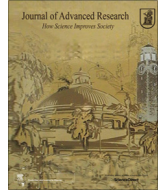




Contents lists available at ScienceDirect

Journal of Advanced Research

journal homepage: www.elsevier.com/locate/jare

Original Manuscript

Industrial Internet of Water Things architecture for data standarization based on blockchain and digital twin technology[☆]



Mazin Abed Mohammed^{a,i,j}, Abdullah Lkhan^{b,i,j}, Karrar Hameed Abdulkareem^{c,d}, Mohd Khanapi Abd Ghani^e, Haydar Abdulameer Marhoon^{f,g}, Seifedine Kadry^{h,*}, Jan Nedomaⁱ, Radek Martinek^j, Begonya Garcia Zapirain^k

^a Department of Artificial Intelligence, College of Computer Science and Information Technology, University of Anbar, Anbar 31001, Iraq

^b Department of Cybersecurity and Computer Science, Dawood University of Engineering and Technology, Karachi City 74800, Sindh, Pakistan

^c College of Agriculture, Al-Muthanna University, Samawah 66001, Iraq

^d College of Engineering, University of Warith Al-Anbiyaa, Karbala 56001, Iraq

^e Department of Software Engineering, Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM), Malaysia

^f Information and Communication Technology Research Group, Scientific Research Center, Al-Ayen University, Thi-Qar, Iraq

^g College of Computer Sciences and Information Technology, University of Kerbala, Karbala, Iraq

^h Seifedine is with the Department of Applied Data Science, Noroff University College, Kristiansand, Norway

ⁱ Department of Telecommunications, VSB-Technical University of Ostrava, 70800 Ostrava, Czech Republic

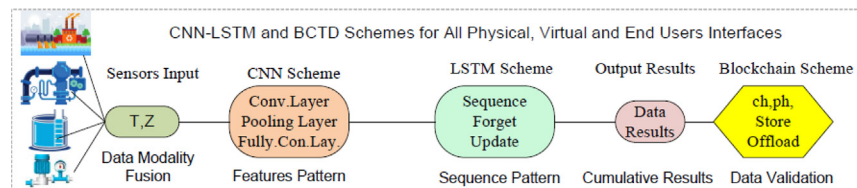
^j Department of Cybernetics and Biomedical Engineering, VSB-Technical University of Ostrava, 70800 Ostrava, Czech Republic

^k Begonya is with the eVIDA Laboratory, University of Deusto, 48007 Bilbao, Spain

HIGHLIGHTS

- We present a hybrid model called CNN-LSTM. The CNN-LSTM model is designed to efficiently extract and process data from diverse sources, thereby addressing the issue of data modality.
- We devise the cross-platform runtime environment based on Message Queuing Telemetry Transport (MQTT) protocol. This environment allows for the seamless processing of diverse data sources and types through heterogeneous computing nodes in IIoWT. By leveraging MQTT, the devised runtime environment ensures interoperability and facilitates the execution of data across heterogeneous computing nodes.
- We devise a blockchain scheme that utilizes SHA-256 for data hashing. This scheme incorporates a proof-of-work mechanism and ensures data validity across heterogeneous nodes in IIoWT.

GRAPHICAL ABSTRACT



[☆] Peer review under responsibility of Cairo University.

* Corresponding author.

E-mail address: mazinalshujaery@uoanbar.edu.iq (M.A. Mohammed).

ARTICLE INFO

Article history:

Received 8 July 2023

Revised 13 September 2023

Accepted 5 October 2023

Available online 13 October 2023

Keywords:

IIoWT

Blockchain

Water management

Digital twin

Transparency

Data modality and standardization

Introduction:: The Industrial Internet of Water Things (IIoWT) has recently emerged as a leading architecture for efficient water distribution in smart cities. Its primary purpose is to ensure high-quality drinking water for various institutions and households. However, existing IIoWT architecture has many challenges. One of the paramount challenges in achieving data standardization and data fusion across multiple monitoring institutions responsible for assessing water quality and quantity.

Objective:: This paper introduces the Industrial Internet of Water Things System for Data Standardization based on Blockchain and Digital Twin Technology. The main objective of this study is to design a new IIoWT architecture where data standardization, interoperability, and data security among different water institutions must be met.

Methods:: We devise the digital twin-enabled cross-platform environment using the Message Queuing Telemetry Transport (MQTT) protocol to achieve seamless interoperability in heterogeneous computing. In water management, we encounter different types of data from various sensors. Therefore, we propose a CNN-LSTM and blockchain data transactional (BCDT) scheme for processing valid data across different nodes.

Results:: Through simulation results, we demonstrate that the proposed IIoWT architecture significantly reduces processing time while improving the accuracy of data standardization within the water distribution management system.

Conclusion:: Overall, this paper presents a comprehensive approach to tackle the challenges of data standardization and security in the IIoWT architecture.

© 2023 The Authors. Published by Elsevier B.V. on behalf of Cairo University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Recently, there have been notable advancements in the Industrial Internet of Things (IIoT) paradigm, which focuses on implementing Internet of Things (IoT) applications for water management [1]. IIoT incorporates automated sensors, sensing devices, and data analytics to improve sustainability, efficiency, data security, and transparency in industrial operations involving multiple stakeholders. IIoT fetches about a revolutionary transformation in water management by facilitating precise calculations, real-time monitoring, and comprehensive profiling of resources within distributed infrastructures [2]. This paradigm opens up many possibilities for diverse water applications, incorporating various data modalities and features. Some notable examples of these applications include using sensors to monitor and assess water quality, track water consumption, analyze water composition, measure water levels, detect leaks, evaluate wastage, and understand consumption patterns. Data standardization plays a vital role in harnessing the potential of IIoT-enabled applications. Establishing uniformity and consistency in data representation and formats across different platforms and systems is imperative [3–5].

IIoT recognizes the potential of blockchain technology in creating new data standards that can seamlessly integrate and operate across various platforms. This standardized data format would benefit multiple stakeholders involved in water management, including end customers, government regulatory bodies, and water distributors. By adopting standardized data formats, valuable insights can be gained from water quality data. For instance, information regarding the precise amount of chlorine dosage and hormonal levels can be efficiently communicated to both government entities responsible for the regulation and end customers who are concerned about water quality and safety. This data standardization ensures transparency, reliability, and interoperability in the exchange of water-related information, ultimately enhancing the efficiency and effectiveness of water management processes. Effective data modalities and standardized data management are key factors in enhancing IIoT performance. Implementing technologies like blockchain and digital twins allows for this improvement. The convergence of the Industrial Internet of Things (IIoT) paradigm with blockchain technology and Digital Twin (DT) concepts brings numerous advantages to the water industry [6,7]. Blockchain technology (BCT) plays a pivotal role in processing and managing the data generated by interconnected water industries.

It operates with transparency, security, and immutability, ensuring the data remains tamper-proof and protected against unauthorized modifications. By leveraging blockchain technology, applications in the water industry can establish a reliable and auditable data trail throughout the supply chain, ensuring data integrity and enabling authenticated data processes. BCT provides a robust framework for data asset management in the water industry. Through the utilization of Digital Twin (DT) technology, physical data from various locations and industries can be replicated and managed virtually. DT offers scalability, cost-efficiency, and interoperability, making it an invaluable tool for water industry applications. With DT, stakeholders in the water industry can create virtual replicas of their physical assets, enabling real-time monitoring, analysis, and simulations. This facilitates improved decision-making, predictive maintenance, and optimization of water management processes. In summary, integrating IIoT, blockchain technology, and Digital Twin concepts within the water industry offers benefits such as transparent and secure data management, traceability throughout the supply chain, authenticated data processes, and efficient data asset management through virtual replication. These technologies empower water industry applications to enhance operations, refine decision-making processes, and achieve cost-effective and scalable solutions [8–11].

Deep learning (DL) has garnered considerable attention in research and academia due to its potential for automating processes within the Industrial Internet of Things (IIoT) framework, particularly in water applications. DL, an integral part of artificial intelligence (AI), encompasses a range of mechanisms and algorithms. Notably, CNN and RNN strategies were utilized in DL. The primary objective is to effectively process and analyze the vast volume of data generated by water sensors in real-time. DL techniques excel in handling diverse data modalities associated with water sensors, including image data captured by sensors, time series data, and numeric and textual data obtained from distributed infrastructures.

However, data standardization in water data management within the Industrial Internet of Things (IIoT) presents numerous challenges that must be addressed. This paper considers these research questions to solve the data standardization problem in IIoWT architecture. (i) The first challenge arises from the diverse data modalities encountered in water management. (ii) The data originates from various sources and is represented in different formats, necessitating suitable processing across multiple nodes. Sec-

only, the attainment of data interoperability and integration poses a significant obstacle. The IIoT involves collecting data from disparate sources, making it arduous to convert and consolidate it into a standardized format that can be seamlessly processed and yield meaningful outcomes on the nodes. (iii) Data security is another pivotal challenge, particularly when distributing data among IIoT water nodes involving end users, industries, and government entities.

To address the aforementioned challenges, this paper presents the Industrial Internet of Water Things (IIoWT) architecture, which includes components such as a water quality monitoring application, computing and wireless communication infrastructure, and data-generating sensors. The application aims to monitor water quality by implementing secure data standardization protocols that enable smooth data exchange among end users, service providers, and government entities. The integration of CNN and blockchain reveals many potential advantages and applications, particularly within the IIoWT (Industrial Internet of Things) industry. CNNs can be employed to analyze various types of data, such as images, text, and content, while blockchain technology ensures robust security and confidentiality, especially for sensitive data like patients' medical records. By storing the training data and predictions of the CNN model on the blockchain, it becomes possible to authenticate and validate the accuracy and reliability of the model's results, thereby fostering increased confidence in the system.

