




Using Zero-Watermarking Algorithm to Protect Arabic Text Documents

Shahbaa Mohammed Abdulmaged¹^{*}, Nadia Mohammed Abdulmaged²^{*},
Aumama Mohammed Farhan¹

¹Al-Iraqia University, Law dep., Baghdad, Iraq.

²University of Baghdad, Computer dep., Baghdad, Iraq.

*Corresponding Author: Shahbaa Mohammed Abdulmaged

DOI:<https://doi.org/10.55145/ajest.2025.04.01.002>

Received May 2024; Accepted July 2024; Available online August 2024

ABSTRACT: In today's digital age where powerful communication media such as the Internet allow for efficient information exchange, copyright protection and authentication of digital content has become an important issue. Plain text is the most widely used means of exchanging information online, and it is important to verify the authenticity of the text. Very limited techniques are available for watermarking and authentication of Arabic text. This paper aims to develop a zero-watermark algorithm to protect the Arabic digital content from being tampering. The algorithm creates an accurate zero watermark based on the characteristics of the content of the Arabic text itself. The watermark is determined by the frequency of appearance of prepositions and their specific locations (في، على، عن، من) in the text. The extraction algorithm used for extracting zero watermark and compared with the created one to prove the authenticity of the Arabic text. This method ensures that any change to the text will lead to a change in the extracted zero watermark, which will facilitate the detection process upon verification. Experimental results demonstrate the effectiveness of the algorithm in discovering text tampering (addition, deletion, and rephrasing), tested on nine different text samples with various lengths and attacks, the proposed approach is robust in detecting modifications to the text, even if they are minor, compared to the existing approach.

Keywords: zero watermarking, security, Arabic text, preserving, protection



1. INTRODUCTION

The digital age has made sharing Arabic documents online easier than ever before. However, with this convenience comes a new critical challenge: ensuring the authenticity and integrity of digital content [1,2], especially for all kinds of Arabic text documents, from academic papers and religious texts to everyday communication, all demanding robust security measures to protect these documents [2-5]. During digital document transfer, there is a risk that someone could tamper with them, which may lead to critical problems. To fight against such threats, various information security techniques could be used, including access controls, content verification methods, and tamper detection [5,6].

Traditional watermarking techniques provide a layer of protection by allowing different types of data, including audio, text, video, or even basic image data, to be embedded within a digital file, this hidden data acts as a secret key and offers several security advantages [7]. However, these techniques can alter the original content [8], which might not be ideal for preserving texts of the Holy Quran, the Prophet's hadiths, historical manuscripts or legal documents. Here, zero watermarking emerges as a revolutionary and non-intrusive solution.

Zero watermarking offers an effective and inconspicuous approach to protect Arabic text documents. Unlike traditional approaches, it does not mess with the original content itself [9]. Rather, it hides a unique digital signature, or watermark, into the document's inherent structure or other features. This watermark allowing for verification of the document's authenticity and detection of any unauthorized modifications.

2. RELATED WORK

This research [1] proposes RCATED-AT, a robust system for Arabic text authentication and tamper detection online. RCATED-AT leverages a 4th-level word order process based on Markov models to extract features used as watermarks.

The research [2] proposed an intelligent text Zero-Watermarking approach (ZWAFWMMM) using fourth level of Markov Model. Experiments have proven the effectiveness of ZWAFWMMM in embedding data and detecting manipulation with high accuracy.

The research [4] proposed a hybrid approach (HNLZWA) for the content authentication and tampering detection of Arabic text. The approach integrates the treatment of Natural Language Processing and Zero-Watermarking to analyze the Arabic text and extract features that support the detection of tampering attacks with high level of accuracy. Experiments on four datasets showed the effectiveness of HNLZWA in detecting all kinds of tampering attacks

The proposed approach [10] for zero watermarking Arabic text achieves watermark embedding and identification without modifying the original document. It does this by combining existing automated zero watermarking techniques with a novel second-layer framework based on the Markov model. This framework acts as a natural language processing tool, analyzing the Arabic text and extracting features that capture the relationships between words within the context. These extracted features then become the watermark itself. Importantly, the entire process is designed to be fragile. Any modifications to the text will disrupt the watermark, making them easily detectable during validation.

This approach [11] avoids modifying the original digital text by embedding and extracting the watermark logically. It achieves this by analyzing the Arabic text using a special technique called the "fourth-level-order and alphanumeric mechanism of the Markov model." This analysis extracts characteristic features from the text, which essentially become the digital watermark. Later, this watermark is used to detect any tampering attempts made to the received Arabic text.

