challenges while translating advertisement text from Arabic to English. The results also indicates that "Rich and complex morphology" and "Low resource" are the other critical challenges as reported in the literature.

Comparison between main challenges in Arabic machine translation is presented as below

1) WORD SENSE DISAMBIGUATION

The Arabic language has many different types of ambiguity; depending on the situation, many words can have many meanings. For instance, the word 'ذهب' can mean two different things: first, as verb such as "go" and second as a noun "gold". A human using common sense can recognise this ambiguity with ease, but a machine translating the text cannot tell the difference. Instead, MT needs more intricate computation and analysis to accurately determine the meaning; this procedure is known as Word Sense Disambiguation (WSD) [32], [33].

Word Sense Disambiguation (WSD) is the challenge of determining a word's sense (meaning) in a certain context. WSD in Natural Language Processing (NLP) is the process of automatically figuring out a word's meaning by taking the surrounding context into account [34]. One illustration of an ambiguous Arabic term is the word 'خال' (Khal), which can be rendered as either "empty," "imagined," or "battalion."

The three meanings are mixed together because of the Arabic writing system, which is undiacritical and vowelless. In general, Arabic is full with polysemous terms. The astounding repetition of names for human bodily parts in Arabic is an intriguing finding. For instance, when thinking of the term ''head,'' one would picture the neck, nose, eyes, ears, and tongue [35].

Arabic letters can also be ambiguous when attached to morphemes to create ambiguous compound words, so the problem is not restricted to Arabic words. For example, adding the letter 'ب' which is equivalent to the ''b in English letter is changing an atomic word into a compound word in case of ''بالمدرسة'' meaning ''in the school'', ''بالمال'' meaning ''by the money'', ''بالباب'' meaning ''at the door'' and ''بالقلم'' meaning ''using the pen''. This is thus because the letter ''ب'' can have any of the following meanings when used as a prefix: through, in, by, for, and at. Only five of the ten functions that the letter ''ب'' can play when prefixed to various nouns [35].

The biggest hurdle for WSD comes from Arabic texts lacking diacritical marks since they increase the amount of possible meanings for a word and consequently make the work of disambiguation much more challenging. According to the Arabic WordNet (AWN), the word ''صوت'' Sawt, for instance, has 11 senses when written without diacritics, but only two when it is written as ''صوّتَ'' Sawata. The word ''مال,'' which contains seven senses, is another example [36]. The works [32], [37], [38], [39], [40] attempt to handle the challenge of WSD. All of the WSD mechanisms make use of words in a sentence to mutually disambiguate each other. The difference between different techniques can be seen in the type and source of knowledge that the lexical units in a sentence convey. Therefore, all of these methodologies can be categorized as knowledge-based or corpus-based approaches.

2) ARABIC NAMED ENTITY

Identifying and categorising proper names within an open-domain text is the goal of the Named Entity Recognition (NER) task. Because of its intricate morphology, this NLP task is generally recognised to be more challenging for the Arabic language. The use of NER has also been demonstrated to improve the performance of NLP tasks like machine translation, information retrieval, and question answering. The lack of capitalization, the extensive lexical variety, and the inconsistent manner in which Arabic names are written all contribute to the difficulty of named entities [41]. For either meaning-based translation or phoneme-based transliteration to yield a trustworthy translation result, named entities must be handled correctly [42]. These issues were solved by Shaalan [41], who developed a productive and robust Arabic named entity recognition system.

Ameuret al. [19] has proposed a translation attempt for the MT between Arabic and English utilising an attention-based encoder-decoder. The outcomes demonstrated the effectiveness of their strategy in compared to some earlier studies. Recently, to improve NE transliteration, Alkhatib and Shaalan [42] used a hybrid deep learning approach based on CNN, followed by Bi-LSTM and CRF. Their findings on the corpora ANERcorp and Kalimat demonstrate the effectiveness of their model for Arabic-English machine transliteration, producing cutting-edge outcomes.