To answer the research questions raised earlier, this paper offers the following research findings and contributions to the existing body of knowledge.

- We present a hybrid model called CNN-LSTM. The CNN-LSTM model is designed to efficiently extract and process data from diverse sources, thereby addressing the issue of data modality.
- We devise the cross-platform runtime environment based on the Message Queuing Telemetry Transport (MQTT) protocol. This environment allows for the seamless processing of diverse data sources and types through heterogeneous computing nodes in IIoWT. By leveraging MQTT, the devised runtime environment ensures interoperability and facilitates the execution of data across heterogeneous computing nodes.
- We devise a blockchain scheme that utilizes SHA-256 for data hashing. This scheme incorporates a proof-of-work mechanism and ensures data validity across heterogeneous nodes in IIoWT. The existing blockchain schemes, such as proof of credibility, byzantine fault-tolerant, and hash validation schemes, consume more resources and support only homogeneous nodes in distributed networks.

The paper is organized into several sections. Section 1 focuses on the related work, providing an overview of previous studies in the field. Section 1 explains the IIoWT architecture, outlining its components and functionality. Section 1 presents the proposed methodology, detailing how it operates and addresses the research objectives. The implementation of the system and performance evaluation are discussed in Section 28. Finally, Section 28 concludes the paper and outlines potential avenues for future research.

Related work

The IIoT paradigm has been widely applied in water management across various smart cities. Its primary goal is to enable water quality monitoring across different gradients and stakeholders. In these studies [1–5], an industrial IoT architecture was presented to collect gradient data from raw water at different plants. These

studies showcased various applications for monitoring raw water quality using plant ultrasonic sensors. The surveys and case studies discussed both application and infrastructure-related challenges. However, data standardization and security about data modality still need to be addressed in these works. To ensure data security standardization across heterogeneous computing nodes in water management, previous studies [6,7] have proposed decentralized architectures based on blockchain technology. The primary objective was to facilitate the secure transfer of valid data among interconnected nodes using blockchain. The key contribution of this architecture lies in its ability to process data modalities in a transparent, immutable, and authenticated format across nodes. The studies presented blockchain-based systems for IoT water management in smart cities. The objective of these systems was to enable the transformation of water quality, sensor, and application data through a distributed ledger among involved stakeholders. Different processing levels implemented proof-of-work-validity, proof-of-credibility, and fault-tolerant schemes within the blockchain. These blockchain schemes aimed to ensure consistency in water data for quality enhancement and compliance with standards across various locations. However, developing and implementing blockchain blocks for data standardization with the aforementioned studies incurred significant resource and cost requirements at different nodes.

Previous studies [8–11] explored the integration of Digital Twin (DT) technology to mitigate the implementation cost associated with blockchain technology in water management systems and applications. They presented a DT-enabled blockchain solution wherein the implementation relied on virtualizing the consensus rules of physical blockchain blocks. The aim was to decrease the implementation cost of water management across various nodes by leveraging the combination of blockchain technology and DT. However, resource.

A case study by Xu et al. [12] presented a water quality monitoring system utilizing underwater sensors in fish ponds. The study's objective was to analyze the water quality and levels in various tanks using IoT technology and examine the flow after storage from different water sources in distributed fog cloud networks. Another study by Smith et al. [13] focused on a drinking water quality system that performs real-time checks on raw materials and their quality. The researchers aimed to develop a lightweight and cost-efficient IoT water management system. Furthermore, several studies [14–18] presented distributed architectures for water management, where raw data collected from diverse sensors is processed at different nodes and aggregated centrally for further analysis. However, it should be noted that these water management systems can only process one data modality at a time.

By leveraging DL, improvements can be made in various aspects of water management. For example, DL algorithms can enhance water quality assessments, enable the detection of vulnerabilities and faults, provide fault tolerance mechanisms, and extract relevant contextual features during the data processing phase. Overall, DL's capabilities make it a valuable tool for addressing the challenges present in water management. DL empowers water industry stakeholders to make informed decisions, optimize processes, and improve water-related operations' overall efficiency and effectiveness through its ability to process and analyze complex and diverse data. In previous studies [18–20], various DL and machine learning schemes were proposed to address data modality processing and feature extraction in distributed water management environments. Specifically, using CNN and LSTM models was widely explored for handling data modality and data fusion challenges, particularly in scenarios where sensor data processing is distributed across multiple nodes. These studies successfully processed diverse data formats, including images, text, numeric, and

other data types. The paper [21] proposed a framework based on blockchain technology for environment and distributed data management. The objective was to establish a trusted water management process among interconnected nodes belonging to different stakeholders. To enhance the dynamic adoption of IoT-based water management in distributed blockchain-enabled networks, several studies [22–25] presented their findings. In these studies, IoT sensors were deployed across various water management components to monitor water conditions at different locations. Implementing Chinese smart water systems based on blockchain technology enabled transparent, secure, and valid record-keeping across diverse locations within smart cities. The studies [26–33] presented architectures for water management in agriculture applications based on blockchain technology. These studies proposed various components to facilitate water management in different architectural contexts. The presented architectures comprised several elements: source linking, processing nodes, pump speed control, water flow management, and scheduling mechanisms across distributed nodes.

Based on our current understanding, unresolved challenges are still related to the standardization of data modalities, security aspects, real-time water quality monitoring across various nodes, and interoperability issues. In light of these challenges, this paper aims to address the issues of data modality security, interoperability, water management, and data validation among nodes.

IloWT architecture

The architecture IloWT, depicted in Fig. 1, comprises various components, including the water source, harvest system, storage

facility, distribution pump, and end-user institutions such as houses, hospitals, schools, and more. Fig. 1 shows the flow of water management from raw material to the distribution of clean water in the different institutions for drinking. The water source and water treatment plant are responsible for the purification of water obtained from diverse sources. The water source gathers water from various points and employs sensors to generate data. These sensors include Residual Chlorine [34], Organic Carbon [35], Turbidity [36], Conductivity [37], pH [38], and others, which are utilized to monitor the quality of water at the collection sources for the water treatment plant. The architecture consists of physical and virtual servers and application interfaces. The physical servers are integrated into the power plant and government monitoring offices, while digital twin-enabled virtual servers are deployed in various water management offices to handle water flow and distribution. The physical servers exhibit heterogeneity, whereas the virtual servers possess homogeneous characteristics. Additionally, the computing node is integrated with various components to monitor the water quality within the architecture. The architecture collects data from diverse sources, each with its unique format. This heterogeneity poses a significant challenge in terms of interoperability and processing. To address this issue, our architecture incorporates a runtime design enabled by Message Queuing Telemetry Transport (MQTT) [39], ensuring seamless interoperability among data formats and runtimes. Furthermore, we have integrated deep learning convolutional neural network (CNN) [40] and least memory (LSTM) [41] schemes into our architecture to effectively process data with varying features and formats. This enables efficient analysis and utilization of the collected data. Our study integrates lightweight blockchain schemes within different

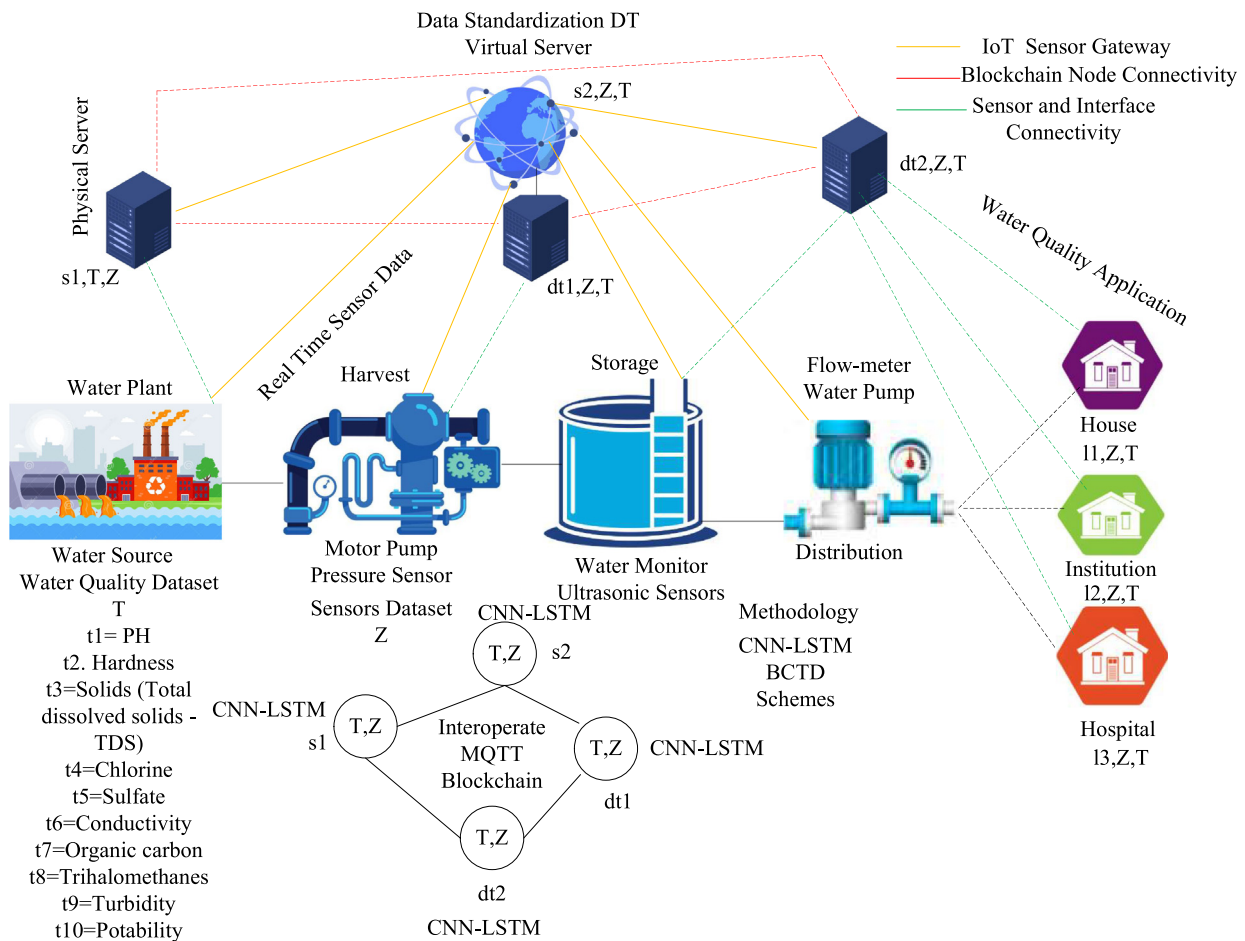


Fig. 1. Industrial Internet of Water Things architecture for data standardization based on blockchain and digital twin technology.