3. PROPOSED ALGORITHM

The proposed approach utilizes the intrinsic properties of the Arabic text document itself to create a zero watermark, eliminating the need for any modifications to the original content. This fragile watermark serves as an authenticity verification tool for Arabic text documents. The generation and extraction processes for this watermark are depicted in Fig. 1. Additionally, the watermark is registered with a CA and employed within the extraction algorithm to confirm the document's legitimacy.

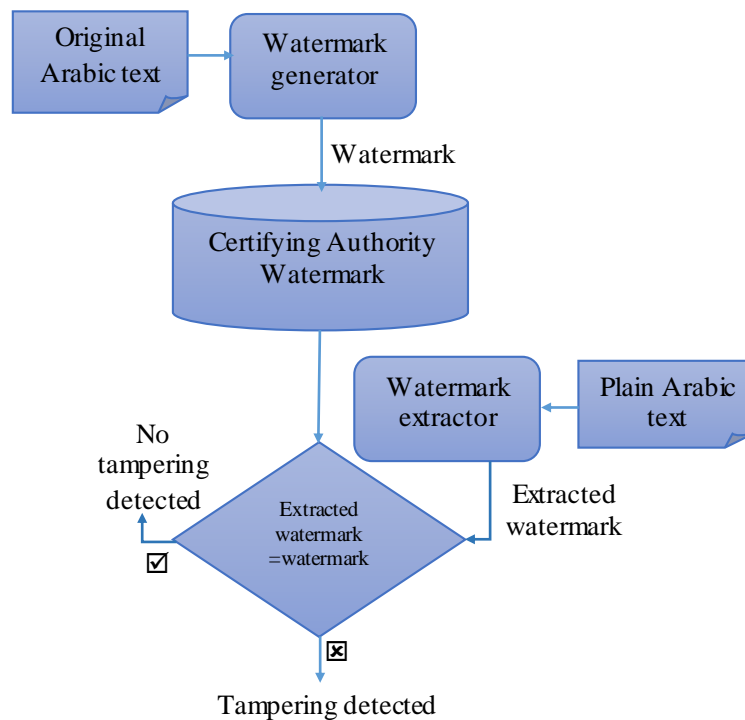


FIGURE 1. - Flowchart of generating watermark and extracting processes

The Arabic text's contents are used by the suggested approach to create a zero watermark. Specifically, it identifies certain prepositions (في، عن، من، على) within the Arabic text and calculates the frequency of their occurrence for each preposition individually. It then stores three distinct locations (if possible), where each preposition appears within the text (e.g., at the beginning, middle and end). Otherwise, the first three locations (if found) are stored. The resulting zero

watermark for the Arabic text combines the frequency of each preposition's occurrence and their respective locations. As depicted in Fig 2, this zero watermark can be used to verify the authenticity of the Arabic text.

يتم استخدام محتويات النص العربي في الخوارزمية المقترحة لحمايته. إذ يتم البحث عن بعض حروف الجر (في، عن، من وعلی) الواردة في النص العربي، ويتم حساب عدد تكرار ظهورها ضمن المتن ولكل حرف علی حدة، مع حفظ ثلاثة مواقع متفرقة ان أمكن من التي ظهر فيها كل حرف من حروف الجر من بين كلمات النص (مثلا موقع في بداية النص، موقع في منتصف النص وموقع في نهاية النص)، وإلا فسيتم الاحتفاظ بأول ثلاث مواقع إن وجدت، ويتم الاحتفاظ بالنتيجة النهائية لعدد تكرار ظهور كل حرف جر ضمن المتن وأول ثلاثة مواقع ظهر فيها حرف الجر، والنتيجة النهائية ستمثل العلامة المائية الصفرية لذلك النص العربي، وكما موضح في هذا الشكل...الخ.

Prepositions	في	عن	من	على
Freq. of occurrence	7	2	3	2
Position	6, 22, 62	13, 18, 0	19, 45, 48	20, 34, 0

Watermark= 4, 6, 22, 62, 2, 13, 18, 0, 3, 19, 45, 48, 2, 20, 34, 0

FIGURE 2. - Watermark Generation

This approach employs zero watermarking, where the watermark is not directly embedded but generated from the inherent characteristics of the Arabic text. The process consists of two stages: embedding and extraction. The original author embeds the watermark, and a trusted CA later extracts it to verify ownership. This CA plays a crucial role, as copyright owners register their watermarks with it. In case of content ownership disputes, this trusted third party serves as the final authority.