3) RICH AND COMPLEX MORPHOLOGY

The morphology of Arabic is rich and complicated, very different from that of English. By reducing the quantity of the source vocabulary and enhancing the accuracy of word alignments, pre- and post-processing of Arabic source text using NLP technologies like segmentation and tokenization has been proven to increase the performance of MT. For example, a single Arabic word ''ومنزله'' translated to ''and his home'' in English is formed by prepending the prefix ''و'' (''and'') to the base lexeme ''منزل'' (''home'') and the prefix ''ه'' (''him''). Numerous publications have addressed this difficult aspect of Arabic [14], [25], [43], [44], [45], [46], [47].

Compared to prior AMT research works, there are few MT investigations exists that address this issue. An attention-based neural machine translation model between Arabic and English is proposed by Almahairi et al. [25] that had the highest level of accuracy.

More recently, Garcia-Martinez et al. [48] examined the results of decomposing the target words of an Arabic-French factored NMT model employing linguistic preprocessing.

Their model predicted the lemma as well as the combination of the following elements at the time of decoding: the POS tag, the tense, the gender, the number, the person, and the case information. To replicate low-resource and rich resource behaviors, the model underwent training utilizing small or large parallel training datasets, accordingly. Both their Factored and conventional NMT architectures used BPE segmentation. The factored NMT models performed significantly better, according to their evaluation results on various test sets.

4) LOW RESOURCE

Arabic is a language with low resources. Since learning depends on the volume of training data, MT performs better for high resource languages than for low resource languages [8]. The millions of data that MT systems train on demonstrate a direct correlation with accuracy. Triangular MT [49], back-translation [49], fine-tuning [50], multilingual NMT [50], and zero-shot transfer [51] are only a few of the strategies that have been researched for handling low-resource languages.

Some works that try to handle low-resource languages are worth mentioning. For instance, to enhance the translation performance of low-resource couples, the authors in [49] presented a novel triangular training architecture. The translation models of the rich language are jointly optimised using a unified bidi-rectional Expectation-Maximization (EM) algorithm in this design, which uses a rich language as the intermediate latent variable. On the MultiUN and IWSLT2012 datasets, their strategy dramatically improves the translation quality of rare languages like Arabic.

TABLE 1. Data sources and search results.

Resource	Research Question 1			Research Question 2		
	Total	Primary Selection	Final Section	Total	Primary Selection	Final Section
IEEEXplore	263	73	14	238	70	11
ACM Digital Library	152	42	07	137	40	6
Google Scholar	213	59	13	210	61	10
ScienceDirect	182	50	08	146	43	7
SpringerLink	243	67	14	219	64	9
Total Search	1053	291	**56**	951	278	**39**

TABLE 2. Challenges of AMT identified through SLR (in descending order of frequency).

S.No.	Challenges	Frequency (n= 56)	Percentage
1.	**Word sense disambiguation**	15	26.8
2.	**Arabic named entity**	10	17.9
3.	**Rich / complex morphology**	7	12.5
4.	**Low resource**	6	10.7
5.	**Arabic vocalization**	4	7.1
6.	**Rare, unknown and spelling mistakes**	4	7.1
7.	**Dialectal variation**	3	5.4
8.	**Free word order**	2	3.6
9.	**Out-of-domain**	2	3.6
10.	**Out of vocabulary and vocabulary size**	2	3.6
11.	**Word alignment**	2	3.6
12.	**Computational overhead**	2	3.6
13.	**Sentence length**	1	1.8
14.	**Accuracy**	1	1.8
15.	**Fluency**	1	1.8
16.	**Performance**	1	1.8
17.	**Computation time**	1	1.8
18.	**Data sparsity**	1	1.8
19.	**Noisy data**	1	1.8

A character-based hybrid NMT model that mixes RNN and CNN networks was presented by Almansor and Al-Ani [52]. They developed their model using only 90K sentence pairs from a very small subset of the TED parallel corpora, including the IWSLT 2016 Arabic-English corpus. Comparing the openNMT word-based NMT model to Arabic-English, they noted considerable gains.