nodes to ensure security and maintain data standards. This ensures the immutability and transparency of data sharing among the nodes, enhancing overall security and data integrity. The architecture incorporates various algorithm schemes, including CNN-LSTM and blockchain transactions data schemes (BCTD). The primary objective is to manage standardized data and process it efficiently across different nodes while ensuring transparency, immutability, and validity.

All the notations of the problem are defined in Table 1.

Problem mathematical formation

The architecture comprises S physical servers, each denoted as s . Each physical server possesses computing speed (ζ_s) and resource capability (ϵ_s). Additionally, virtual servers are denoted as DT , shared components among the physical servers. Each virtual server, represented by dt , has its own computing speed (ζ_{dt}) and resource capability (ϵ_{dt}). Moreover, computing nodes exist in houses and institutions, denoted as L , which are shared components among the physical servers. Each local interface, referred to as l , possesses its own computing speed (ζ_l) and resource capability (ϵ_l).

The architecture encompasses a collection of T sensors, where each sensor t is responsible for monitoring specific data within a particular water plant and storage facility. These T sensors generate numeric data through various sensing mechanisms. Specifically, sensors such as those detecting water leakage, pump water flow, and faults are denoted as Z . In our analysis, we examine a total of B blockchain blocks, each denoted by ba , representing the specific attributes associated with the blockchain block in the architecture. We establish the blockchain validation mechanism using the formula presented below.

$$Enc = \sum_{b=1}^B \sum_{t=1}^T \sum_{z=1}^Z SHA - 256 \sim t, z, pk. \quad (1)$$

Eq. (1) determines the encryption of all datasets data based on SHA-256 [42].

$$Dec = \sum_{b=1}^B \sum_{t=1}^T \sum_{z=1}^Z SHA - 256 \sim Enc, t, z, pv. \quad (2)$$

Table 1

Notations and definition.

Notations	Definitions
CNN	Convolutional Neural Networks
RPC	Remote Procedure Call
T	Number raw source data of water
F, f	Features of dataset and particular feature T
M	Features of dataset Z
t	Particular water ingredient data
Z	Total number of sensors data
z	Particular sensor
S	Number of physical computing nodes
s	Particular physical computing nodes
ζ_s	Computing capability of node
ϵ_s	Resource capability of node
DT	Number of virtual DT computing nodes
s	Particular virtual DT computing nodes
ζ_{dt}	Computing capability of DT node
ϵ_{dt}	Resource capability of DT node
L	Number of virtual DT computing nodes
l	Particular virtual DT computing nodes
ζ_l	Computing capability of DT node
ϵ_l	Resource capability of DT node
δ_p	Determine the distance between different nodes
B, A, bid	Number of blockchain blocks and Attributes and blockid
N, n, b_{up}, b_{down}	Network of communication networks and particular network bandwidth
b_a	Attributes of particular blockchain block
pk, pv, ch, phs	Public key, private key, current hash and previous hash

Eq. (2) determines the decryption of all datasets data based on SHA-256 [42].

The data validation among nodes is determined in the following way.

$$Validation = \sum_{s=1}^S \sum_{dt=1}^{DT} \sum_{l=1}^L ch \sim s, dt, l \leftarrow ch \sim s, dt, l, \quad t = 1, \dots, T, z = 1, \dots, Z. \quad (3)$$

Eq. (3) determines the validation of data transactions of all datasets between nodes based on current and previous hash transactions. The blockchain validation determines for all nodes in the following way.

$$bc = Enc + Dec + Validation. \quad (4)$$

Eq. (4) determines the encryption, decryption and validation between nodes.

The physical servers process the sensor datasets data in the following way.

$$P_s = \sum_{s=1}^S \sum_{t=1}^T \sum_{z=1}^Z \left(\frac{t, z}{\zeta_s} \right) + bc, \quad t = 1, \dots, T, z = 1, \dots, Z. \quad (5)$$

Eq. (5) determines the physical server processing for all datasets. The sensor communication processing and distance calculation are determined in the following way.

$$\delta_{distance} = \sum_{n=1}^N \sum_{t=1}^T \sum_{z=1}^Z \left(\frac{n, t, z}{bw_{up}} \right) + \left(\frac{n, t, z}{bw_{down}} \right), \quad t = 1, \dots, T, z = 1, \dots, Z. \quad (6)$$

Eq. (6) determines the communication time and distance during reconvening and sending data between sensors and nodes.

$$P_{dt} = \sum_{dt=1}^{DT} \sum_{t=1}^T \sum_{z=1}^Z \left(\frac{t, z}{\zeta_{dt}} \right) + bc, \quad t = 1, \dots, T, z = 1, \dots, Z. \quad (7)$$

Eq. (7) determines the DT virtual server processing on different nodes.

$$P_l = \sum_{l=1}^L \sum_{t=1}^T \sum_{z=1}^Z \left(\frac{t, z}{\zeta_l} \right) + bc, \quad t = 1, \dots, T, z = 1, \dots, Z. \quad (8)$$

Eq. (8) determines the local processing in different institutions. The total processing time is calculated as follows.

$$Total = \sum_{t=1}^T \sum_{z=1}^Z P_s + P_{dt} + P_l + \delta_{distance}. \quad (9)$$

Eq. (9) determines the total processing for all datasets in architecture.

Proposed algorithm methodology: CNN-LSTM and BCTD Schemes

The study integrates techniques, including CNN, LSTM, and blockchain, to handle water management's diverse data modalities (e.g., image, text, numeric data) in IloWT architecture [45–47]. We combine both CNN-LSTM to process the diversity of data generated from different sensors on the different computing nodes. The main reason is that CNN extracts the features of data, and LSTM sequences the data properly. In IloWT architecture, the computing nodes, such as physical servers, virtual servers, and end-user nodes, the BCTD scheme, share the validated data among nodes during processing in IloWT.

To address this, we employ a combination of methods in our approach, enabling data feature extraction of sensor data using CNN. The sensors' data requires capturing data dependencies. There-

fore, we exploit LSTM for time series analysis and ensuring data validation across nodes (leveraging blockchain). The CNN scheme is utilized for extracting data features, while another scheme, like LSTM, captures data dependencies among various data points. Additionally, considering the heterogeneity of the computing nodes involved in processing, analyzing, and monitoring water from different perspectives, we propose data validation based on blockchain technology. This work showcases the methodology, highlighting the distinctions. The study illustrates two methodologies: the Flow Management of CNN-LSTM and BCTD Schemes. These schemes are integrated and operate collaboratively in distinct steps. The process begins with the data modality, which encompasses various data sensors. The data is then processed using CNN to extract the relevant features. In addition, for capturing data dependencies, the output layer of CNN serves as the input for LSTM. The LSTM module accumulates the results and shares them across different nodes based on the blockchain scheme implemented in our work.

Data sensors modality process based on CNN and LSTM Schemes

The combination of CNN and LSTM serves different purposes. We utilize the CNN scheme to extract and identify water level, leakage, and flow in various nodes. On the other hand, LSTM is employed to monitor water quantity and analyze its gradients

within the architecture. We provide a detailed description of this scenario of the proposed methodology. Initially, we monitor water activity and data from different sensor modalities, denoted as T . The CNN scheme extracts features from diverse datasets, such as T . The output of the CNN model serves as the input for LSTM. This allows LSTM to analyze the data sequence based on time series analysis within the architecture. Please note that the mentioned description is a general example, and the details may vary depending on the context and data being analyzed.

We are currently monitoring the water pump components and collecting data from various sensors, represented as Z . Due to heterogeneous components in the IIoWT architecture, we use the common runtime platform protocol such as X86 for communication. CNN is responsible for extracting features from different types of inputs, including images, numeric data, and text data from diverse sensors. LSTM is used to generate sequences based on the output of CNN. For example, a water pump is depicted with a blue color and consists of multiple components. The water tank is also connected to various sensors and provides output accordingly. The storage tank contains 30,000 liters of water and includes parameters such as pH level, chloride concentration, and others. The water is distributed to different residential houses, institutions, and hospitals. The CNN schemes analyze the features of various data inputs, denoted as T and Z . Each CNN scheme

Algorithm 1. CNN-LSTM data feature extractions

Algorithm 1: CNN-LSTM Data Feature Extractions

Input : S, DT, T, Z

Output: $R[]$

```

1 begin
2   for  $(T, Z)$  do
3     Generate the matrix label based on Word2VEC [45]. Initialize 3D matrix of sensors dataset
       features ;
4     for  $t=1, z=1, to T, Z$  do
5       Determined the physical server processing based on equation (5);
6       extract  $t.F, z.M$ ;
7       Append  $R[t, f, z, m]$ ;
8        $R[] \leftarrow$  Train and Test;
9       Apply LSTM scheme;
10      Exploited water pre-trained time series model [46]. Apply backpropagation network based on
        Vanish gradient problem [47]. for  $R[t, f, z, m]$  do
11        Input state, hidden state, and output states;
12        previous timestep;
13         $CR[t, f, z, m] \leftarrow R[t, f, z, m]$ ;
14        current timestep;
15      Generated Commutative results;
16    Call Blockchain Validation Schemes for storage and offload;
17 End Main;
```

incorporates distinct hyper-parameters, including filters, padding, and size. CNN and LSTM are supervised learning algorithms that process different data sources across layers. In our methodology, we employ pre-trained data models [43] for CNN and avoid any data pre-processing. This approach allows us to handle non-linear data with diverse formats. The CNN architecture consists of multiple layers, with the convolutional layer of size 28×28 responsible for extracting features based on the trained and tested models. These features are then passed to the pooling layer. The fully connected layers and softmax activation function [44] generate the final feature results.