3.1 EMBEDDING ALGORITHM

The watermark embedding algorithm takes the original Arabic text file as input and generates a unique watermark based on the text's characteristics. This watermark, along with the timestamp, author name, and the original text document, is then registered with a trusted CA for ownership verification. Fig. 3 shows how the algorithm works.

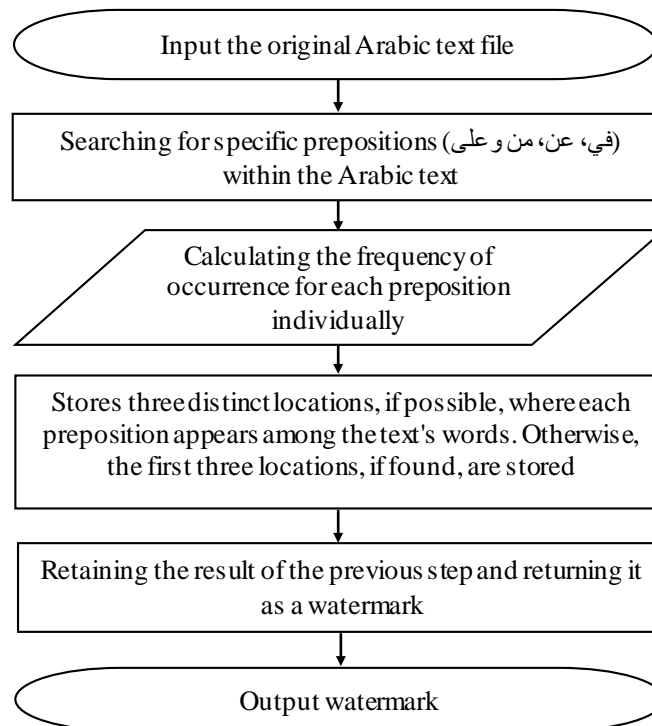


FIGURE 3. - Flowchart of embedding algorithm

First, the author provides the original Arabic text, and the frequency of occurrence of each preposition (في، عن، من، على) in the text is analyzed. A digital watermark is generated based on the total occurrence count of each preposition within the text and the three locations where each preposition appears. Subsequently, the CA registers this watermark with the current time and date.

3.2 EXTRACTING ALGORITHM

This section details the watermark extraction algorithm. The plain Arabic text (potentially tampered with) and the original watermark it used as input. The algorithm analyzes the text and generates a watermark based on its characteristics. The generated watermark is then compared to the original watermark that was registered with the CA, along with the author’s name and timestamp. Conflicts arising from multiple watermark registrations are resolved by considering the timestamps, where the author with the earlier registration is considered the rightful owner

In the absence of tampering, the algorithm accurately detects the watermark, thereby verifying the authenticity of the Arabic text document. However, if the text has been tampered with (through insertions, deletions or rephrasing), the watermark becomes distorted. Fig. 4 will likely describe the specific steps involved in the watermark extraction algorithm.

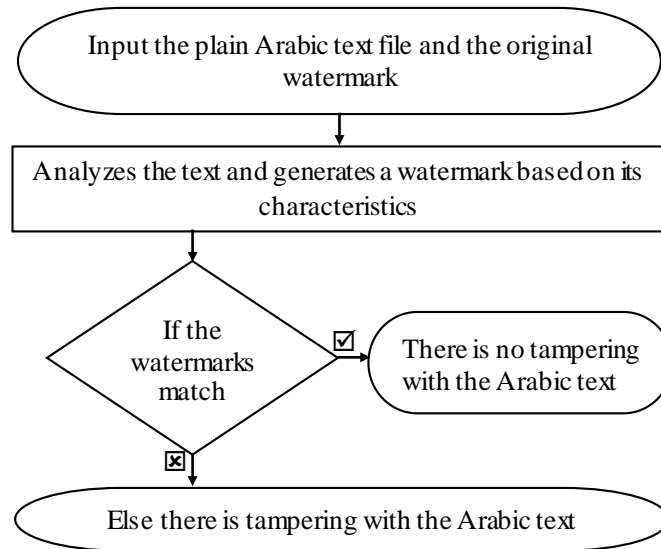


FIGURE 4. - Flowchart of extraction algorithm

4. EXPERIMENTAL RESULTS

To assess the robustness of the proposed zero watermarking approach, nine Arabic text samples of varying size were employed from e-books and webpages. These samples were subjected to different tampering methods: word insertions, deletions, and rephrasing. The attacks were applied at random locations within the text. Table 1 details the sample identification, the word count of the original text, the volume attack of insertions, deletions, and rephrasing applied and the final word count after the attack.