Abid [53] developed an NMT model as a way to enhance MT models without using any outside sources of data. To achieve solid baselines, the author bootstrapped already-existing parallel sentences and combined them with multilingual training. They produced a benchmark dataset in four languages—Egyptian, Levantine, Arabic, and English—and made it available to the public for free. According to the findings of the studies, a multilingual dialect model with MSA and bootstrapping produces the best outcomes.

AraBench [54] system is developed that is used to perform machine translation evaluation suite for Arabic dialect to English. They used a variety of training settings, including fine-tuning, back-translation, and data augmentation. The assessment suite opens a wide range of research frontiers on low-resource machine translation, such as Arabic dialect translation. Both the dialectal system and the assessment suite are accessible to the general public for academic research. Data augmentation techniques are investigated for synthesizing dialectal Arabic-English code-switching (CS) text [55].

The quality of the generated sentences are accessed through human evaluation and evaluate the effectiveness of data augmentation on machine translation (MT), automatic speech recognition (ASR), and speech translation (ST) tasks. Results showed that data augmentation achieved much improvement in perplexity, relative improvement on WER for ASR task, BLEU points on MT task, and BLEU points on ST over a baseline trained on available data without augmentation. Transformation mechanism is proposed to augment data during training that extended the distribution of authentic data [56]. In particular, it uses augmented data as auxiliary tasks to provide new contexts when the target prefix is not helpful for the next word prediction. This enhances the encoder and steadily increases its contribution by forcing the Grammatical Error Correction (GEC) model to pay more attention to the text representations of the encoder during decoding. The impact of these approach was investigated using the Transformer-based for low-resource Arabic GEC. Experimental results showed that the proposed approach outperformed the baseline, the most common data augmentation methods, and classical synthetic data approaches.

V. DISCUSSION

In this study, the challenges encountered during translating Arabic text to English are categorized based on its reporting in research papers. The most significant study areas that have been examined in the studies that have been made will be highlighted in this part.

Machine translation is an NP-hard problem that aims to produce accurate translations. Although technology has advanced significantly over the last decade, more effort still has to be done in its development. As a result, even after post-processing, the original text meaning is still not precise. Despite numerous attempts to improve word alignment in AMT systems, it still falls short of the mark and exhibits divergence in terms of fluency. As a result, to improve fluency in the target language, an effective alignment procedure is needed after translation. One method for doing this is through an induced alignment during the decoding step, as was done for English.

In terms of computation, it is essential to find a way to accelerate neural network training at both the computation and memory levels, especially for rich morphological words, to enable the use of much larger vocabularies for both the source and target languages, long sentences, and low-frequent words. With existing MT approaches, OOV, uncommon and unknown words, ambiguous words, and misspelt words are difficult to manage. One of the fundamental issue while working with the Arabic language is really its intricate and rich morphology, which differs greatly from that of Indo-European languages (such as English). These experiments demonstrated that morphological segmentation and Arabic tokenization can significantly enhance the results of overall translation.

VI. CONCLUSION AND FUTURE WORK

A Systematic Literature Review (SLR) is used to identify the challenges during translating Arabic text to English. The four main challenges encountered while rendering Arabic text into English are word sense disambiguation, Arabic named entities, rich and complex morphology, and low resource availability. The four most important problems are carefully examined, and other researchers have suggested solutions. Other challenges includes Arabic vocalization, unknown and spelling mistakes, dialectal variation, free word order, out-of-domain, out of vocabulary, word alignment, computational overhead, sentence length, accuracy, fluency, performance, computation time, data sparsity and noisy data. Some of these challenges depend on other challenges during translation process. Although technology has advanced significantly in the decade, much effort needs to be done in order to obtain high accuracy and good fluency.

Various machine translation techniques to translate from Arabic to English text are examined in this research work. We intend to create a new Arabic to English machine translation system in the future that will handle the main challenges discussed in this research paper.

Conflicts of Interest: The authors declare no conflict of interest.

ACKNOWLEDGMENT

(Shahab Ahmad Almaaytah and Soleman Awad Alzobidy contributed equally to this work.)