The combination of CNN and LSTM Algorithm 1 incorporates various steps to achieve optimal results. The CNN's convolutional layer, pooling layer, and output layer are combined with the different states of LSTM. This integration enables the determination of water flow, analysis of water quality and its constituents, identification of water flow patterns, and detection of component leakage in real-time within distributed computing components. In this paper, we present a combined CNN and LSTM Algorithm 1 that involves the following steps:

- The combined algorithm schemes take several parameters as inputs, such as S, DT, T, Z .
- The initial results, denoted as R , are generated by CNN. These results involve the extraction of features from datasets based on pre-trained models.
- Steps 1 to 3 focus on extracting features and labeling them on 3D images. For numeric and text data, matrixes are generated using the Word2Vec algorithm.
- Steps 4 to 8 involve appending all the generated features and storing them into the initial results generated by CNN-LSTM, represented as $R[t, f, z, m]$. The results are appended using the operation "Append $R[t, f, z, m]$ " and then undergo training and testing.
- For handling the sequence of water quantity, LSTM is applied as described in step 9.
- From steps 9 to 14, the existing trained and tested model is utilized.
- We consider this problem a vanishing gradient problem, where backpropagation aims to achieve minimum errors in different time series analyses.
- Different states occur during the LSTM process, including the input, hidden, and output.
- The previous time stamp and current timestamp are utilized to obtain accurate results from the inputs provided by CNN.
- The algorithm produces results with cumulative accuracy, denoted as $CR[t, f, z, m]$, as shown in steps 15 to 17.

Blockchain transactional data validation (BCTD)

This paper's computing nodes encompass various data processing and analysis entities, including physical servers, digital twin servers, and end users' computing interfaces. Physical and virtual servers are responsible for processing data modalities, leveraging their respective resource capabilities and computing power. On the other hand, user end nodes solely receive data based on the validation scheme. To ensure interoperability across different nodes, we have designed a runtime environment based on MQTT for all cross-platform computing nodes. As mentioned earlier, our approach involves processing data modalities and fusion, employing a combination of CNN and LSTM for monitoring water level feature extraction and sequencing data dependencies in the algorithm. Physical servers are established exclusively at the power plant, while identical virtual servers are deployed at various monitoring institutions, such as government water municipal

institutes. To ensure the scalability and effective deployment of the IIoWT architecture as the complexity and scale of the water management network grow over time. We set the deadlines for each transaction in the blockchain network as shown in Algorithm 2. The main concern is ensuring the resource's scalability and overall execution time (e.g., computation time and communication time) must be less than the given deadlines on all transactions. All nodes have sufficient resources for scaling up and down in the blockchain networks. Consequently, physical servers share all processing data with virtual servers in different locations. Additionally, the water distributor digitally shares water quality data with end users based on their applications and interfaces throughout the processing phase within the architecture. However, ensuring security is crucial when dealing with heterogeneous nodes. Therefore, we propose utilizing blockchain technology to generate hash data before offloading it to another node, employing asymmetric SHA-256 hashing schemes.

To ensure secure communication, we implemented a combination of public and private keys for encryption and decryption, allowing for the transformation of data into readable forms at different computing nodes. In addition, we developed the Blockchain Transaction Data (BCTD) scheme, where encryption and decryption processes on current nodes are treated as blocks. Each block contains essential information such as a unique identifier (*bid*), current hashing (*ch*), previous hashing (*phs*), data transactions, offloading details, validation rules, and acknowledgments during the validation process between data transfers among nodes. This scheme transforms data transactions into a transparent and immutable format, enhancing the security and efficiency of our study. The study introduces the BCTD Algorithm 2 for data validation among nodes in the blockchain, which encompasses the following steps.

- The blockchain scheme utilizes the input from the combined results obtained from the CNN and LSTM schemes on the given data modalities within the architecture. For example, the blockchain input is $CR[t, f, z, m], B$ across all blockchain blocks.
- Steps 1 to 8 illustrate the blockchain's encryption, decryption, and data validation phases, implemented across different nodes.
- Steps 9 to 13 involve data validation in its immutable form, based on predetermined rules.
- These rules consist of generated hashing patterns, where the miner node establishes the blockchain rules for the other nodes.
- Steps 10 to 21 demonstrate the processing of blockchain data at various nodes, following the predefined rules and utilizing the current and previous hash values.
- The miner node, which in our case is the government institution, holds the authority to add or remove blocks within the blockchain architecture.
- Steps 22 to 24 outline the matching of each transaction across nodes, using a designed proof-of-work pattern and rules during offloading and generating acknowledgments if the transaction validations are successful.
- In cases where transactions are deemed invalid, the blockchain identifies them as suspicious within the architecture.

Time complexity and space complexity of CNN-LSTM for considered problem

In this study, we exploited different schemes such as CNN, LSTM, and blockchain validation for the data standardization process. We determined the time complexity of CNN based on $N \times N$. The N shows the number of convolutional layers, pooling layer, fitness, and sigmoid function of the processing. The LSTM has three distinct gates: input, output, and forget gates. Therefore, time complexity becomes N^3 . The blockchain has many blocks for

Algorithm 2. BCTD scheduling scheme**Algorithm 2:** BCTD Scheduling Scheme

```

Input :  $CR[t, f, z, m], B;$ 
1 begin
2   for  $(S, DT, L)$  do
3     For blockchain validation at physical before offloading;
4     Determined the physical server processing based on equation (5);
5     Apply the encryption on the commutative result data based on equation (1);
6      $\sum_{b=1}^B \sum_{t=1}^T \sum_{z=1}^Z SHA - 256 \sim t, z, pk;$ 
7     Apply the decryption on the commutative result data based on equation (2);
8      $\sum_{b=1}^B \sum_{t=1}^T \sum_{z=1}^Z SHA - 256 \sim Enc, t, z, pv;$ 
9     Determined the data validation between nodes based on (3);
10     $\sum_{s=1}^S \sum_{dt=1}^{DT} \sum_{l=1}^L ch \sim s, dt, l \leftarrow ch \sim s, dt, l, \quad t = 1, \dots, T, z = 1, \dots, Z;$ 
11    Apply the blockchain scheme and determined based on equation (4);
12    Determined the blockchain validation on virtual server based on equation (7);
13     $Enc + Dec + Validation;$ 
14    Determined the data transactions between based on equation (6);
15    for  $(bid \text{ as } B)$  do
16      Rules;
17      if  $b1.CR[t, f, z, m], B \leftarrow s1 \sim s2, dt, l$  then
18        Validated the transactions between nodes;
19        Offloading, Validation, Acknowledgment;
20        Determined total processing based on equation (9);
21         $\sum_{t=1}^T \sum_{z=1}^Z P_s + P_{dt} + P_l + \delta_{distance};$ 
22      else
23        Validated the transactions between nodes;
24        Offloading, Validation, Acknowledgment;
25      End validation;
26    End blockchian;
27  End Offloading;
28 End Main;

```

validation. Therefore, the time complexity is determined in $O(N)$ number of blocks for data standardization on heterogeneous nodes. The space complexity is determined in heterogeneous nodes based on available resources in IIoWT architecture. These classes include raw water ingredient monitoring, supply tank management, storage tank management, water flow control, water motor pump distribution, and water volume calculation based on data collected from different sensors. The digital twin system involves physical and virtual components, utilizing Amazon Lambda services. It incorporates blockchain technology to enhance

the security and transparency of the virtual water management process. The physical and virtual server resources are determined in ϵ_s and ϵ_{DT} .

Performance evaluation and implementation

In this part, we designed the simulation environment based on various components. We mainly focused on the IIoWT (Industrial Internet of Water Things) architecture components we prepared for the simulation. We compared these components with existing

architectures, algorithms, data validations, and performance for distributed water management in smart cities. The IloWT architecture comprises various components, including sensors, computing nodes, wireless communication, and different programming application interfaces (API). We defined and listed the simulation parameters in Table 2.

In IloWT, different types are implemented with different characteristics. Table 2 displays the configuration of nodes in the IloWT architecture. It encompasses various computing nodes, including physical servers, digital twin virtual servers, and end-user interfaces. All the nodes in the architecture utilize Amazon’s virtualization technology. The physical and digital twin servers have adaptive and scalable storage, allowing for flexible expansion or reduction based on specific resource requirements. On the other hand, the end-user devices or interfaces are fixed and possess resource-constrained characteristics.

