Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

ARABIC TEXT STEGANOGRAPHY: LITERATURE SURVEY

Nadia Abdul-rahman[1], Salwa Shakir[2]

[1] University of Al-Qadisiyah/College of Computer Science and Information Technology/Department of Computer Science/Iraq
*Corresponding author e-mail: com21.post10@qu.edu.iq

https://doi.org/10.46649/fjiece.v3.2.1a.10.5.2024

Abstract. *The use of strong encryption and hiding techniques is essential for secure digital communication. The aim of this article is to give a thorough study of advanced steganography methods designed especially for embedding information in Arabic text, which may be sensitive in nature and needs to be transmitted securely. This review is focused on four main criteria; embedding rate, invisibility, robustness against detection, and text quality. We conducted a comprehensive survey of text steganography and examined its distinct characteristics, including the nature of the script, accents, and linguistic rules. Several researchers have utilized these distinct characteristics of Arabic writing for developing a number of techniques for embedding information. This article gives a detailed analysis of the strengths, weaknesses, possible applications, and major challenges. The conclusions of this review can help researchers and practitioners to select an ideal method according to their own needs regarding storage capacity, security, and text quality. Moreover, the article underscores the need for giving priority to Arabic text steganography which has become highly important as the use of hidden communication increases in Arabic-speaking communities, while also preserving the cultural and linguistic integrity.*

Keywords: *Arabic text , Steganography, Text steganography, Data hiding, Arabic characteristics*

1. INTRODUCTION

Most of the modern civilizations are technologically dependent. A lot of data is produced whether someone uses the internet for shopping, browsing, mobile banking, or saving bank account details. Sometimes specific information—personal appointments, national secrets, secret mission codes, and banking credentials—is classified as confidential. As such, sending various data kinds calls for keeping security and confidentiality. Information is deliberately buried inside a media to make its existence known in steganography[1]. This method captures private data in text, images, audio, and video files. The concept originates in Greek, most precisely in the words "Steganos" and "Graptos," which both imply "covered writing" or "hidden writing" [2]. Even although both steganography and cryptography need data concealing, they have different properties. Secret information concealed in the covering media is kept out of the hands of criminals by steganography. It is the goal to keep undesired access by concealing information beneath a cover object [3]. With text steganography, a text file serves as the cover to hide personal information. Because text steganography carries less superfluous information than speech, video, and pictures, it is crucial [4].Four characteristics are necessary for the steganographic approach to be successfully applied: imperceptibility, information embedding capacity, durability, and security. Steganographic technology implementation would be ideal since it could preserve all of these characteristics [5].

2. BASIC OVERVIEW OF STEGANOGRAPHY

Steganography refers to the presence of concealed or disguised written content. Steganography is employed to achieve clandestine communication by hiding a message from an external observer. Steganography is the practice of hiding the existence of a secret communication, without making the message unreadable. In contrast, cryptography is focused on making a message unreadable to unauthorized individuals. While steganography and cryptography are separate ideas, they do exhibit significant parallels. Steganography, like cryptography, involves covert communication and can be seen as secretive writing methods [2].The practice of steganography often entails embedding a concealed message into a specific medium of transmission, referred to as the carrier. The concealed communication is intricately incorporated within the host to create the steganographic medium. A stego key can be used to encrypt the concealed message and/or as a randomization seed for the stego algorithm [6]. To summarize, the following equation shows how the stego_medium is calculated:

Stego_medium = hidden_ message + carrier+ stego key (1)

Steganography is classified into two distinct categories: technical steganography and linguistic steganography. Figure 1 elaborates on the categorization of the steganographic method [7].

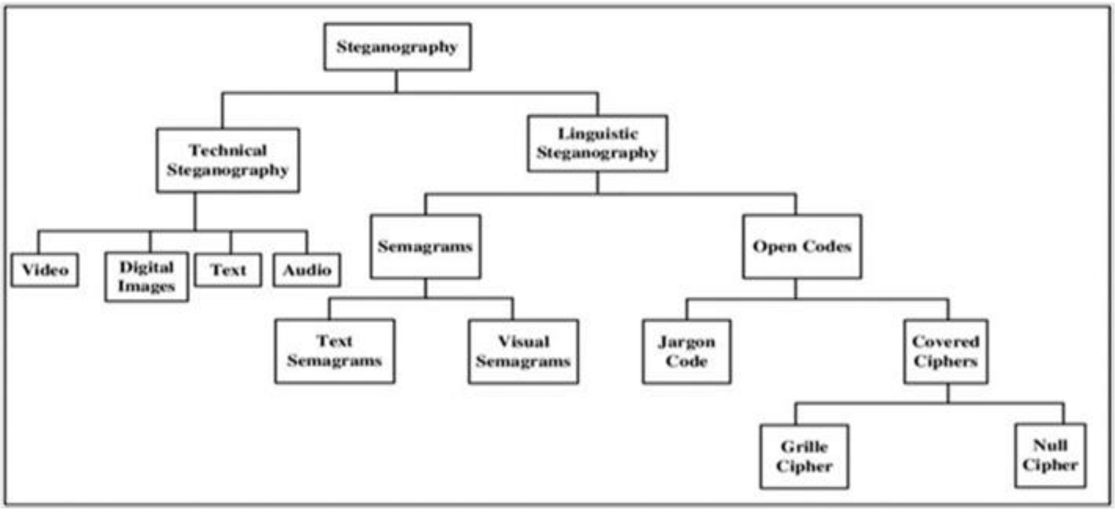


Figure 1. Classification of steganography techniques [7]

2.1 Linguistic steganography

The message is concealed within the carrier using methods that are not immediately apparent, and it is further classified as either semagrams or open codes. Summarized as follows:

1- Semagrams conceal information by the utilization of symbols or signs. A visual semagram employs seemingly innocuous or commonplace tangible objects to communicate a message such as sketches, the arrangement of items on a desk or website, or the positioning of a flag on a balcony (reminiscent of Deep Throat from the Watergate scandal). A text semagram conceals a message by altering the visual characteristics of the original text, such as making slight adjustments to the font size or type, inserting more gaps, or incorporating distinct embellishments in the letters or handwriting [8].

2- Open codes conceal messages within legitimate carrier messages in ways that are not readily apparent to a casual observer. In communication, concealed messages are known as covert messages, while carriers are known as overt messages. Jargon codes and covered ciphers are part of this category.

3-As its name implies, jargon code consists of language that is understood by a particular group but not by others. Warchalking (symbols for wireless network signals and underground terminology) or innocent conversations convey special meaning because the facts that are known only to the speaker are common jargon codes. Jargon codes can be divided into cue codes, which contain prearranged phrases [9].

4-In covered ciphers, the message is hidden openly in the carrier medium, so the message can be recovered without knowledge of how it was hidden. With a grille cipher, the carrier message is hidden in the template while the words within the openings reveal the hidden message. A null cipher hides the message according to some prearranged set of rules, such as “read every fifth word” or “look at the third character in every word” [3,5,7,8,9].

2.2 Technical steganography

A steganography system is secured when observers cannot identify the cover object from the steganogram object. This type of steganography uses different types of cover. Different media types can be used in cover objects including images, videos, audio, text, and networks [8]. Figure 2 shows the technical steganography types.