REFERENCES

[1] N. T. Alsohybe, N. A. Dahan, and F. M. Ba-Alwi, "Machine-translation history and evolution: Survey for Arabic-English translations," 2017, *arXiv:1709.04685*.
[2] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," 2014, *arXiv:1409.0473*.
[3] L. Alkhawaja, H. Ibrahim, F. Ghnaim, and S. Awwad, "Neural machine translation: Fine-grained evaluation of Google translate output for English-to-Arabic translation," *Int. J. English Linguistics*, vol. 10, no. 4, p. 43, Apr. 2020.
[4] A. Alqudsi, N. Omar, and K. Shaker, "Arabic machine translation: A survey," *Artif. Intell. Rev.*, vol. 42, no. 4, pp. 549–572, Dec. 2014.
[5] O. Jabak, "Assessment of Arabic-English translation produced by Google translate," *Int. J. Linguistics, Literature Transl.*, vol. 2019, pp. 1–12, Jul. 2019.
[6] M. H. Al-Khresheh and S. A. Almaaytah, "English proverbs into Arabic through machine translation," *Int. J. Appl. Linguistics English Literature*, vol. 7, no. 5, pp. 158–166, 2018.
[7] M. E. Marouani, T. Boudaa, and N. Enneya, "Statistical error analysis of machine translation: The case of Arabic," *Computación Sistemas*, vol. 24, no. 3, pp. 1053–1061, Sep. 2020.
[8] P. Koehn and R. Knowles, "Six challenges for neural machine translation," 2017, *arXiv:1706.03872*.
[9] R. Sennrich, B. Haddow, and A. Birch, "Neural machine translation of rare words with subword units," 2015, *arXiv:1508.07909*.
[10] W. Ling, I. Trancoso, C. Dyer, and A. W. Black, "Character-based neural machine translation," 2015, *arXiv:1511.04586*.
[11] M.-T. Luong and C. D. Manning, "Achieving open vocabulary neural machine translation with hybrid word-character models," 2016, *arXiv:1604.00788*.
[12] N. Y. Habash, "Introduction to Arabic natural language processing," *Synth. Lectures Hum. Lang. Technol.*, vol. 3, no. 1, pp. 1–187, 2010.
[13] A. Alqudsi, N. Omar, and K. Shaker, "A hybrid rules and statistical method for Arabic to English machine translation," in *Proc. 2nd Int. Conf. Comput. Appl. Inf. Secur. (ICCAIS)*, May 2019, pp. 1–7.
[14] N. Habash and F. Sadat, "Arabic preprocessing schemes for statistical machine translation," in *Proc. Hum. Lang. Technol. Conf. NAACL, Companion Volume, Short Papers*, 2006, pp. 49–52.
[15] M. Oudah, A. Almahairi, and N. Habash, "The impact of preprocessing on Arabic-English statistical and neural machine translation," 2019, *arXiv:1906.11751*.

[16] M. Ellouze, W. Neifar, and L. H. Belguith, "Word alignment applied on English-Arabic parallel corpus," in *Proc. LPKM*, 2018, pp. 1–9.

[17] M. E. Marouani, T. Boudaa, and N. Enneya, "Incorporation of linguistic features in machine translation evaluation of Arabic," in *Proc. Int. Conf. Big Data, Cloud Appl.* Cham, Switzerland: Springer, 2018, pp. 1–18.

[18] H. M. Elsherif and T. R. Soomro, "Perspectives of Arabic machine translation," *J. Eng. Sci. Technol.*, vol. 12, no. 9, pp. 2315–2332, 2017.

[19] M. S. H. Ameur, F. Meziane, and A. Guessoum, "Arabic machine translation: A survey of the latest trends and challenges," *Comput. Sci. Rev.*, vol. 38, Nov. 2020, Art. no. 100305.

[20] M. Alkhatib and K. Shaalan, "The key challenges for Arabic machine translation," in *Intelligent Natural Language Processing: Trends and Applications*. Cham, Switzerland: Springer, 2018, pp. 139–156.

[21] B. Kituku, L. Muchemi, and W. Nganga, "A review on machine translation approaches," *Indonesian J. Elect. Eng. Comput. Sci.*, vol. 1, no. 1, pp. 182–190, 2016.