IloWT architecture implementation based on Amazon baseline architecture

The open-source code for industrial IoT digital twin in water management, designed for simulation and design purposes, is available on GitHub by Amazon. This code provides a baseline application programming interface (API) that comprises various classes essential for water management, as shown in Fig. 2. These classes include raw water ingredient monitoring, supply tank management, storage tank management, water flow control, water motor pump distribution, and water volume calculation based on data collected from different sensors. The digital twin system involves physical and virtual components, utilizing Amazon Lambda services. It incorporates blockchain technology to enhance the security and transparency of the virtual water management process. Additionally, different sensors within the system generate

Table 2
Experiment configuration.

Parameter	Description	Parameter	Description
PH, chlorine	METTLERS	Hardness,Solid	Ultrasonic Sensor
Sulfate, Conductive	HYDROS, Sonar	Organic-carbons	Sonars
Trihalomethanes,Turbidity	Sonars	Potability	Portable sonars
Raspberry-Pi	Water Pump	IoT-Clock	GPIO-Module
Languages	JAVA, Python, XML	Ineroperateability	MQTT Client Server Broker
Layer (type)	Output Shape	Param	Learning Rate
input(Input Layer)	(None, 224, 224, 1)	0	0.5
model-T (Model)	(None, 7, 7, 1)	492897	0.3
model-Z (Model)	(None, 224, 224, 1)	62711	0.5
Total params	555,608	Trainable params	553,702
Non-trainable params	1,906	Convolutional layer	128×128

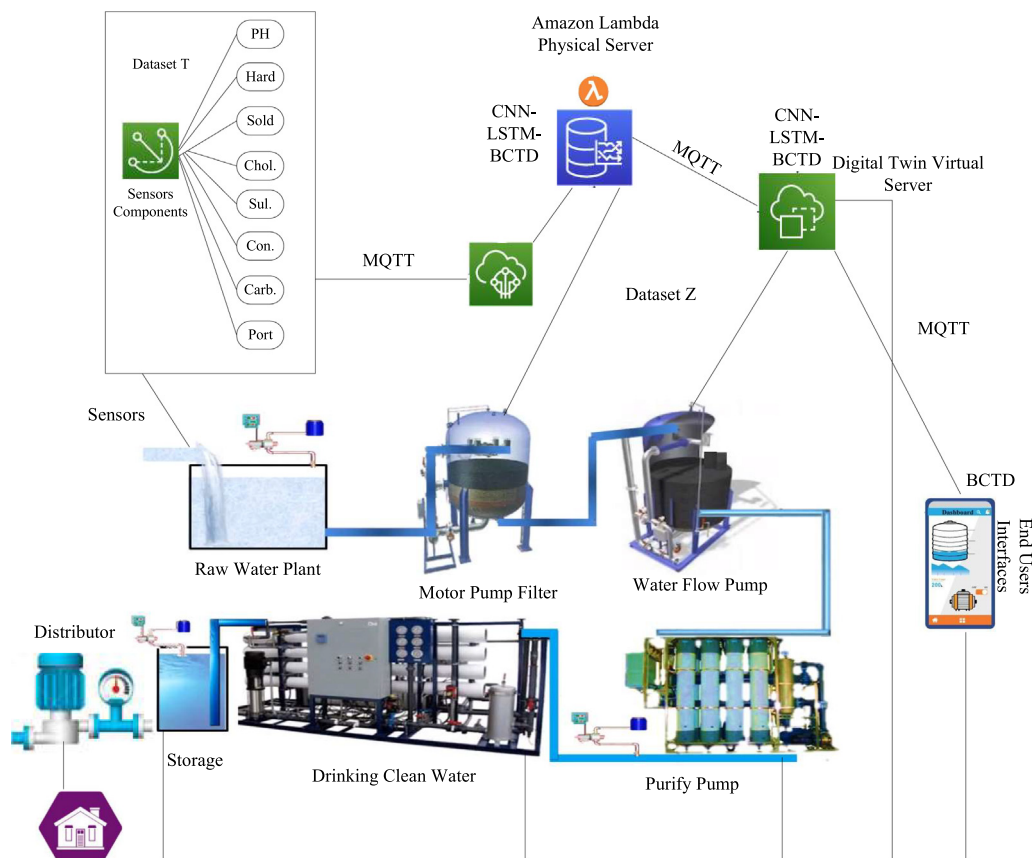


Fig. 2. IloWT implementation with different components.

images to detect and monitor water leakage. Overall, the digital twin code developed by Amazon for water management demonstrates the integration of virtual and physical components, leveraging blockchain technology to ensure efficient and secure operations in water management processes. The code is publicly available on Github for further exploration and utilization. The Amazon Cloud Development Kit (CDK) is a simulator that enables the creation of virtual physical servers and digital twin servers. This functionality is made possible through the Amazon Virtual Lab, which provides the necessary hardware and software resources. It offers an open environment where mobile applications for industrial water management can be seamlessly connected to mobile devices. The environment is based on Docker containers, allowing the encryption of complete image data within blockchain blocks within the IIoWT architecture. In our research, we utilized two datasets that had been trained and tested: one focused on water quality and the other on water leakage detection. In our study, we utilized trained and tested datasets for our analysis. The CNN-LSTM algorithm collected data from sensors and made decisions based on the models that had been trained and tested. Blockchain technology was implemented on both physical and virtual digital twin servers. However, in the case of end-user applications or interfaces, blockchain was solely utilized to ensure the validity of drinking water data within the home environment. All nodes within the system were interconnected through distributed networks, ensuring interoperability and seamless communication between them.

Water sensors datasets benchmark

To evaluate existing studies' performance, we exploited the same datasets with different features to compare the proposed work with the existing studies. The datasets consisted of different sensor data as the benchmark with the data modality as shown in Table 3. We applied blockchain technology to sensor data based on the SHA-256 scheme to collect valid sensor data among fog and cloud-enabled water components. Drinking water plays a vital role in human life, and therefore, government healthcare agencies always prioritize ensuring clean drinking water in smart cities. At national and international levels, water is generated from raw materials such as seawater and other resources, collected at power plants, and treated following predefined government regulations. Consequently, it is imperative to adhere to these regulations to

purify drinking water at the power plants. To monitor compliance with these regulations, we continuously monitor the readings of various sensors in real-time using other computing nodes. We collected the real-time water quality and monitoring data in different time zones with different parameters. We generated sensor values using various hyperparameters, such as Sensor, Count, Mean, STD (standard deviation), Min (minimum), 12 am-8 am, 9 am-5 pm, 6 pm-1.55 pm, and Max (maximum). Among these sensors, the pH sensor is of utmost importance as it measures the acidity level in water. As per the guidelines set by the government and the World Health Organization, the pH values must fall within the range of 6 to 9. Hardness in water consisted of different gradients such as magnesium and calcium salts. These salts are observed in water randomly during flows in various sources. Solid in water consisted of other ingredients. For instance, bicarbonates, chlorides, sulfates, calcium, magnesium, sodium, and potassium. Therefore, it has a standardized threshold to make the water into drink form. For instance, 499 mg/l to 999 mg/l were incorporated in water during processing at the power plants. The study adjusted the chlorine ratio to clean water at a concentration of 4 to 6 milligrams per liter. A certain percentage of ammonia was added to the water to render it suitable for drinking. Additionally, the water was supplemented with a sulfate level ranging from 2500 to 3000 milligrams per liter. The water leakage datasets included images such as Raw Water Plant, Motor Pump Filter, Water Flow Pump, Purify Pump, Drinking Clean Water, and Storage. The water leakage and flow and storage level of water tanks are determined and monitored in the IIoWT based on trained and tested models of dataset Z.

Result analysis and discussion

The study assessed the performance of existing industrial IoT water management architectures and compared them with the proposed IIoWT architecture, which utilized different components. Baseline approaches, namely IoT-Water [13–15] and random IoT-Water (RIoT-Water) [16–19], were considered in this evaluation. These baseline architectures were employed to manage water quality in smart cities' drinking water distribution. The architectures comprised computing nodes, sensors, wireless networks, and raw material sources for water, all simulated in the study. The architecture played a crucial role in determining the flow of raw water, the ingredients present in it, and the additional ingredi-

Table 3
Data modality samples from different sensors for water quality.

Sensor	Count	Mean	STD	Min	12am-8am	9am-5 pm	6 pm-11.55 pm	Max
t1	310112.0	3.372221	0.512227	0.500000	1.438831	3.456539	3.499826	1.549016
t2	319951.0	27.591611	2.296666	1.000000	26.310760	28.133678	39.479160	26.727430
t3	090301.0	3.8392	1.88820	7.159720	8.0620	9.9300	23.57770	2.032990
t4	10301.0	1.752481	1.418887 1	5.640620	32.838539	12.227428	19.2500	12.220490
t5	9301.0	6.673936	4.023912	0.798032	0.620400	0.638916	0.615723	10.000000
t6	490301.0	1.396414	0.298247	0.44440	1.976260	0.6790	1.912150	0.9998
t7	7890301.0	2.396414	0.558247	0.67440	0.976260	0.99790	0.912150	0.9998
t8	84869.0	0.843152	0.201155	1.0000	1.907120	5.167530	0.7950	1.596640
t9	111213.0	4.200721	2.037390	0.8935	12.183740	13.494790	11.697340	15.348960
t10	31213.0	2.200721	5.037390	0.3935	1.183740	7.494790	7.697340	2.348960
t11	41213.0	3.200721	7.037390	0.4935	2.183740	5.494790	8.697340	3.348960
t12	71213.0	5.200721	6.037390	0.5935	17.183740	7.494790	3.697340	4.348960
t13	21213.0	6.200721	7.037390	0.6935	10.183740	1.494790	6.697340	6.348960
t14	231213.0	8.200721	5.037390	0.9935	9.183740	2.494790	8.697340	7.348960
t15	3241213.0	9.200721	3.037390	0.7935	8.183740	3.494790	1.697340	9.348960
ph	Hard	Solids	Chloramines	Sulfate	Conduct	carbon	methanes	Turbidity
0.1	1890455	2.318981	1.300212	123.516441	8.308654	1.379783	8.990970	3.963135
1.716080	45.422921	5450.057858	2.635246	0.1	2.885359	1.180013	12.329076	1.500656
2.099124	12.236259	5454.541732	3.275884	0.1	46213	16.868637	66.420093	3.055934
3.316766	45.373394	5644.417441	4.059332	4.886136	1.266516	18.436524	5.341674	4.628771
5.092223	18.101509	342.986339	2.546600	3.135738	12.410813	1.558279	4.997993	4.075075

ents added to the water plant. In the experiments, we implemented component calibration using multi-ANOVA testing techniques. The random data and solution were assumed to follow normal distributions.