Table 1. - Original and attacked Arabic text samples with volume attack of insertions, deletions, and paraphrases applied

Sample No.	word count of original text	volume attack			word count of attacked text
		Insertion	Deletion	rephrasing	
1 [ASST1]	193	3%	2%	1%	195
2 [ASST3]	578	50%	25%	10%	722
3 [ASST6]	802	7%	5%	3%	818
4 [ASST5]	847	45%	35%	15%	847
5 [ASST8]	896	12%	10%	8%	986
6 [AMST2]	1068	4%	4%	2%	1068
7 [AMST4]	1378	15%	11%	5%	1433
8 [AMST9]	1550	35%	30%	9%	1627
9 [ALST7]	6959	25%	20%	4%	7307

This study examined how often the prepositions (على، من، عن، في) appeared in both the attacked text samples and the original one. These specific prepositions were chosen for the reason that they are commonly used across all types of writing. The next formulas will explain how two metrics, Watermark Distortion Rate (WDR) and Accuracy Rate (WAR) are computed:

$$\text{WAR} = \text{Characters accurately identified} / \text{Characters in the watermark}$$

$$\text{WDR} = 1 - \text{WAR}$$

Where,

$$0 \leq \text{WAR} \leq 1; \text{WAR should be close to 1.}$$

$$0 \leq \text{WDR} \leq 1; \text{WDR should be close to 0.}$$

The effectiveness of watermark extraction was evaluated by comparing its WAR with the original watermark. This analysis helped determine whether or not the text had been tampered with. Tab. 2 displays the WAR values alongside the count and order of the chosen prepositions in both the original and attacked text.

Table 2. - Accuracy of extracted watermark

Sample No.	The prepositions count and its order in original text (على، من، عن، في)	The prepositions count and its order in attacked text (على، من، عن، في)	Tampering detected	WAR
1 [ASST1]	12, 48, 56, 60, 0, 0, 0, 0, 3, 50, 143, 186, 6, 17, 46, 86	12, 54, 62, 66, 0, 0, 0, 0, 3, 51, 144, 187, 6, 18, 47, 87	yes	0.9674
2 [ASST3]	4, 27, 45, 229, 2, 280, 368, 0, 2, 163, 294, 0, 5, 24, 177, 299	6, 9, 28, 100, 3, 60, 353, 441, 2, 236, 367, 0, 5, 97, 250, 372	yes	0.8239
3 [ASST6]	18, 103, 109, 122, 3, 167, 667, 762, 18, 24, 38, 50, 22, 9, 27, 54	18, 119, 124, 147, 3, 179, 679, 774, 20, 7, 13, 37, 23, 11, 23, 40	yes	0.9891
4 [ASST5]	21, 71, 85, 102, 4, 214, 368, 411, 17, 202, 468, 506, 15, 9, 25, 30	21, 87, 105, 143, 4, 229, 383, 426, 17, 217, 483, 512, 15, 24, 40, 45	yes	0.9262
5 [ASST8]	15, 7, 29, 62, 8, 228, 299, 570, 7, 53, 266, 409, 19, 57, 326, 341	15, 14, 36, 68, 8, 235, 306, 577, 7, 60, 273, 416, 19, 64, 333, 348	yes	0.9701
6 [AMST2]	14, 34, 118, 172, 79, 214, 222, 231, 4, 205, 749, 1032, 8, 42, 47, 67	14, 32, 116, 170, 79, 212, 220, 229, 4, 203, 747, 1030, 8, 40, 45, 65	yes	0.9506
7 [AMST4]	47, 13, 37, 62, 13, 58, 400, 460, 23, 117, 121, 125, 12, 8, 162, 292	48, 16, 45, 71, 13, 92, 432, 492, 23, 149, 153, 157, 12, 40, 194, 324	yes	0.8624
8 [AMST9]	39, 12, 141, 147, 10, 9, 50, 65, 15, 537, 573, 646, 38, 198, 607, 708	39, 12, 147, 153, 15, 9, 50, 65, 10, 537, 573, 646, 32, 198, 607, 708	yes	0.9984
9 [ALST7]	171, 28, 33, 42, 20, 741, 1441, 1706, 141, 25, 59, 67, 172, 15, 61, 103	171, 31, 36, 45, 20, 741, 1441, 1706, 141, 25, 59, 67, 172, 15, 61, 103	yes	0.9981