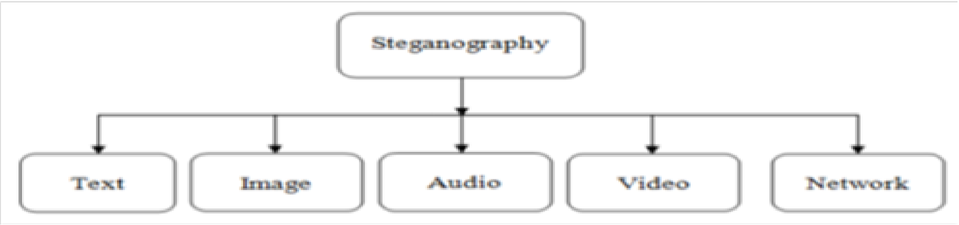


Figure 2. Types of technical steganography [8]

2.2.1 Image Steganography

A stenography type called image steganography uses an image carrier file as its carrier. Here, the image files can be in different formats (e.g., JPEG, GIF, BMP, and PNG) that are utilized to cover the sensitive message. A combination of image format steganographies, spatial domain steganography, and adaptive steganography are used for creating image steganography. In recent years, a data mapping technique has been used as a technique for image steganography that minimizes the amount of information in each pixel changes. There are four important bits of the cover pixel that are mapped to the four hidden data bits [10].

2.2.2 Video Steganography

MPEG, AVI, and MP4 video files can be a combination of images and sounds to conceal a great deal of sensitive information. Embedding sensitive information in every frame of a video, which is applied to image steganography, can also be applied to video steganography using computed tomography (CT) scanning. A technique called Least Significant Bit (LSB) uses the pixel values of the most significant bit of the frame to embed the secret bit within the frame, as does bit-plane complexity segmentation (BPCS) which embeds the secret bit inside the MPEG frame [11] .

2.2.3 Audio Steganography

It is possible to protect secret messages by shifting the binary sequence of a file that saves digital sound (for instance, MP3 or WAV) by utilizing a file that saves the digital. Modern steganography methods consist of the use of error diffusion to change the LSBs. Using inaudible frequencies is an additional method that can be used to conceal secret messages [12].

2.2.4 Network Steganography

This type of steganography embeds the secret bits in a single network protocol. Throughout this article, a list of network protocols can be found that have been used for covert purposes. These protocols are Transmission Control Protocol (TCP), Protocol Data Unit (PDU), User Datagram Protocol (UDP), Internet Control Message Protocol (ICMP), and Internet Protocol (IP). A network steganography system can provide a high level of security and robustness [13].

2.2.5 Text Steganography

In comparison with a picture or a sound file, text steganography is considered more difficult because there is less redundant information in a text file [1]. Text documents frequently exhibit a structure that closely resembles what is visible, while the structures of audio, picture, and video cover media are very different in nature, enabling the concealment of information beyond text without any noticeable alterations. Text steganography makes it easier to conceal information from texts, simplifies communication, allows transmission of more information, and reduces printing costs. Computer systems are involved in this process, as a result of technological development, there has also been a wide variety of uses for hiding information in text. Electronic documents, web pages, and electronic text in general are some of the most important technologies in this field. Text steganography, based on the concealment of information in the cover text, falls into three basic categories [3, 4], explained and summarized as follows:

I. Random and Statistical Generation

Cover texts are generated randomly and statistically using statistical properties. This method is based on character and word sequences. The hidden information is embedded in a random sequence of characters, with the information to be hidden implicitly embedded in that information. If any threat intercepts the message, the sequence of characters is interpreted as random. By leveraging the same statistical numbers as the original word, it is also possible to generate "words" (without lexical value) from the statistical properties of word length and letter frequency. Several bits and dictionary words have been mapped between lexical items and a codebook that hides the information. In addition, codebooks are used to disguise the information [12].

II. Linguistic Steganography

In linguistic steganography, secret information is concealed by using the language of words or other linguistic features that can be incorporated into words to conceal it. The linguistic method can be divided into two categories [14]:

1- Syntactic methods: as a syntactic method, punctuation is used to determine the way a sentence is composed.
2- Synonymic methods: by using synonyms instead of interactive words in the dictionary (by some carrier file words), where the hidden bits have been passed between the words (in the dictionary).

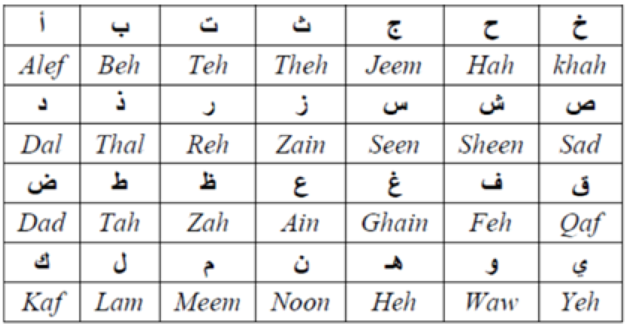
III. Format-Based Steganography

Physical documents are formatted in such a way as to hide secret information from this group. There are several format-based methods used in text steganography including deliberate misspellings, font resizing, and space injection. It may seem that these format-based methods are successful at tricking the human eye, but they miss out on tricking computer systems or extending the length of the Stego text. In addition, these methods are not robust enough to defend against text-retyping attacks. Feature-based and word-rule-based methods are divided into the format-based category [13, 14].

3. ARABIC TEXT STEGANOGRAPHY

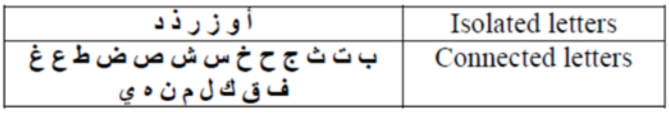
3.1 Characteristics of the Arabic Language

People from all over the world use Arabic as a language. There are 29 letters in the Arabic alphabet as shown in Figure 3. The Arabic alphabet is traditionally thought to contain 28 letters, but many linguists believed that the letters Hamza and Aleph should be treated as one letter (Unicode.org, 2023). The Arabic alphabet is based on a series of connected letters, whereas the English alphabet uses separate letters. It is possible to divide the Arabic letters into two groups (Unicode.org, 2023): i) connected letters, and ii) isolated ones, as shown in Figure 4. The first group contains 22 letters, each of which can take four different forms depending on where it appears in the word, as shown in Figure 5, while in the second group, it consists of only some letters including the final, the middle, and the initial. There is a connection between each letter of a word and its following letter. For example, the connected letter (Beh) can be written in four different ways, as shown in Figure 5.



أ	ب	ت	ث	ج	ح	خ
Alef	*Beh*	*Teh*	*Theh*	*Jeem*	*Hah*	*khah*
د	ذ	ر	ز	س	ش	ص
Dal	*Thal*	*Reh*	*Zain*	*Seen*	*Sheen*	*Sad*
ض	ط	ظ	ع	غ	ف	ق
Dad	*Tah*	*Zah*	*Ain*	*Ghain*	*Feh*	*Qaf*
ك	ل	م	ن	هـ	و	ي
Kaf	*Lam*	*Meem*	*Noon*	*Heh*	*Waw*	*Yeh*

Figure 3. Arabic letters



أ و ز ر ذ د	Isolated letters
ب ت ث ج ح خ س ش ص ض ط ع غ ف ق ك ل م ن ه ي	Connected letters

Figure 4. Types of Arabic letters

General unicode	Contextual forms				Name
	Isolated	Final	Middle	Initial	
0628 ب	FE8F ب	FE90 ـب	FE92 ـبـ	FE91 بـ	Beh

Figure 5 . Unicode Arabic lettrer of Beh

Thal, Dal, Reh, Zain, Waw, and Alef are the remaining six letters of the Arabic alphabet, and these letters have only two variants: as first letters in words, and as last letters in words (Figure 5). No matter how many letters are within a word, these characters cannot be merged with each other. After these six letters, each letter following them must be written in an initial form, unconnected from the previous letter. Moreover, the first letter of a word has the same properties as the rest of the words in the word. Right-side letters consist of those first letters, but they are not joined together. To create a steganographic method, this set of letters can be used as a hidden key. Unicode-based algorithms have many drawbacks that have previously been discussed. We can avoid some drawbacks by using this approach.