We compared IoT water architectures using different sensor datasets, namely T and Z . These sensors' values were collected and presented in Table 3. In Fig. 3, the x-axis represents the sensor data values ranging from 100 to 1200, while the y-axis displays the total processing time involved in receiving raw data at the power plant and processing it using the given dataset in the architectures. This figure illustrates the performance evaluation of the architectures when water, obtained from various sources, undergoes processing based on different inputs outlined in Table 3. Remarkably, the IloWT architecture outperformed all existing water architectures regarding raw data processing. This superiority can be attributed to monitoring water flow at different intervals, considering the availability of water flow in various time zones. We can effectively minimize delays by adjusting the water flow and raw material mechanism during nighttime compared to daytime. Consequently, IloWT demonstrated superior performance in reducing total delays, measured in seconds, compared to other existing architectures.

The powerful physical servers at the power plant retrieve data from various sources to transform raw water into drinkable form. These physical servers are solely responsible for executing all dataset tasks related to data modality, as depicted in Fig. 3. With minimal processing time, measured in seconds, these physical servers efficiently handle the data modality data, even when presented in different formats. In Fig. 3, the proposed IloWT architecture surpasses all other existing water architectures in data modality processing while adhering to the established standards in distributed water management. In IloWT, the training and testing time is adaptive and provides enhanced accuracy for various modalities. By employing a CNN, the architecture effectively extracts water ingredient features from sensors and utilizes the trained and tested data for decision-making regarding water level and quality management. As illustrated in Fig. 3, IloWT demonstrates significantly improved training and testing validation times compared to existing IoT water and RIoT water architectures in physical and virtual server environments during the processing phase.

We have leveraged digital twin technology to create virtual servers for managing distributed water processing. This innovation has fostered seamless collaboration among government entities, companies, and users, enabling the sharing of data and components. Instead of relying on physical servers in different places, we have opted for virtual servers in the distributed water system. These integrated virtual servers, which are real and constantly updated, ensure the sharing of up-to-date data. In summary, digital twin technology facilitates the smooth sharing of components and data, leading to improved collaboration, real-time monitoring, enhanced efficiency, and better decision-making through simulations and predictive analysis. The virtual servers in digital twins replicate the physical servers and can be easily scaled up or down during dataset processing. In Fig. 3, it is evident that the processing and data sharing time of the digital twin virtual servers is significantly reduced compared to physical servers. Consequently, the IloWT architecture yields optimal results, and the successfully shared trained and tested models outperform existing IoT and RIoT water architectures.

In addition to physical and virtual servers, the system incorporates other nodes, such as end-processing nodes. The end-user interface retrieves information about drinking water from various datasets, which include gradient details. End users can access a limited set of dataset sensors, specifically catering to incoming water data. These sensors provide parameters such as water cleanliness, pH, salt content, chlorine level, and other relevant factors. Such data points assist end users in evaluating the effectiveness of water in maintaining good health, aligning with predefined guidelines. Fig. 4 showcases the IloWT architecture, specifically designed with distinct data modalities for end users. This design reduces processing time and real-time data updates compared to IoT architectures. Consequently, end users receive timely and accurate information about drinking water quality.

Standardizing data is crucial for effective distributed water management in smart cities. We have implemented data standardization across various nodes, including physical servers, virtual servers, and end users. Maintaining data processing and sharing accuracy is fundamental to the IloWT architecture. Data standardization ensures that data conforms to a predefined set of standards across different nodes while maintaining data privacy and security

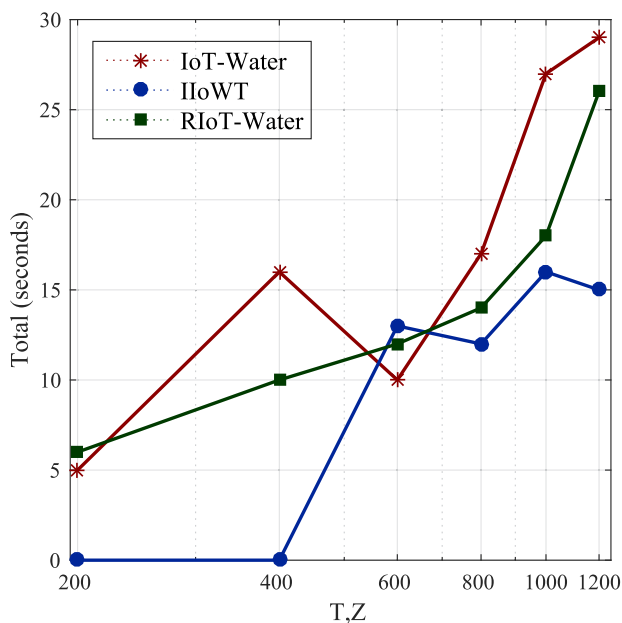


Fig. 3. Water plant trained and tested datasets processing total time (seconds).

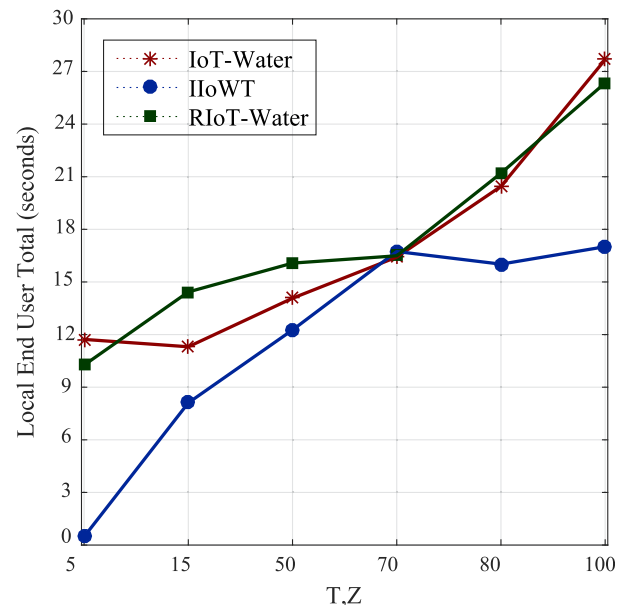


Fig. 4. Water drinking quality ingredients at end users interfaces with optimal processing.

(as depicted in Fig. 4). To assess the accuracy of data standardization, we conducted evaluations using different IoT architectures and varied datasets, such as *T* and *Z*. Fig. 4 demonstrates that IloWT achieves a high level of accuracy in data standardization and sharing, nearing 99.9%, surpassing existing architectures. This achievement can be attributed to the robust data validation processes and the interoperability of the architecture, which are optimized components of the IloWT framework. Standardizing data has different data formats and distributed water management in smart cities. We have implemented data standardization with the given datasets on different nodes, including physical servers, virtual servers, and end users. Maintaining data processing and sharing accuracy is fundamental to the IloWT architecture. Data standardization ensures that data conforms to a predefined set of standards across different nodes while maintaining data privacy and security (as depicted in Fig. 4). To assess the accuracy of data standardization, we conducted evaluations using different IoT architectures and varied datasets, such as *T* and *Z*. Fig. 4 demonstrates that IloWT achieves a high level of accuracy in data standardization and sharing, nearing 95% to 99%, surpassing existing architectures. These optimal results and robustness in data validation processes are obtained due to the interoperability of the architecture, which are optimized components of the IloWT framework.

Interoperability is the seamless cross-platform in heterogeneous nodes, where technologies and protocols are suggested in existing IoT water architectures. The existing interoperate technologies, such as socket programming interface [48] and remote procedure call [49] were widely exploited in cross-platform water with heterogeneous nodes. These are client and server message and data broker call protocols, where different data formats and data standardization are followed by nodes. However, socket and RPC are slow when they travel and offload the different formats in IoT water architectures. Fig. 5 (a) shows the accuracy of data sharing with the different formats among physical servers and virtual servers during water monitoring and processing in the architectures. Fig. 5 (b) shows the data format standardization among physical and virtual and end users interface with the higher accuracy with the MQTT broker application programming interface. Fig. 5 shows the proposed MQTT interoperability with IloWT architecture has more optimal results than socket and RPC protocols. The main reason is that the existing protocols only share and convert the data into one form, such as extensible markup language and java object notation protocols. Therefore, they lose their accuracy when different data formats are shared among nodes. Therefore, MQTT obtained the optimal accuracy as compared to existing protocols. The research incorporated blockchain validation at the end of the CNN-LSTM process. Data encryption and decryption were performed using SHA-256, and transaction validation

relied on proof of work validation across different nodes. Fig. 6(a) illustrates the three security schemes employed in IoT water architectures. Specifically, the integration of CNN-LSTM security and machine learning (ML) blockchain was explored in IoT water management to analyze data security through blockchain technology.