[22] L. Han, "Machine translation evaluation resources and methods: A survey," 2016, *arXiv:1605.04515*.

[23] Y. Salem, A. Hensman, and B. Nolan, "Implementing Arabic-to-English machine translation using the role and reference grammar linguistic model," in *Proc. 8th Annu. Conf. Inf. Technol. Telecommun.*, Galway, Irelan, 2008, pp. 103–110.

[24] F. J. Och and H. Ney, "The alignment template approach to statistical machine translation," *Comput. Linguistics*, vol. 30, no. 4, pp. 417–449, 2004.

[25] A. Almahairi, K. Cho, N. Habash, and A. Courville, "First result on Arabic neural machine translation," 2016, *arXiv:1606.02680*.

[26] A. Alrajeh, "A recipe for Arabic-English neural machine translation," 2018, *arXiv:1808.06116*.

[27] L. H. Baniata, I. K. Ampomah, and S. Park, "A transformer-based neural machine translation model for Arabic dialects that utilizes subword units," *Sensors*, vol. 21, no. 19, p. 6509, 2021.

[28] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, and I. Polosukhin, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–11.

[29] A. I. E. Farouk, "Transformer model and convolutional neural networks (CNNs) for Arabic to English machine translation," in *Proc. 5th Int. Conf. Big Data Internet Things*. Cham, Switzerland: Springer, 2022, pp. 399–410.

[30] B. Kitchenham, O. P. Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering-a systematic literature review," *Inf. Softw. Technol.*, vol. 51, no. 1, pp. 7–15, Jan. 2009.

[31] L. Shamseer, D. Moher, M. Clarke, D. Ghersi, A. Liberati, M. Petticrew, P. Shekelle, and L. A. Stewart, "Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: Elaboration and explanation," *Bmj*, vol. 349, pp. 1–12, Jan. 2015.

[32] M. Hadni, S. E. A. Ouatik, and A. Lachkar, "Word sense disambiguation for Arabic text categorization," *Int. Arab J. Inf. Technol.*, vol. 13, no. 1A, pp. 215–222, 2016.

[33] S. A.-A. Mussa and S. Tiun, "Word sense disambiguation on English translation of holy Quran," *Bull. Electr. Eng. Informat.*, vol. 4, no. 3, pp. 241–247, Sep. 2015.

[34] R. Navigli, "Word sense disambiguation: A survey," *ACM Comput. Surv.*, vol. 41, no. 2, p. 1–69, 2009.

[35] E. Abuelyaman, L. Rahmatallah, W. Mukhtar, and M. Elagabani, "Machine translation of Arabic language: Challenges and keys," in *Proc. 5th Int. Conf. Intell. Syst., Modeling Simulation*, Jan. 2014, pp. 111–116.

[36] N. Bouhriz, F. Benabbou, and E. Habib, "Word sense disambiguation approach for Arabic text," *Int. J. Adv. Comput. Sci. Appl.*, vol. 7, no. 4, pp. 1–10, 2016.

[37] W. A. Gale, K. W. Church, and D. Yarowsky, "A method for disambiguating word senses in a large corpus," *Comput. Humanities*, vol. 26, nos. 5–6, pp. 415–439, Dec. 1992.

[38] I. Dagan and A. Itai, "Word sense disambiguation using a second language monolingual corpus," *Comput. Linguistics*, vol. 20, no. 4, pp. 563–596, 1994.

[39] F. Ahmed and A. Nürnberger, "Arabic/English word translation disambiguation using parallel corpora and matching schemes," in *Proc. 12th Annu. Conf. Eur. Assoc. Mach. Transl.*, 2008, pp. 6–11.

[40] D. Yarowsky, "Unsupervised word sense disambiguation rivaling supervised methods," in *Proc. 33rd Annu. Meeting Assoc. Comput. Linguistics*, 1995, pp. 1522–1531.

[41] K. Shaalan, "A survey of Arabic named entity recognition and classification," *Comput. Linguistics*, vol. 40, no. 2, pp. 469–510, Jun. 2014.