3.2 Challenges in Arabic Text Steganography Compared to English text

While text steganography presents challenges across languages due to the limited redundancy and strict structure of text files, concealing information in Arabic texts poses unique complexities compared to English texts, explained as follows [5]:

A. Structural Complexity

- The Arabic script is a form of handwriting that employs cursive, where each letter is linked to the letter before it and takes on a distinct form according to its position within a word (e.g., initial, medial, final, or isolated). The complex configuration of this structure and the abundance of letter forms and connections present challenges for steganography techniques. Conversely, the discrete nature of the letters in the English alphabet improves the efficacy of steganographic methods [1].

-A variety of vowel marks (harakat) and other markings, including shadda (gemination marks) and tanwin (nunation marks), are present in the Arabic language [26]. By manipulating or capitalizing on these diacritics, one can obscure the existence or nonexistence of data. Due to these advancements, steganography methods must cautiously manage additional complexities in order to preserve the integrity of the text and evade detection. The English diacritical system is considerably more straightforward, encompassing only accents and other symbols present in foreign names and loanwords [18].

-The Arabic script is unique in that it contains kashidas (horizontal extension characters) and irregular letters, both of which pose challenges and opportunities for steganography in a way that is impracticable when compared to the English language. Numerous methodologies can be employed to preserve the authentic visual representation of text, such as kashidas insertion and dot manipulation [31].

B. Linguistic differences

- Steganography techniques used in Arabic require careful attention to grammatical features and traditions, namely distinguishing between the letters representing the sun and the moon. Because of these subtle differences in language, the way words that have the definite article "al" as a prefix are spelled and spoken may vary. The linguistic nuances mentioned here bring about further intricacies that are absent in the English language [35].

-Arabic is highly inflected, involving a complex morphological system involving affixes, patterns and roots.This can pose difficulties for certain steganography methods that rely on word-level modifications or substitutions, as they must maintain the integrity of the inflectional system to avoid detection [15].

C. Cultural and Scriptural Considerations

- Some Arabic text steganography techniques specifically target sacred texts like the Quran, which requires extra sensitivity and care due to the revered status of text in Islamic culture. Modifying or tampering with such texts, even for steganographic purposes, may raise ethical and cultural concerns that are generally absent when working with non-religious English texts .
- Arabic calligraphy and the Arabic script hold a respected position in Islamic and Arab cultures, which may impact the acceptability and perception of certain steganography techniques that can visibly modify the appearance of texts [36].

3.3 Arabic text steganography techniques

We can use several features of the Arabic language to conceal secret texts. Modern steganography methods evaluate Arabic language texts for steganography in the following manner [15]:

A. *Using the Arabic Astrology*

Farah R. Shareef [16] presented a novel Arabic text steganography method based on Arabic astrology. This method aims to significantly improve data hiding capacity and security within Arabic text. This method divides Arabic letters into four astrological groups: fiery, watery, earthy, and airy. Then it combines these groups with special characters like kashida and various Unicode characters like zero-width no joiner (ZWNJ), hair right-to-left (RTL), and zero-width no-break space (ZWS) to conceal messages. This method can embed two secret bits into the Arabic text for each character used, enhancing the traditional steganography techniques that usually embed only a single bit. Implementing this method further complicates the task of detecting and examining steganography, while also augmenting the quantity of concealed information contained within a specified word count. An evaluation was undertaken to ascertain the efficacy of diverse Arabic text coverings, some of which concealed messages were printed in languages other than Arabic, Persian, and English. It enhanced the level of concealment that could be achieved while decreasing the need for covers.

B. *Using lunar and Solar diacritics*

Rawaa et al. [17] have developed an innovative algorithm that leverages the unique attributes of the Arabic script—such as the solar and lunar characters, in conjunction with their respective diacritics—to obscure sensitive information within Arabic texts. By manipulating the diacritical marks that follow the prefixes "Al" or "ال" this methodology entails the concealment of confidential information within Arabic terms. The diacritical marks are modified in accordance with the classification of the third letter as solar or lunar. Therefore, due to their subtle characteristics, these alterations are imperceptible to casual observers. By implementing this approach, it becomes possible to maintain the text's initial visual appeal and legibility. Additional advantages of the algorithm include its robustness against discovery and its reliance on inherent linguistic characteristics of Arabic. In terms of capability, security, and concealment, the method was compared to other Arabic text-hiding strategies during an evaluation. Experiments indicate that the cover text is virtually undetectable. In determining the capacity of a cover text, the number of syllables beginning with the prefix "Al." is utilized.

C. *Using Diacritics*

Nooruldeen and Mohammed [18] put forth a methodical approach to obfuscate delicate content within Arabic scriptures, with particular emphasis on the Holy Quran. A technique is implemented to obscure two bits of information through the identification and modification of the Unicode values of designated Arabic characters (–ز–ىٰ–د–إ). Various bit values may be concealed by the algorithm, contingent upon the combination of letters and diacritics employed. The algorithm alters the Unicode based on the search for terms that meet the stated criteria within the chosen cover text (Sura Al-Falaq) to encode the hidden message. To effectively conceal information in the text, this method takes advantage of the unique characteristics of Arabic script and diacritics. By utilizing this approach, it is feasible to enhance the concealment of information while maintaining the integrity of the sacred text. The researchers utilized a varied assortment of chapters (surahs) from the Holy Quran as the fundamental text for their study. The outcomes clearly demonstrated a notable level of embedding capabilities. Although no changes have been made to the Holy Quran text, such as adding, removing, or replacing letters or diacritical marks, the authors claim that the proposed technique maintains its integrity.

Malak and Adnan [19] introduced a two-tier security method to conceal important Arabic text data, specifically on devices with restricted processing capabilities. This approach integrates the concealment of Arabic text within another medium, known as steganography, with a cryptography process that is efficient and not resource-intensive. The initial layer employs lightweight encryption algorithms such as AES, DES, and IDEA to secure sensitive Arabic textual material. Steganography is employed on the secondary level to conceal the encrypted information within the vowel markings (diacritics) of an Arabic text cover medium. The encryption of the data is controlled by the presence or absence of specific diacritics. DES exhibits superior data concealment capabilities when compared to AES and IDEA, rendering it well-suited for systems with restricted storage capacity. The use of the Peak Signal-to-Noise Ratio (PSNR) measure in quantitative and visual analysis revealed that AES outperformed DES and IDEA in terms of security. The performance of the system is categorized into three phases according to the duration of the secret message. Messages exceeding 16 letters in length are not recommended owing to their significantly reduced capacity. In order to bolster the security of sensitive data on resource-constrained devices, it is more efficacious to employ both LWC and Arabic text steganography in contrast to relying exclusively on a single variant of steganography.