By integrating machine learning techniques, blockchain technology becomes more adaptable and efficient in validating data formats and hashes among nodes. Throughout the experiment, the transparency and immutability of shared data demonstrated higher accuracy when utilizing the CNN-LSTM-BCTD scheme. Fig. 6(a) showcases the blockchain validation process between virtual servers and physical servers after implementing the CNN-LSTM schemes on the data before offloading them to each other. The study involved executing 1200 transactions per block, and the results displayed in Fig. 6(a) indicated that the CNN-LSTM-BCTD approach achieved an optimal accuracy of approximately 99% compared to existing security and blockchain technologies. Moreover, the CNN-LSTM-BCTD method attained higher accuracy in data validation among servers and end-user interfaces, as depicted in Fig. 6(b). CNN-LSTM-BCTD achieves higher data validation among different nodes compared to existing studies due to several key factors. One crucial aspect is the validation process through multiple interrogations within the CNN-LSTM-BCTD framework. By leveraging this approach, the miner node has the capability to add and delete blockchain blocks, ensuring the integrity of the transactional data. An important feature integrated into the CNN-LSTM-BCTD approach is validating the "vanish ingredient" transaction. This validation process involves verifying each node's original and transactional data before transferring them to other nodes. By implementing this step, the proposed method significantly enhances data validation across different nodes. As a result, CNN-LSTM-BCTD offers superior data validation among various nodes compared to existing studies. Its rigorous validation process and inclusion of the "vanish ingredient" transaction verification contribute to the higher data integrity and accuracy achieved in the proposed method.

Finding and Limitation

We presented the IloWT architecture incorporating blockchain and digital twin technologies and processed the data modality using the CNN-LSTM approach. The main finding of IloWT architecture is to examine the existing IoT water systems and discuss their limitations. We presented the simulation parameters and described the implementation of IloWT. We discussed each deployment step in detail. Furthermore, we analyzed the obtained results, demonstrating the efficiency and effectiveness of the proposed IloWT and CNN-LSTM-BCTD schemes for water management

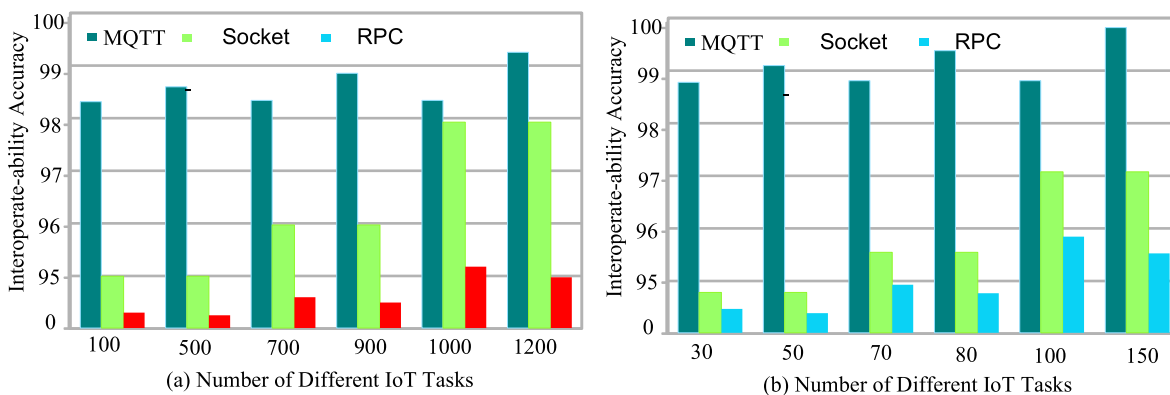


Fig. 5. Accuracy of interoperability for water quality and level monitoring in physical, virtual and end users nodes.

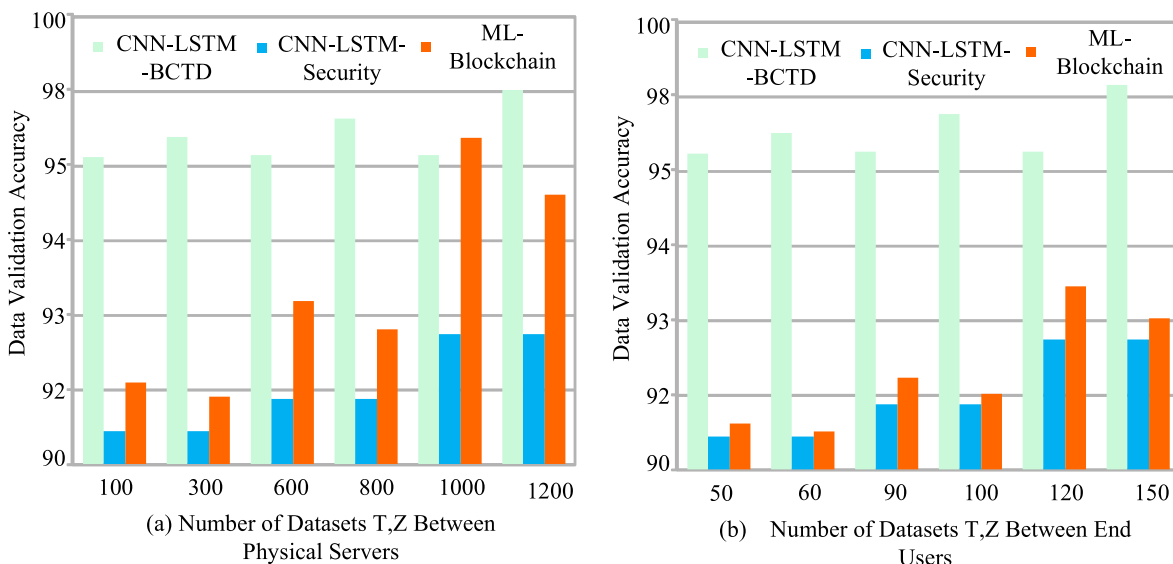


Fig. 6. Data validation and data standardization water quality and level monitoring in physical, virtual and end users nodes based on blockchain.

across various computing nodes. There are many limitations of the proposed IloWT architecture. In our proposed IloWT architecture, we did not consider the failure ratio of different components in the large and distributed setup. The work did not consider the motor’s energy and power consumption of computing nodes.

Conclusion and future work

In this paper, we have presented the IloWT architecture incorporating blockchain and digital twin technologies and processed the data modality using the CNN-LSTM approach. We aimed to address the research challenges by formulating relevant research questions and highlighting the contributions of this study. We thoroughly examined the existing IoT water systems and discussed their limitations. We also provided a comprehensive overview of the problem formulation and proposed algorithm, detailing the steps and scenarios involved. In the evaluation section, we presented the simulation parameters and described the implementation of IloWT. We discussed each deployment step in detail. Furthermore, we analyzed the obtained results, demonstrating the efficiency and effectiveness of the proposed IloWT and CNN-LSTM-BCTD schemes for water management across various computing nodes.

However, there is still considerable room for improvement in the existing IloWT architecture. The current system faces challenges such as an increasing failure ratio of different components in the large and distributed setup. The processing cost has also become prohibitively high due to the numerous components involved. We will address these issues by considering new constraints such as fault tolerance, processing cost optimization, and data safety within the proposed IloWT architecture. These improvements will enhance the reliability and efficiency of the system, ensuring its smooth operation in the future. We will also improve the motor’s energy and power consumption of computing nodes in the proposed architecture IloWT.

Code and data availability

The study exploited the existing Amazon industrial IoT water management simulator and integrated the digital twin and blockchain to meet the requirements of data security and standardization in IloWT. The source code is available on the following URL.

<https://github.com/aws-samples/water-tank-digital-twin>. The study exploited water quality and water leakage datasets available on the following url. <https://www.kaggle.com/code/miklgr500/leakage-detection-about-8-test-dataset/input>.

The water leakage dataset is available on the following URL. <https://www.kaggle.com/datasets/adityakadiwal/water-potability>.

Funding

This article was co-funded by the European Union under the REFRESH -Research Excellence For REgion Sustainability and High-tech Industries project number CZ.10.03.01/00/22_003/0000048 via the Operational Programme Just Transition. Also, this work was supported by the Ministry of Education, Youth and Sports of the Czech Republic conducted by VSB -Technical University of Ostrava, Czechia under Grants SP2023/039 and SP2023/042.

CRediT author statement: Mazin: Idea Building, Abdullah: Methodology, Simulation, Karrar: Data curation, Khanapi: Writing-Original Manuscript preparation, Haydar: Visualization, Kadyr: Simulation and Experiment Design, Jan and Radek: English Grammar and Editing, Begonya: Cloud Service Validation.

Compliance with ethics requirements

This article was supported by the Ministry of Education of the Czech Republic (Project No. SP2023/039 and No. SP2023/042).

This article does not contain any studies involving human participants performed by any authors.

This article does not contain any studies involving animals performed by any of the authors.