[42] M. Alkhatib and K. Shaalan, "Boosting Arabic named entity recognition transliteration with deep learning," in *Proc. 33rd Int. Flairs Conf.*, 2020, pp. 484–487.

[43] Y.-S. Lee, K. Papineni, S. Roukos, O. Emam, and H. Hassan, "Language model based Arabic word segmentation," in *Proc. 41st Annu. Meeting Assoc. Comput. Linguistics*, 2003, pp. 399–406.

[44] A. Hatem and N. Omar, "Syntactic reordering for Arabic-English phrase-based machine translation," in *Database Theory and Application, Bio-Science and Bio-Technology*. Cham, Switzerland: Springer, 2010, pp. 198–206.

[45] E. A. Mohammed and M. J. A. Aziz, "English to Arabic machine translation based on reordring algorithm," *J. Comput. Sci.*, vol. 7, no. 1, pp. 120–128, Jan. 2011.

[46] W. Antoun, F. Baly, and H. Hajj, "AraBERT: Transformer-based model for Arabic language understanding," 2020, *arXiv:2003.00104*.

[47] N. Zalmout and N. Habash, "Optimizing tokenization choice for machine translation across multiple target languages," *Prague Bull. Math. Linguistics*, vol. 108, no. 1, pp. 257–269, Jun. 2017.

[48] M. García-Martínez, W. Aransa, F. Bougares, and L. Barrault, "Addressing data sparsity for neural machine translation between morphologically rich languages," *Mach. Transl.*, vol. 34, no. 1, pp. 1–20, Apr. 2020.

[49] S. Ren, W. Chen, S. Liu, M. Li, M. Zhou, and S. Ma, "Triangular architecture for rare language translation," 2018, *arXiv:1805.04813*.

[50] P. Shapiro and K. Duh, "Comparing pipelined and integrated approaches to dialectal Arabic neural machine translation," in *Proc. 6th Workshop NLP Similar Lang., Varieties Dialects*, 2019, pp. 214–222.

[51] W. Lan, Y. Chen, W. Xu, and A. Ritter, "GigaBERT: Zero-shot transfer learning from English to Arabic," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, 2020, pp. 1–5.

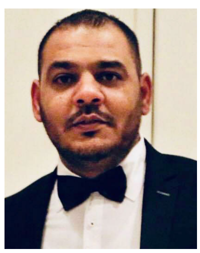
[52] E. H. Almansor and A. Al-Ani, "A hybrid neural machine translation technique for translating low resource languages," in *Proc. Int. Conf. Mach. Learn. Data Mining Pattern Recognit.* Cham, Switzerland: Springer, 2018, pp. 347–356.

[53] W. Abid, "The SADID evaluation datasets for low-resource spoken language machine translation of Arabic dialects," in *Proc. 28th Int. Conf. Comput. Linguistics*, 2020, pp. 6030–6043.

[54] H. Sajjad, A. Abdelali, N. Durrani, and F. Dalvi, "AraBench: Benchmarking dialectal Arabic-English machine translation," in *Proc. 28th Int. Conf. Comput. Linguistics*, 2020, pp. 5094–5107.

[55] I. Hamed, N. Habash, S. Abdennadher, and N. T. Vu, "Investigating lexical replacements for Arabic-English code-switched data augmentation," 2022, *arXiv:2205.12649*.

[56] A. Solyman, M. Zappatore, W. Zhenyu, Z. Mahmoud, A. Alfatemi, A. O. Ibrahim, and L. A. Gabralla, "Optimizing the impact of data augmentation for low-resource grammatical error correction," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 35, no. 6, Jun. 2023, Art. no. 101572.



SHAHAB AHMAD ALMAAYTAH received the Ph.D. degree in linguistics and translation. He is currently a Coordinator with the English Language Department, King Faisal University. His research interests include machine translation (MT), machine learning (ML), natural language processing, linguistics, and translation studies.

SOLEMAN AWAD ALZOBIDY received the Ph.D. degree in linguistics from Mysore University, India, in 2015. He is currently an Assistant Professor in linguistics with Saudi Electronic University, where he has been a Faculty Member, since 2016. His research interests include computational linguistics, machine translation, and translation studies.