D. *Using Unicode*

Allah Ditta et al. [20] devised a method that integrated steganography and cryptography in order to bolster the integrity of the cover text and improve the efficacy of data concealment. Sophisticated data is concealed utilizing linguistic steganography techniques and Arabic text carriers.The implementation of Zero-Width-Joiners (ZWJ) and Zero-Width-Characters (ZWC), which are Unicode characters, is necessary for this purpose.In order to augment the confidentiality of the covert communications, the data was initially encrypted and then subsequently concealed via bit inversion. The cover text effectively conceals and communicates any concealed information while retaining its aesthetic appeal through the use of this methodology.The evaluation algorithm exhibits respectable levels of coverage, robustness, security, and moderate capacity, according to the results of the simulations. Consequently, it can be regarded as an exceptionally valuable tool in guaranteeing the preservation of data and facilitating secure communication. The approach being evaluated successfully obscures a maximum of two sensitive details per character in cover Arabic text. This methodology substantially improves the ability to obscure data by a factor of 200 when compared to previous approaches. An additional layer of security was achieved by employing bit inversion as a means to obscure the confidential information. To reduce suspicion, it is essential to ensure that the stego-text cannot be visually distinguished from the cover text so as to preserve the visual integrity of the cover text.

Norah et al. [21] introduced a novel approach that sought to improve the performance and reliability of steganography specifically designed for Arabic text. In order to represent concealed bits, Arabic characters are converted from their original Unicode format to a form that is contextualized according to their particular usage context. The incorporation of supplementary characters, including Kashida, Zero-Width Joiners (ZWJ), Medium Mathematical Spaces (MMSPs), and Zero-Width Non-Joiners (ZWNJ), guarantees the maintenance of data integrity while augmenting operation efficiency. By an average of more than fifty percent, the efficacy of the implemented strategy was considerably greater than that of comparable approaches. Moreover, the system demonstrates resilience in the face of diverse types of electronic text manipulation, such as duplication, cloning, and replication, while simultaneously upholding an exceptionally stringent level of security. Moreover, the algorithm could potentially be implemented in additional Arabic dialects, including Urdu and Farsi. This is achievable as a result of the widespread adoption of Unicode characters, which constitute the foundation of the majority of writing systems. The researchers conducted the investigations utilizing nine Arabic texts of varying dimensions. In comparison to alternative methodologies, the proposed approach demonstrated a significant enhancement. Furthermore, an elevation was implemented in the security ratio. The utilization of the suggested approach in conjunction with the remaining benchmark techniques did not produce any statistically significant results when it came to copying, pasting, or formatting texts. Placing reliance on the internal structure of letters or non-printing characters made them vulnerable to retyping, while the implementation of printing/OCR operations weakened their resilience.

Adeel et al. [22] focused on the "non-printable characters" technique for Arabic text steganography. This technique hides secret messages by inserting non-printable Unicode characters like zero-width joiner and zero-width non-joiner into an Arabic cover text. When taking an Arabic cover text, that will be used to hide the secret message after that, it convert the secret message (in Arabic or any other language) into a binary-bit stream of 0s and 1s. Inserting a non-printable Unicode characters such as zero-width joiner (U+200D) and zero-width non-joiner (U+200C) into the Arabic cover text to represent the 0s and 1s of the secret binary message. The resulting text with embedded non-printable characters is the stego-text containing the hidden message. To retrieve the secret message, the recipient removes the non-printable characters from the stego-text, which reveals the original binary stream that can be converted back to the plaintext message. There was a thorough evaluation of various Arabic text steganography techniques, identification of their limitations, and suggestions for improving them. The strategy of non-printable characters shown superior performance in both capacity and keeping the appearance of the cover text.

E. Using Deep learning

Omer and Seyed [24] devised an innovative method to conceal sensitive data in the Arabic language by combining artificial intelligence (AI) and deep learning. The researchers investigated the feasibility of utilizing Long Short-Term Memory (LSTM) neural networks to encode categorized data in order to generate genuine Arabic literary compositions. The research successfully enhanced the participants' ability to recall hidden information by employing structural elements commonly seen in Arabic poetry, such as rhyme and meter, resulting in a 45% improvement. The method was done with linguistic precision. The Baudot Code method employs the replacement of words with letters to safeguard sensitive data.As a result, the poetry that is created has the ability to hide a greater amount of information. Implementing this approach not only improves the ability to hide data, but also guarantees that the Arabic text maintains its aesthetic and grammatical integrity. Consequently, alterations made to steganography techniques are not observable or noticeable by unintentional recipients.

F. Using Dotted letters

Abdennour et al. [25] introduced an approach for embedding secret data within Arabic text using a novel steganography technique based on the dotted features of Arabic letters. To hide confidential information, the Kashida extension character and the position and number of points in Arabic letters are used.The approach employs novel hidden bit patterns to obfuscate three bits of sensitive information each letter. Compared to traditional methods, this leads to improved capacity performance. The data is securely buried using a mechanism that strategically inserts Kashida characters within Arabic letters, following specific circumstances based on hidden bit patterns. Experimental assessments revealed that the proposed technology surpassed existing steganography methods. Therefore, it may be possible in the future to enhance the security of written Arabic communication. Several websites incorporated text elements of different dimensions in order to hide confidential data. The capacity ratios displayed promising indications.

The proposed steganography technique, which takes advantage of the unique features of dotted Arabic characters, seems to be highly efficient at concealing a significant amount of sensitive data within texts.

Ahlam et al. [26] developed a technique to hide English communications under cover texts using Arabic symbols, so making them difficult to understand. The hidden message is compressed using the T-5BE 5-bit encoding method to enhance the chances of effectively concealing it within another message. Arabic characters with dots are used to convert distinct dot configurations in the compressed message as part of the secret message. The program ensures accurate retrieval of the hidden message by analyzing the dots in each word and finding the concealed components within the encoded text. Furthermore, an Arabic semantic lexicon is utilized to enhance the steganographic cover text and address any potential discrepancies that may occur during the embedding process. The suggested concealment strategy has proven to be beneficial by highlighting its significant storage capacity and excellent precision in hiding. The researchers achieved exceptional outcomes as a direct consequence of their experimental approaches. An invisible message was cleverly incorporated into the narrative text about the historical atmosphere of Baghdad using eight carefully chosen phrases and a total of 53 characters.The adept concealment of the sensitive data greatly reduced the potential for weaknesses or ambiguity. By employing an algorithm that modified the weights of the letters in the cover text, the hidden communication successfully avoided detection through visual and electronic examination.

G. *Using Kashida & Diacritics*

Afra & Mohamed [27] have developed a performance-based tool for Arabic text steganography, which encrypts and conceals secret messages.The tool utilises two main techniques for concealing the secret message, including i) Kashida technique: This involves adding one or two kashida characters (Arabic extension characters) to certain Arabic letters in the cover text to represent the binary bits (0 or 1) of the secret message. ii) Diacritics technique: This uses the different diacritical marks in Arabic text to represent the binary bits. They used the most common diacritic (fatha) to represent bit 1, and the other seven diacritics represents bit 0. The tool has three modules: Embedding module to hide the secret message using the above techniques; performance calculation module to evaluate the stego-object (cover text with hidden message) based on capacity, security, and robustness metrics; decoding module to extract the hidden message. The user can input their preferences for the most important performance parameter, and the tool will select the stego-object with the highest overall performance score for transmitting the secret message securely. The researchers conducted experiments with different cover texts and user-selected performance parameters. The hiding capacity achieved varied based on the cover text and the selected performance parameter.