CRediT authorship contribution statement

Mazin Abed Mohammed: Conceptualization, Methodology, Formal analysis, Writing – Review & Editing. **Abdullah Lakhani:** Conceptualization, Methodology, Formal analysis, Writing – Review & Editing. **Karrar Hameed Abdulkareem:** Conceptualization, Methodology, Formal analysis, Writing – Review & Editing. **Mohd Khanapi Abd Ghani:** Formal analysis, Visualization, Supervision. **Haydar Abdulameer Marhoon:** Formal analysis, Visualization.

tion, Supervision. **Sefedine Kadry:** Formal analysis, Visualization, Supervision. **Jan Nedoma:** Project administration, Funding acquisition. **Radek Martinek:** Project administration, Funding acquisition. **Begonya Garcia Zapirain:** Formal analysis, Visualization, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Júnior ACDS, Munoz R, Quezada MDLÁ, Neto AVL, Hassan MM, De Albuquerque VHC. "Internet of water things: a remote raw water monitoring and control system. *IEEE Access* 2021;9:35 790–800.
- Suresh A, Nandagopal M, Raj P, Neeba E, Lin J-W. *Industrial IoT application architectures and use cases*. CRC Press; 2020.
- Khan S, Altayar M. *Industrial internet of things: Investigation of the applications, issues, and challenges*. *Int J Adv Appl Sci* 2021;8(1):104–13.
- Sundaresan S, Kumar KS, Kumar TA, Ashok V, Julie EG. Blockchain architecture for intelligent water management system in smart cities. In: *Blockchain for smart cities*. Elsevier; 2021. p. 57–80.
- Xia W, Chen X, Song C. A framework of blockchain technology in intelligent water management. *Front Environ Sci* 2022;10:753.
- Bhattacharya S, Victor N, Chengoden R, Ramalingam M, Selvi GC, Maddikunta PKR, Donta PK, Dustdar S, Jhaveri RH, Gadekallu TR. Blockchain for internet of underwater things: State-of-the-art, applications, challenges, and future directions. *Sustainability* 2022;14(23):15659.
- Agamal ZY, Abonazel MR, Lukman AF. Modified jackknife ridge estimator for beta regression model with application to chemical data. *Int J Math, Stat, Comput Sci* 2022;1:15–24.
- Sasikumar A, Vairavasundaram S, Kotecha K, Indragandhi V, Ravi L, Selvachandran G, Abraham A. Blockchain-based trust mechanism for digital twin empowered industrial internet of things. *Future Gener Comput Syst* 2023;141:16–27.
- Suhail S, Hussain R, Jurdak R, Hong CS. Trustworthy digital twins in the industrial internet of things with blockchain. *IEEE Internet Comput* 2021;26(3):58–67.
- Sasikumar A, Vairavasundaram S, Kotecha K, Indragandhi V, Ravi L, Selvachandran G, Abraham A. Blockchain-based trust mechanism for digital twin empowered industrial internet of things. *Future Gener Comput Syst* 2023;141:16–27.
- Feng X, Wu J, Wu Y, Li J, Yang W. Blockchain and digital twin empowered trustworthy self-healing for edge-ai enabled industrial internet of things. *Inf Sci* 2023;119169.
- Manoj M, Dhillip Kumar V, Arif M, Bulai E-R, Bulai P, Geman O. State of the art techniques for water quality monitoring systems for fish ponds using iot and underwater sensors: A review. *Sensors* 2022;22(6):2088.
- Jan F, Min-Allah N, Düşteğör D. Iot based smart water quality monitoring: Recent techniques, trends and challenges for domestic applications. *Water* 2021;13(13):1729.
- Pasika S, Gandla ST. Smart water quality monitoring system with cost-effective using iot. *Heliyon* 2020;6(7):e04096.
- Huan J, Li H, Wu F, Cao W. Design of water quality monitoring system for aquaculture ponds based on nb-iot. *Aquacult Eng* 2020;90:102088.
- Bhardwaj A, Dagar V, Khan MO, Aggarwal A, Alvarado R, Kumar M, Irfan M, Proshad R. Smart iot and machine learning-based framework for water quality assessment and device component monitoring. *Environ Sci Pollut Res* 2022;29(30):46 018–36.
- F.J.J. Joseph. "Iot based aquarium water quality monitoring and predictive analytics using parameter optimized stack lstm," in 2022 6th International Conference on Information Technology (InCIT). IEEE, 2022, pp. 342–346.
- Chopade S, Gupta HP, Mishra R, Oswal A, Kumari P, Dutta T. A sensors-based river water quality assessment system using deep neural network. *IEEE Internet of Things Journal* 2021;9(16):14 375–84.
- Loc HH, Do QH, Cokro A, Irvine KN. Deep neural network analyses of water quality time series associated with water sensitive urban design (wsud) features. *J Appl Water Eng Res* 2020;8(4):313–32.
- Kaur G, Braveen M, Krishnapriya S, Wawale SG, Castillo-Picon J, Malhotra D, Osei-Owusu J, et al. Machine learning integrated multivariate water quality control framework for prawn harvesting from fresh water ponds. *J Food Qual* 2023;2023.
- U.W. Chohan, "Blockchain and environmental sustainability: Case of ibm's blockchain water management", Notes on the 21st Century (CBRI), 2019.
- E.M. Dogo, A.F. Salami, N.I. Nwulu, and C.O. Aigbavboa, "Blockchain and internet of things-based technologies for intelligent water management system," *Artificial intelligence in IoT*, pp. 129–150, 2019.
- A. Poberezhna, "Addressing water sustainability with blockchain technology and green finance," in *Transforming climate finance and green investment with blockchains*. Elsevier, 2018, pp. 189–196.
- Zhang Y, Luo W, Yu F. Construction of chinese smart water conservancy platform based on the blockchain: technology integration and innovation application. *Sustainability* 2020;12(20):8306.
- Zhang Y, Luo W, Yu F. Construction of chinese smart water conservancy platform based on the blockchain: technology integration and innovation application. *Sustainability* 2020;12(20):8306.
- Pee SJ, Nans JH, Jans JW. A simple blockchain-based peer-to-peer water trading system leveraging smart contracts. In: *Proceedings on the International Conference on Internet Computing (ICOMP)*. The Steering Committee of The World Congress in Computer Science, Computer? p. 63–8.
- Li H, Chen X, Guo Z, Xu J, Shen Y, Gao X. Data-driven peer-to-peer blockchain framework for water consumption management. *Peer-to-Peer Network Appl* 2021;14:2887–900.
- M. Pincheira, M. Vecchio, R. Giuffreda, and S.S. Kanhere, "Exploiting constrained iot devices in a trustless blockchain-based water management system," in 2020 IEEE International Conference on Blockchain and Cryptocurrency (ICBC). IEEE, 2020, pp. 1–7.
- Liu Y, Shang C. Application of blockchain technology in agricultural water rights trade management. *Sustainability* 2022;14(12):7017.
- A. Predescu, D. Arsene, B. Pahon?u, M. Mocanu, and C. Chiru, "A serious gaming approach for crowdsensing in urban water infrastructure with blockchain support," *Applied Sciences*, vol. 11, no. 4, p. 1449, 2021.
- Xia W, Chen X, Song C. A framework of blockchain technology in intelligent water management. *Front Environ Sci* 2022;10:753.
- B. Pahon?u, D. Arsene, A. Predescu, and M. Mocanu, "Application and challenges of blockchain technology for real-time operation in a water distribution system," in 2020 24th International Conference on System Theory, Control and Computing (ICSTCC). IEEE, 2020, pp. 739–744.
- Chang Y, Xu J, Ghafoor KZ. An iot and blockchain approach for the smart water management system in agriculture. *Scalable Comput: Pract Exp* 2021;22(2):105–16.
- Li T, Wang Z, Wang C, Huang J, Zhou M. Chlorination in the pandemic times: The current state of the art for monitoring chlorine residual in water and chlorine exposure in air. *Sci Total Environ* 2022;838:156193.
- Joshi ID, Stramski D, Reynolds RA, Robinson DH. Performance assessment and validation of ocean color sensor-specific algorithms for estimating the concentration of particulate organic carbon in oceanic surface waters from satellite observations. *Remote Sens Environ* 2023;286:113417.
- Cui M, Sun Y, Huang C, Li M. Water turbidity retrieval based on uav hyperspectral remote sensing. *Water* 2022;14(1):128.
- Gao C, Huang C, Wang J, Li Z. Modelling dynamic hydrological connectivity in the zoigê area (china) based on multi-temporal surface water observation. *Remote Sens* 2021;14(1):145.
- Manjakkal L, Szwagierczak D, Dahiya R. Metal oxides based electrochemical ph sensors: Current progress and future perspectives. *Prog Mater Sci* 2020;109:100635.
- Patel C, Doshi N. A novel mqtt security framework in generic iot model. *Proc Comput Sci* 2020;171:1399–408.
- Cong S, Zhou Y. A review of convolutional neural network architectures and their optimizations. *Artif Intell Rev* 2023;56(3):1905–69.
- Nasser AA, Rashad MZ, Hussein SE. A two-layer water demand prediction system in urban areas based on micro-services and lstm neural networks. *IEEE Access* 2020;8:147 647–61.
- Fotohi R, Aliee FS. Securing communication between things using blockchain technology based on authentication and sha-256 to improving scalability in large-scale iot. *Comput Netw* 2021;197:108331.
- Tan T-H, Huang C-J, Gochoo M, Chen Y-F. Activity recognition based on fr-cnn and attention-based lstm network. 2021 30th Wireless and Optical Communications Conference (WOCC). IEEE; 2021. p. 146–9.
- Dong C-C, Tian M-H, Zhao W-C. Research on water quality prediction model of rf-softmax based on ga optimization. *Hubei Agric Sci* 2022;61(7):60.
- Di Gennaro G, Buonanno A, Palmieri FA. Considerations about learning word2vec. *J Supercomput* 2021;1–16.
- Yuan Y, Lin L, Liu Q, Hang R, Zhou Z-G. Sits-former: A pre-trained spatio-spectral-temporal representation model for sentinel-2 time series classification. *Int J Appl Earth Obs Geoinf* 2022;106:102651.
- Abuqaddom I, Mahafzah BA, Faris H. Oriented stochastic loss descent algorithm to train very deep multi-layer neural networks without vanishing gradients. *Knowl-Based Syst* 2021;230:107391.
- Nida HS, Bura RO, Abdurahman. Multi-network transmission using socket programming to support command and control systems. *Cyber physical, computer and automation system: a study of new technologies* 2021;p. 59–68.
- Zhao B, Wu W, Xu W. {NetRPC}: Enabling {In-Network} computation in remote procedure calls. In: 20th USENIX symposium on networked systems design and implementation (NSDI 23); 2023. p. 199–217.