Table 2 shows that any attempt to alter the text is always identified. The lower watermark accuracy, the more text has been tampered with. Looking at samples 1, 4, 5, 6, and 9, it can see that even if the number of times each preposition appears (in this case, the prepositions are من، عن، في، and على) stays the same in both the original and modified texts, it doesn't guarantee that the prepositions will be in the same three locations they were originally.

Fig 5 illustrates the WDR for the prepositions (على، من، عن، في) across all text samples. It is evident that the WDR remains high even with minimal volume attacks (as seen in samples 6 and 7) for all prepositions. Any changes made by the attacker can easily affect the text. A significant distortion rate is a sign that the text has been altered and is no longer genuine. This illustrates how even slight modifications have a big impact on the watermark's accuracy, and the watermark's fragility serves as a clear indication that the text has been attacked.

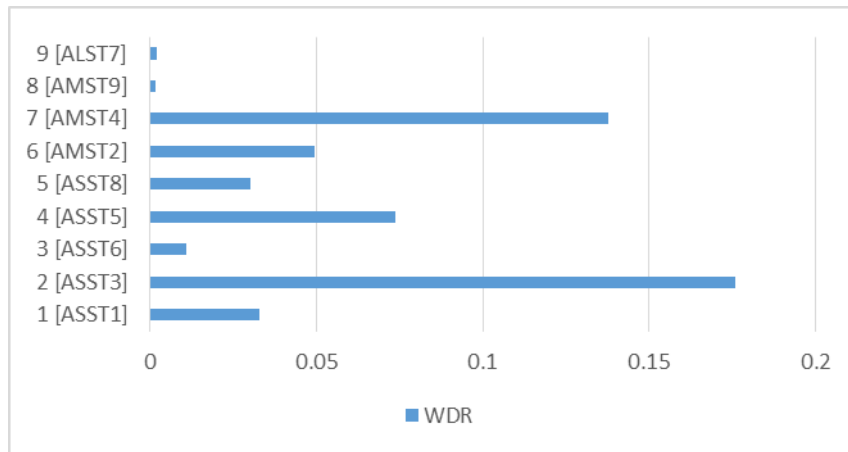


FIGURE 5. - shows the WDR with prepositions (في، عن، من، على) on all text samples

5. COMPARATIVE RESULTS

The proposed approach demonstrates higher accuracy in detecting watermark fragility compared to FATZWNLP [10] and RCATD-AT [1] (as shown in Tab. 3). This is particularly evident in rephrasing attacks. Since rephrasing attacks automatically trigger insertions and deletions, it significantly affects the watermark's integrity. In simpler terms, the proposed approach is better at identifying when the order of words is tampered with.

Table 3. - Watermark fragility accuracy comparison.

Approach	Attack volume	Attack		
		Insertion	Deletion	rephrasing
FATZWNLP	5%	93.32	88.14	77.49
	10%	88.03	82.50	64.10
	20%	80.29	66.30	44.51
	50%	62.12	38.73	25.49
RCATD-AT	5%	94.83	88.88	76.70
	10%	88.84	84.30	62.99
	20%	81.14	66.48	43.45
	50%	62.48	42.97	24.98
The proposed approach	5%	95.84	90.74	80.55
	10%	90.81	84.62	74.22
	20%	82.91	76.42	66.62
	50%	77.01	62.39	52.09

One potential future direction for the proposed approach is to build upon the findings of this research [12] to enhance Arabic character recognition for digital content and to develop a more advanced algorithm.

6. CONCLUSION

Traditional text watermarking techniques used for verifying document authenticity have two major weaknesses. They are not effective against random changes, especially when dealing with different text formats. Additionally, even slight alterations can go unnoticed, making it impossible to be certain of the information's accuracy. To address these issues, a novel zero watermarking algorithm has been created. This innovative approach uses the text's content to generate a unique watermark. By extracting this watermark later, it is possible to determine the document's legitimacy. The algorithm's performance was evaluated against random tampering attacks on a set of nine text samples with varying lengths. The results are impressive, showing that the proposed algorithm consistently detects tampering, no matter how minor the changes are. As you can see from table above, the proposed approach consistently outperforms the other approaches across all attack types and volumes. For instance, it exhibits a significant enhancement in handling insertion attacks, with improvements ranging from 11.83% to 33.43% compared to FATZWNLP.