H. *Using MSCUKT and BloodGroup method*

Farah R. [28] presented a secure communication framework that conceals encrypted data through the integration of text steganography techniques into Arabic text cover messages, as well as cryptography employing the Diffie-Hellman key exchange and the AES encryption algorithm. The cryptographic component establishes a shared secret key between the sender and receiver via the Diffie-Hellman key exchange prior to employing AES to encrypt the secret message. Two unique steganographic methodologies are evaluated in the steganography segment.The first method involves inserting a message covertly into Arabic text. Subtly integrating an Arabic text containing a classified message pertaining to a particular blood type constitutes the second approach. Following the encrypted message's transmission, the steganographic text that was utilized to obfuscate it is also transmitted. By employing the mutually distributed Diffie-Hellman key, the recipient gains the capability to decrypt and revers the encoded data. The researchers examined four separate instances, each concerning confidential communications in Persian, English, and Arabic scripts, which varied in length. The vastly superior processing speed of the MSCUKAT method over the BloodGroup technique was evident from the analytical results. The increased concealment capability and reduced processing time of this technology are advantageous for real-time encrypted messaging applications.

I. *Using Noorani and Darkness letters*

Farah R. [29] developed a steganographic method that utilized Arabic letters referred to as "Darkness letters" and "Noorani letters," along with Kashida (an extension of Arabic letters) and Unicode space characters, to hide hidden bits within an Arabic text overlay. Depending on the type of Arabic letter, it defines five types of embedding scenarios, such as Noorani, Darkness, Isolated-Noor, and Isolated-Dark.In each case, it uses different combinations of inserting Kashida (two types), Unicode space characters (Narrow-No-Break, Hair space, etc.), or leaving the letter unmodified to represent different 2-bit combinations like '00', '11', '01', '10'. It also utilises existing spaces in the cover text to embed bits by replacing them with Zero-Width Joiner, Unicode spaces like PRE-space, thin space, etc. The embedding process checks each letter, determines its type, and then embeds two secret bits based on the defined scenarios for that letter type. The researchers evaluated their method using various secret messages of different sizes (English, Arabic, and Persian) and Arabic text covers. Their method has a high capacity to conceal large secret messages in Arabic text covers, and it is adaptable to any language.

J. *Using Traid-bit method*

Roshidi et al. [30] presented a novel method that integrates three Arabic text steganography techniques: i) Kashida technique (extension): Adds or does not add an extension to Arabic words to represent secret bits '1' or '0; ii) shifting point technique: replaces the dot or point in Arabic letters vertically or horizontally to represent '1' or '0'; and iii) sharp-edges technique: utilizes the number of sharp edges in Arabic letters to conceal bits '0', '1', '01', etc. By integrating these three techniques, the triad-bit method aims to overcome the limitations of each individual technique and provide better embedding performance. The paper evaluates the proposed method based on metrics like capability of stego text size, usage ratio of cover text characters, usability ratio, fitness performance of stego keys, and running time. The experimental results showed that the triad-bit method can improve the embedding capacity while utilizing fewer characters from the cover text compared to using any single technique alone. The method showed improved size capability, with the stego text size being lower than expected and using a relatively small percentage of cover text characters including 8% for Kashida, 9% for Shifting Point, and 4% for Sharp-edges. However, it had high usability.

K. *Using Kashida and Small Space Characters method*

Ahmed et al. [31] proposed a new algorithm that aims to improve the hiding capacity (the amount of secret data that can be embedded) while preserving the quality of the cover text as much as possible, where it utilizes the Arabic extension character "Kashida" to hide 1 bit per character by adding or not adding the Kashida after letters that can take it. It also utilizes three different types of small space characters (thin space, hair space, and six-preem space) in addition to normal spaces between words. The combination of these three small spaces allows for hiding three bits per word spacing. Embeds data by: i) inserting or not inserting Kashida after eligible letters based on the secret bit; ii) inserting specific small space combinations after each word based on the next 3 secret bits. The researchers have demonstrated the effectiveness of their algorithm through experimental results and comparisons with existing methods. They tested the algorithm on different cover text lengths (155, 600, 1100, and 1602 characters). The proposed algorithm achieved higher hiding capacity compared to other algorithms. The Jaro-Winkler similarity score between the cover text and stegotext is relatively high.

L. Using Character Property Method (CPM)

Nuur et al. [32] proposed a new Arabic text steganography method. It utilizes three features of Arabic characters—sharp edges, dots, and calligraphic typographical proportions—as potential places to hide secret message bits within each Arabic character. It employs a maximum positioning algorithm that sorts the Arabic characters based on the descending order of their number of character property elements (sharp edges, dots, etc.). The algorithm then sequentially concatenates the secret bits into the characters according to this order. It uses a secret shared key between sender and receiver to create a unique random number mapping table. For added security, the positioning of the concealed bits within the cover text is randomized. They combined the CPM with a biometric multifactor authentication (BMA) scheme using fingerprint and heartbeat sensors to prevent unauthorized access and address fake identity issues that could lead to modification of the concealed message. As a result of the system, the cover and the stego have the same appearance, due to their high hiding capability and high perceptual transparency. The cover text can be modified via OCR, or retyped, and the file format can be modified.

M. Using Pseudo-spaces with Kashida

Safia and Adnan [33] presented a method for embedding secret data within Arabic text by strategically placing pseudo-space characters. It is through the use of these pseudo-spaces that the secret information is concealed in the text. Various groups of lengths are used to categorize the secret data. These groups range from two bits to eight bits in length. In order to determine how many concealment spaces are required, it converts them into decimal numbers. During transmission over insecure channels, the message was intended to be concealed from unauthorized parties. Cover texts should contain sufficient redundant data so that secret information can be substituted for it. According to the method, this is very important. Consequently, unintended recipients will not be able to see the modifications. Additionally, the work discussed the importance of destroying covers after use in order to prevent message reconstruction by potential attackers. This method achieved high capacity, perfect perceptual transparency, and no visible differences between the original and stego-text.

N. Using Letter Shaping

A.F. Al Azzawi [34] proposed a method for hiding information in Arabic Unicode text documents. It utilizes the different shapes that Arabic letters can take depending on their position in a word (beginning, middle, end, or isolated). Each Arabic letter has an unshaped code in Unicode that is rendered into the proper shaped form by software. To hide data, it replaces the original unshaped Unicode letter with a different Unicode code that renders the same letter shape but indicates a different position (beginning, middle, etc.). This allows hiding 1 bit per letter. The cover text is segmented into layers based on part-of-speech tagging, with all words with the same POS tag in one layer. A random sequence of layer indices is generated as a stego key. The message is hidden by traversing through the words in the randomly selected layer sequence, reshaping letters to hide the message bits. No characters are inserted, so the stego text has the same length and visual appearance as the cover text. The researchers compared this method to other existing Arabic text steganography techniques, such as Kashida, pseudo-space, diacritics, and so on. The method provides a good balance of capacity, imperceptibility, and security for Arabic text steganography without compromising the quality or size of the cover text.

O. Using Statistical Text Steganography

Nujud and Lamia [35] proposed a method that uses Markov chains (MC) and Huffman coding (HC), which generates an MC model from an Arabic text corpus to capture Arabic text's statistical properties. It encodes the secret message using Huffman coding to compress it into a binary sequence. It divides the binary sequence into blocks and convert each block to a decimal number. For each decimal number, it uses the MC model to generate a sequence of states (words) such that the transitions between states for encoding that decimal number. The final stego-text is the concatenation of all the state sequences generated for each decimal block. The encoding process essentially generates text that mimics the statistical properties of natural Arabic text while hiding the secret message in the sequences of words used. The method also explores its capacity by determining the length of the stego-text based on parameters such as block size and the number of outgoing states in the MC model. It was found that increasing block size and the number of outgoing states up to a threshold can increase the embedding capacity. Using Huffman coding on the secret message can also improves capacity.

P. Using Diacritics and Unicode

Huda et al. [36] proposed two novel steganography algorithms, specifically using the Holy Quran as the cover text. Taking the advantage of the existence of sun letters (Arabic letters that cause the preceding definite article "al" to assimilate) and moon letters (that don't cause assimilation) in Arabic grammar rules. Exploiting the Arabic diacritical marks (Harakat) which represent vowel sounds and are compulsory in texts like the Quran using a specific Unicode representation of the isolated Arabic letter Alif (ا) to mark the locations of hidden bits without perceptibly altering the cover text. Algorithm 1 hides 1 bit per word that are starting with "al" followed by a sun or moon letter, by modifying the Unicode of the isolated alif. Algorithm 2 hides 2 bits per eligible word, by combining the sun/moon letter rule with checking the presence of the diacritical mark Fathah after the "al" prefix. The researchers evaluated the algorithms using seven surahs from the Holy Quran as cover texts. The hiding capacity obtained was not extremely high, is reasonable given the unique constraints of using the Holy Quran text, which cannot be modified.

Table 1 summarizes the above works including the advantages and limitations

Ref.	Authors And Years	The Proposed Methods	Advantages	Limitations
[16]	Farah R. Shareef 2023	*Arabic Astrology*	Higher hiding capacity, better invisibility, complicated decryption detection, the ability to work with any language, the ability to handle large secret messages.	Increased Stego text size and limited test evaluation, more extensive testing on diverse datasets may be needed to comprehensively validate performance claims,User experience is moderately difficult, especially for non-technical users.
[17]	Authors 2023	*lunar and Solar diacritics*	The system conceals data in a robust, imperceptible, and size-preserving manner for Arabic text files,The user experience is relatively easy and straightforward.	Limited capacity, a lack of generality beyond Arabic text, the semi-automated nature of the process.
[18]	Authors 2023	*Special letters and Diacritics*	Provides a secure, efficient, and culturally sensitive approach to steganography within Arabic text,	Limited Applicability,cover text limitation and capacity constraints.

			particularly in the context of the Holy Quran,The user experience is relatively easy and user-friendly, especially for non-technical users.	
[20]	Authors 2022	*Unicode*	The method enhances data hiding capacity and security in cover text, making it a valuable tool for secure communication and information concealment. Users will find the interface to be intuitive and easy to use.	Using Unicode characters will result in an increase in the size of the steg-text file, so certain applications may experience a decrease in storage and transmission efficiency.
[24]	Authors 2022	*Artificial Intelligence*	The steganography of Arabic text improves embedding capabilities, security, and linguistic quality compared with previous word-level steganography. It is relatively simple to use.	The method is specifically designed for the Arabic language and may not be easily extensible to other languages with different linguistic structures and properties.
[21]	Authors 2022	*Unicode*	The method offers a balance between capacity, security, and readability, making it a valuable approach for secure individual uses of text steganography in Arabic, The user experience is relatively easy and straightforward. The interface is user-friendly.	The approach is designed exclusively for steganographic Arabic text and lacks generalizability to other languages or formats.Consequently, its restricted versatility renders it unsuitable for use in specific circumstances.
[25]	Authors 2021	*Dotted letters*	This approach offers a highly effective and resilient method of hiding sensitive information while communicating. The user experience is relatively straightforward.	This approach is specifically tailored for Arabic text and may not be suitable for other text formats or languages. As a result of this constraint, its versatility between platforms and programming languages is restricted.
[27]	Afra & Mohamed 2021	*Kashida and Diacritics*	High capacity, customizable security, automated workflow, and Arabic language-specificity,The user experience is relatively easy and straightforward.	In the Arabic language, concealing information through the use of kashida and diacritics is a unique practice. It is not practical to implement this technology in its entirety across various dialects and text formats.
[22]	Authors 2021	*Unicode*	The high capacity, complete imperceptibility, flexibility across texts, and conceptual simplicity, the user experience is easy and seamless	This approach avoids the use of encryption or obfuscation techniques and instead exclusively depends on the concealment of data. The decryption of any encryption is unnecessary in the event that the plaintext communication is detected.
[26]	Authors 2021	*Dotted letters*	The method offers a combination of security, capacity, and accuracy in concealing secret messages within text. the user experience user-friendly and efficient.	May not be directly applicable to other languages, limiting its versatility in a multilingual context.
[28]	Farah 2021	*MSCUKT and BloodGroup method*	Enhanced security through multiple layers, real-time secure communication, good hiding capacity in Arabic texts, and a user-friendly and efficient user experience.	Intentionally created to hide information under Arabic text covers.This makes the method language-dependent and may not be directly applicable to non-Arabic text covers.

[19]	Malak & Adnan 2021	*Diacritics*	Enhanced security, good capacity, and imperceptibility , the user experience is relatively easy and straightforward	It is specifically designed for Arabic text covers that use diacritics,it cannot be directly applied to non-Arabic text formats.
[29]	Farah R 2021	*Noorani and Darkness letter*	Increased hiding capacity, reduced visibility of modifications, and utilised the unique properties of Arabic script,the user experience is clear and straightforward	The method may not perform optimally when the cover text contains non-Arabic words, numbers, or symbols, as the rules won't apply to them.
[30]	Authors 2021	*Traid-bit*	Increased capacity, flexibility in technique selection, and better usability	The method's tailored design for Arabic text restricts its direct adaptability to other languages or text formats. In comparison to employing a single methodology, the embedding procedure is more intricate for the user.
[31]	Authors 2020	*Kashida and Small Space Characters*	High hiding capacity and maintaining text quality,provided a good user experience.	It has limitations in terms of applicability and robustness in certain scenarios.
[32]	Authors 2020	*Character Property Method*	Increased capacity, enhanced security, enhanced optimal positioning of secret bits within the cover text, a high level of transparency, and robustness against OCR, the user experience is clear and straightforward	Specifically designed for Arabic text steganography, which may limit its applicability to other languages or text formats.
[23]	Authors 2020	*Unicode*	The method can be applied to languages related to Arabic, such as Persian and Urdu. In addition, the user experience is relatively easy	The placement of the concealed elements is not randomised.
[33]	Safia & Adnan 2020	*Pseudo-spaces with Kashida*	Increased capacity and improved security are applied not only to Arabic text but also to languages similar to Arabic, such as Urdu and Persian, the user experience is relatively easy for genuine users.	The placement of the concealed elements is not randomised.
[34]	A.F. Al Azzawi 2019	*Letter Shaping*	Strikes a good balance between capacity, imperceptibility, and maintaining original text quality and naturalness, the user experience is clear and straightforward	Lower embedding capacity
[35]	Nujud and Lamia 2019	*Statistical method*	High embedding capacity, statistically similar to cover, the user experience is complex.	Lack of semantic meaning
[36]	Authors 2019	*Diacritics and Unicode*	High Capacity, Robustness, and Imperceptibility ,the user experience is easy for both the sender and receiver to use.	Relies on using Holy Quran, as cover text ,may restrict the applicability of the technique to a broader range of texts or contexts.