FUNDING

None

ACKNOWLEDGEMENT

The authors would like to thank the anonymous reviewers for their efforts.

CONFLICTS OF INTEREST

The authors declare no conflict of interest

REFERENCES

- [1] A. S. Oleiwi, M. A. Alkhafajj, R. R. Ali, E. Ali, M. Al-Tahee, and M. Almusawi, "Robust content authentication and tampering detection of Arabic text transmitted through Internet," in *Proc. 2023 6th Int. Conf. Engineering Technology and its Applications (IICETA)*, Al-Najaf, Iraq, 2023, pp. 790–796, doi: 10.1109/IICETA57613.2023.10351462.
- [2] F. N. Al-Wesabi, K. Mahmood, and N. Nemri, "A zero watermarking approach for content authentication and tampering detection of Arabic text based on fourth level order and word mechanism of Markov model," *Journal of Information Security and Applications*, vol. 52, p. 102473, 2020, doi: 10.1016/j.jisa.2020.102473.
- [3] N. A. A. Usop and S. I. Hisham, "A review of digital watermarking techniques, characteristics and attacks in text documents," in *Proc. Int. Conf. Innovative Technology, Engineering and Science*, 2020, pp. 256–271, doi: 10.1007/978-3-030-70917-4_25.
- [4] F. N. Al-Wesabi, "Proposing high-smart approach for content authentication and tampering detection of Arabic text transmitted via Internet," *IEICE Trans. Inf. Syst.*, vol. 103, no. 10, pp. 2104–2112, 2020, doi: 10.1587/transinf.2020EDP7011.
- [5] A. A. Alkhafaji, N. N. A. Sjarif, M. A. Shahidan, N. F. M. Azmi, H. M. Sarkan, et al., "Tamper detection and localization for Quranic text watermarking scheme based on hybrid technique," *Computers Materials & Continua*, vol. 68, no. 1, pp. 771–802, 2021, doi: 10.32604/cmc.2021.015770.
- [6] R. Thabit, N. I. Udzir, S. M. Yasin, A. Asmawi, N. A. Roslan, et al., "A comparative analysis of Arabic text steganography," *Applied Sciences*, vol. 11, no. 15, p. 6851, 2021, doi: 10.3390/app11156851.
- [7] A. Bastani and E. Fatemi Behbahani, "A self-recovery digital watermarking approach for tamper detection of handwritten and printed electronic documents," *Library and Information Sciences*, vol. 24, no. 1, pp. 174–193, 2021, doi: 10.30481/LIS.2020.235473.1727.
- [8] S. Y. Chen, H. Ma, and Q. J. Chen, "Tamper detection of batch websites based on text comparison," in *Proc. Computer Science and Technology: Int. Conf.*, pp. 573–579, 2017, doi: 10.1142/9789813146426_0065.
- [9] U. Khadam, M. M. Iqbal, M. Alruily, M. A. Al Ghamdi, M. Ramzan, et al., "Text data security and privacy in the Internet of Things: threats, challenges, and future directions," *Wireless Communications and Mobile Computing*, 2020, doi: 10.1155/2020/7105625.
- [10] F. N. Al-Wesabi, A. Abdelmaboud, A. A. Zain, M. M. Almazah, and A. Zahary, "Tampering detection approach of Arabic text based on contents interrelationship," *Intell. Automat. Soft Comput.*, vol. 27, no. 2, pp. 483–498, 2021, doi: 10.32604/iasc.2021.014322.
- [11] A. M. Hilal, F. N. Al-Wesabi, M. A. Hamza, M. Medani, K. Mahmood, et al., "Content authentication and tampering detection of Arabic text: an approach based on zero-watermarking and natural language processing," *Pattern Analysis and Applications*, vol. 25, no. 1, pp. 47–62, 2022, doi: 10.1007/s10044-021-01032-5.
- [12] A. H. Ali, M. A. Mohammed, and M. A. Ahmed, "Character recognition by implementing FPGA-based artificial neural network," *Mesopotamian Journal of Computer Science*, vol. 2021, pp. 13–17, 2021, doi: 10.58496/MJCSC/2021/003.