4-ARABIC TEXT STEGANOGRAPHY KEY PERFORMANCE CRITERIA

The key performance criteria for steganography considered in this study are capacity, invisibility, robustness, and security. To sustain high effectiveness, steganography algorithms must adhere to these criteria, particularly invisibility. We define the capacity of a cover medium as the amount of secret information that can be conceal. Next, invisibility denotes the eavesdropper's difficulty in detecting secret data. Eavesdroppers cannot detect or alter secret communication which eventually leads to ensuring its security. Finally, the robustness of a steganography method denotes its resistance to change or destruction by the stego medium.

Table 2 presents the details of each the most related article that considered the main key performance criteria. Upon detection of the secret data, the steganography communication breaks down. Each method, along with its key performance, is depicted in Figure 6.

Table 2: The evaluation criteria for each previous Arabic text steganography method

No.	Title Paper	Year of publication	Evaluation Criteria			
			Capacity	Invisibility	Robustness	Security
1	[16]	2023	/			/
2	[17]	2023	/	/		/
3	[18]	2023	/			
4	[19]	2021	/	/	/	
5	[20]	2022	/			
6	[21]	2022	/			/
7	[22]	2021	/			
8	[23]	2020	/			/
9	[24]	2022	/			
10	[25]	2021	/			
11	[26]	2021	/			
12	[27]	2021	/		/	/
13	[28]	2021	/			
14	[29]	2021	/			
15	[30]	2021	/			
16	[31]	2020	/	/		
17	[32]	2020	/	/		/
18	[33]	2020	/	/		
19	[34]	2019	/			
20	[35]	2019	/			
21	[36]	2019	/			

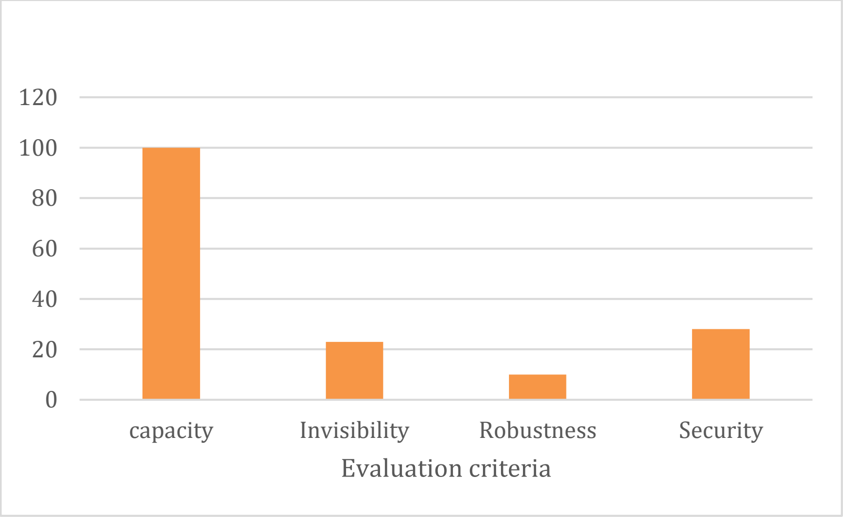


Fig. 6. Arabic text steganography methods key performance evaluation

Notably, all the surveyed articles focused on capacity, and 23% of them looked at invisibility measurement performance. While 10 % of the research work improved the robustness key performance. Lastely, 28 % of them looked into security measurement performance.

5-CONCLUSIONS

This survey study has explored the state-of-the-art Arabic text steganography techniques. The unique characteristics of the Arabic script including its cursive writing style, diacritical marks, and linguistic rules have inspired researchers to develop innovative approaches for covert communication. The reviewed methods leverage features like kashidas, dotted letters, diacritics, Unicode manipulation, and deep learning to effectively conceal data within Arabic texts. While achieving promising results in terms of capacity, security, and imperceptibility, the techniques face trade-offs and limitations specific to the Arabic language. Future research should aim to further improve the robustness and cross-language applicability of different steganography methods, ensuring seamless integration with modern communication channels. Additionally, addressing cultural considerations and ethical implications will be crucial for the responsible adoption of the Arabic text steganography in diverse contexts. Overall, this domain holds significant potential for enabling secure information exchange while preserving the linguistic and cultural integrity of the Arabic script.

REFERENCES

[1] A. Mohammed Alshamsi, S. Matar Albaloushi, M. Younes Alkhoori, H. Ahmed Almheiri, and N. Ababneh, “A review of arabic text steganography,” *ACM Int. Conf. Proceeding Ser.*, pp. 6–11, 2021, doi: 10.1145/3456146.3456148.

[2] N. A. A. Mustafa, “Text hiding in text using invisible character,” *Int. J. Electr. Comput. Eng.*, vol. 10, no. 4, pp. 3550–3557, 2020, doi: 10.11591/ijece.v10i4.pp3550-3557.

[3] R. Thabit, N. I. Udzir, S. Md Yasin, A. Asmawi, N. A. Roslan, and R. Din, “A comparative analysis of arabic text steganography,” *Appl. Sci.*, vol. 11, no. 15, 2021, doi: 10.3390/app11156851.

[4] E. Sai, P. Chunduru, and N. Rao Koppolu, “Hiding in the Plain Text: A Critical Analysis of Whitespace Steganography,” *Int. Res. J. Eng. Technol.*, p. 4129, 2008, [Online]. Available: www.irjet.net

[5] M. T. Ahvanooey, Q. Li, J. Hou, A. R. Rajput, and Y. Chen, “Modern text hiding, text steganalysis, and applications: A comparative analysis,” *Entropy*, vol. 21, no. 4, 2019, doi: 10.3390/e21040355.

[6] E. Saad and A. Azzam, “A New Algorithm to Hide a Secret Text in another Text,” vol. 31, pp. 1–18.

[7] O. S. Adewale, O. K. Boyinbode, and S. E. Adekunle, “Visual semagram: An enhanced technique for confidentiality requirement of electronic voting system,” *Int. J. Comput. Netw. Inf. Secur.*, vol. 12, no. 4, pp. 51–59, 2020, doi: 10.5815/ijcnis.2020.04.05.

[8] S. Chaudhary, M. Dave, and A. Sanghi, “Review of Linguistic Text Steganographic Methods,” *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 4, no. 7, pp. 377–381, 2016.

[9] L. Liu, L. Tang, and W. Zheng, “Lossless Image Steganography Based on Invertible Neural Networks,” *Entropy*, vol. 24, no. 12, 2022, doi: 10.3390/e24121762.

[10] R. J. Mstafa and K. M. Elleithy, “Compressed and raw video steganography techniques: a comprehensive survey and analysis,” *Multimed. Tools Appl.*, vol. 76, no. 20, pp. 21749–21786, 2017, doi: 10.1007/s11042-016-4055-1.

[11] M. M. Taher, A. R. B. H. J. Ahmad, R. S. Hameed, and S. S. Mokri, “a Literature Review of Various Steganography Methods,” *J. Theor. Appl. Inf. Technol.*, vol. 100, no. 5, pp. 1412–1427, 2022.

[12] M. H. Muhammad, H. S. Hussain, R. Din, H. Samad, and S. Utama, “Review on feature-based method performance in text steganography,” *Bull. Electr. Eng. Informatics*, vol. 10, no. 1, pp. 427–433, 2021, doi: 10.11591/eei.v10i1.2508.

[13] A. Ali, W. Mohamed, and M. Sayed, “A Survey Paper of Information Hiding by Using Steganography Techniques,” *Kafrelsheikh J. Inf. Sci.*, vol. 3, no. 2, pp. 1–23, 2022, doi: 10.21608/kjis.2022.280155.

[14] A. M. Alhusban and J. Q. O. Alnihoud, “A Meliorated Kashida Based Approach for Arabic Text Steganography,” *Int. J. Comput. Sci. Inf. Technol.*, vol. 9, no. 2, pp. 99–112, 2017, doi: 10.5121/ijcsit.2017.9209.

[15] A. A., F. Ridzuan, and S. Ali, “Text Steganography using Extensions Kashida based on the Moon and Sun Letters Concept,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 8, pp. 286–290, 2017, doi: 10.14569/ijacsa.2017.080838.
[16] F. R. Shareef Taka, “Journal of Information Hiding and Multimedia Signal Processing Arabic Text Steganography based on Arabic Astrology,” *Ubiquitous Int.*, vol. 14, no. 3, pp. 124–135, 2023.
[17] R. H. Ali, B. N. Dhannoon, and M. I. Hamel, “Arabic text steganography using lunar and solar diacritics,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 31, no. 3, pp. 1559–1567, 2023, doi: 10.11591/ijeecs.v31.i3.pp1559-1567.
[18] N. Subhi Shakir and M. Salih Mahdi, “Using Special Letters and Diacritics in Steganography in Holy Quran,” *Iraqi J. Comput. Informatics*, vol. 49, no. 2, pp. 1–8, 2023, doi: 10.25195/ijci.v49i2.417.
[19] M. Alkhudaydi and A. Gutub, “Securing Data via Cryptography and Arabic Text Steganography,” *SN Comput. Sci.*, vol. 2, no. 1, 2021, doi: 10.1007/s42979-020-00438-y.
[20] A. Ditta, M. Azeem, S. Naseem, K. Gulzar Rana, M. Adnan Khan, and Z. Iqbal, “A secure and size efficient algorithm to enhance data hiding capacity and security of cover text by using unicode,” *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 5, pp. 2180–2191, 2022, doi: 10.1016/j.jksuci.2020.07.010.
[21] N. Alanazi, E. Khan, and A. Gutub, “Inclusion of Unicode Standard seamless characters to expand Arabic text steganography for secure individual uses,” *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 4, pp. 1343–1356, 2022, doi: 10.1016/j.jksuci.2020.04.011.
[22] A. Alshamsi, S. Albaloushi, M. Alkhoori, H. Almheiri, and N. Ababneh, “Enhancing Arabic Text Steganography Based on Unicode Features,” *Int. J. Comput. Digit. Syst.*, vol. 11, no. 1, pp. 685–693, 2022, doi: 10.12785/ijcds/110155.
[23] N. Alanazi, E. Khan, and A. Gutub, “Functionality-Improved Arabic Text Steganography Based on Unicode Features,” *Arab. J. Sci. Eng.*, vol. 45, no. 12, pp. 11037–11050, 2020, doi: 10.1007/s13369-020-04917-5.
[24] O. F. A. Adeeb and S. J. Kabudian, “Arabic Text Steganography Based on Deep Learning Methods,” *IEEE Access*, vol. 10, no. May, pp. 94403–94416, 2022, doi: 10.1109/ACCESS.2022.3201019.
[25] A. Boulesnane, A. Beggag, and M. Zedadik, “A New Steganography Technique Based on Dotted Arabic Letters Features,” *5th Int. Conf. Netw. Adv. Syst. ICNAS 2021*, no. December, 2021, doi: 10.1109/ICNAS53565.2021.9628914.
[26] A. R. Khekan, H. M. W. Majeed, and O. F. Ahmed Adeeb, “New text steganography method using the arabic letters dots,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 21, no. 3, pp. 1784–1793, 2021, doi: 10.11591/ijeecs.v21.i3.pp1784-1793.
[27] A. I. Alaqeel and M. S. Saleh, “Developing a Performance-based Tool for Arabic Text Steganography,” *Proc. - 2021 IEEE 4th Natl. Comput. Coll. Conf. NCCC 2021*, no. 2, 2021, doi: 10.1109/NCCC49330.2021.9428837.
[28] F. R. Shareef Taka, “Secure communication by combined diffe-hellman key exchange based AES encryption and Arabic text steganography,” *J. Inf. Hiding Multimed. Signal Process.*, vol. 12, no. 4, pp. 186–198, 2021.
[29] F. R. S. Taka, “Text steganography based on noorani and darkness,” *J. Inf. Hiding Multimed. Signal Process.*, vol. 12, no. 3, pp. 127–139, 2021.
[30] R. Din, R. A. Thabit, N. I. Udzir, and S. Utama, “Traid-bit embedding process on arabic text steganography method,” *Bull. Electr. Eng. Informatics*, vol. 10, no. 1, pp. 493–500, 2021, doi: 10.11591/eei.v10i1.2518.
[31] A. Taha, A. S. Hammad, and M. M. Selim, “A high capacity algorithm for information hiding in Arabic text,” *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 32, no. 6, pp. 658–665, 2020, doi: 10.1016/j.jksuci.2018.07.007.
[32] N. A. ROSLAN, N. I. UDZIR, R. MAHMOD, Z. A. ZUKARNAIN, M. I. H. NINGGAL, and R. THABIT, “Character property method for arabic text steganography with biometric multifactor authentication using liveness detection,” *J. Theor. Appl. Inf. Technol.*, vol. 98, no. 24, pp. 4140–4157, 2020.
[33] S. M. A. Al-Nofaie and A. A. A. Gutub, “Utilizing pseudo-spaces to improve Arabic text steganography for multimedia data communications,” *Multimed. Tools Appl.*, vol. 79, no. 1–2, pp. 19–67, 2020, doi: 10.1007/s11042-019-08025-x.
[34] A. A. A.F, “A Multi-Layer Arabic Text Steganographic Method Based on Letter Shaping,” *Int. J. Netw. Secur. Its Appl.*, vol. 11, no. 01, pp. 27–40, 2019, doi: 10.5121/ijnsa.2019.11103.
[35] N. Alghamdi and L. Berriche, “Capacity investigation of Markov chain-based statistical text steganography: Arabic language case,” *ACM Int. Conf. Proceeding Ser.*, pp. 37–43, 2019, doi: 10.1145/3314527.3314532.
[36] H. K. Tayyeh, M. S. Mahdi, and A. S. A. AL-Jumaili, “Novel steganography scheme using Arabic text features in Holy Quran,” *Int. J. Electr. Comput. Eng.*, vol. 9, no. 3, pp. 1910–1918, 2019, doi: 10.11591/ijece.v9i3.pp1910-1918.