Cluster Computing
https://doi.org/10.1007/s10586-022-03582-7

Blockchain for deep learning: review and open challenges

Muhammad Shafay[1] · Raja Wasim Ahmad[1,2] · Khaled Salah[1] · Ibrar Yaqoob[1] · Raja Jayaraman[3] · Mohammed Omar[3]

Received: 13 October 2021 / Revised: 11 January 2022 / Accepted: 3 March 2022

Abstract

Deep learning has gained huge traction in recent years because of its potential to make informed decisions. A large portion of today's deep learning systems are based on centralized servers and fall short in providing operational transparency, traceability, reliability, security, and trusted data provenance features. Also, training deep learning models by utilizing centralized data is vulnerable to the single point of failure problem. In this paper, we explore the importance of integrating blockchain technology with deep learning. We review the existing literature focused on the integration of blockchain with deep learning. We classify and categorize the literature by devising a thematic taxonomy based on seven parameters; namely, blockchain type, deep learning models, deep learning specific consensus protocols, application area, services, data types, and deployment goals. We provide insightful discussions on the state-of-the-art blockchain-based deep learning frameworks by highlighting their strengths and weaknesses. Furthermore, we compare the existing blockchain-based deep learning frameworks based on four parameters such as blockchain type, consensus protocol, deep learning method, and dataset. Finally, we present important research challenges which need to be addressed to develop highly trustworthy deep learning frameworks.

Keywords Deep learning · AI · Machine learning · Federated learning · Blockchain · Ethereum · Smart contracts · Security · Transparency

1 Introduction

The potential of deep learning has been witnessed in almost all industrial sectors. For example, in the healthcare sector, deep learning models are used by physicians to correctly diagnose the disease of the patient from the symptoms. During the recent pandemic caused by the spread of coronavirus disease (COVID-19), deep learning models have been employed to predict the disease spread rate in a particular region and assist the authorities in managing the pandemic using the forecasted results [1–3]. Also, novel deep learning techniques have assisted health physicians in diagnosing COVID-19 patients using the dataset of CT and X-ray images [4, 5]. Apart from deep learning applications in the healthcare industry, it has been employed by security officers at airports to identify and verify banned items in passengers' luggage or safeguarding software from vulnerabilities [6–8]. Using biometric security and face recognition features, deep learning models can assist the authorities in recognizing any physical dangers in real-time. The efficacy and efficiency of a deep learning system basis on the quality of the data used during the model training phase [9]. The majority of the deep learning techniques have considered centralized storage and processing for training the model that is prone to a single point of failure and data alteration by the adversaries. Any alteration of the data used for deep learning operations can corrupt the training model. Blockchain is a decentralized technology that can efficiently handle data integrity, security, and confidentiality [2, 10, 11, 11]. The integration of blockchain with deep learning can bring several benefits, e.g., automated and trusted decision making, efficient data

✉ Ibrar Yaqoob
ibraryaqoob@ieee.org

[1] Department of Electrical Engineering and Computer Science, Khalifa University, Abu Dhabi 127788, UAE

[2] College of Engineering and Information Technology, Ajman University, Ajman, UAE

[3] Department of Industrial & Systems Engineering, Khalifa University, Abu Dhabi 127788, UAE

Published online: 14 March 2022

market management, data security, better model building for prediction purposes, model sharing, and enhancement of the robustness of the deep learning-based systems.

The data collection stage which involves data acquisition, labeling, and improvement, is of utmost important as it can significantly affect the quality and performance of the developed deep learning models. Data collection ensures that large-sized, diverse, and high-quality training data is gathered from multiple sources to enable the training models to perform well on real-world testing data [9]. The model trained on a small amount of data is often overfit and offers limited performance on testing data [12, 13]. The quality of an algorithm is highly affected by the quality of the supplied data; hence models trained using high-quality data offer higher predictive accuracy. Figure 1 enlists important applications and benefits of deep learning techniques along with a description of the role of each application. The support of deep learning in image analysis and recognition, as highlighted in Fig. 1 enables identifying objects, humans, or any action in the image. Such image analysis techniques can be useful for detecting the number of faces in an image or locating the nearby area of an autonomous vehicle using image segmentation techniques. Similarly, voice recognition techniques are mostly used for controlling smartphones or smart homes by recognizing voice-based commands. Text prediction models are often used in web services to predict the text in received emails, summarize messages, or paraphrase essays. Not only is a prediction made, but the deep learning models are also capable of generating human-like text in online web services based on intelligent interpretation of the input data [14]. Apart from these primary benefits of deep learning, it has the potential in terms of efficient data processing for sensory data analytic and optical character recognition (OCR)-based applications. Traditional deep learning systems have leveraged cloud-based servers to store and manage large-sized data to train the algorithms. Cloud computing can speed up the training process of deep learning algorithms by leveraging clusters of GPUs and CPUs for quickly executing compute-intensive tasks [13, 15, 16]. For instance, in telehealth and telemedicine-based services, wearable devices can collect and transfer large-sized healthcare data to the cloud servers through a trusted edge server [2]. In the next phase, such healthcare data is analyzed on resource-rich cloud servers to identify the patterns in the data using deep learning models. However, many of the traditional systems are incapable to fully exploit their potential because of the hindrance created by centralized architecture that is followed by the existing systems. Furthermore, centralized-based data storage and processing raise the risk of a single point of failure problem. Since data has the utmost importance in a deep learning system, hence such data requires high protection against any external or internal attack. Furthermore, the data saved on a centralized system will be less reliable because it is susceptible to changes and fraud.

Blockchain technology has the potential to efficiently handle the aforementioned problems caused by centralized-

Fig. 1 Key benefits of deep learning techniques in various fields

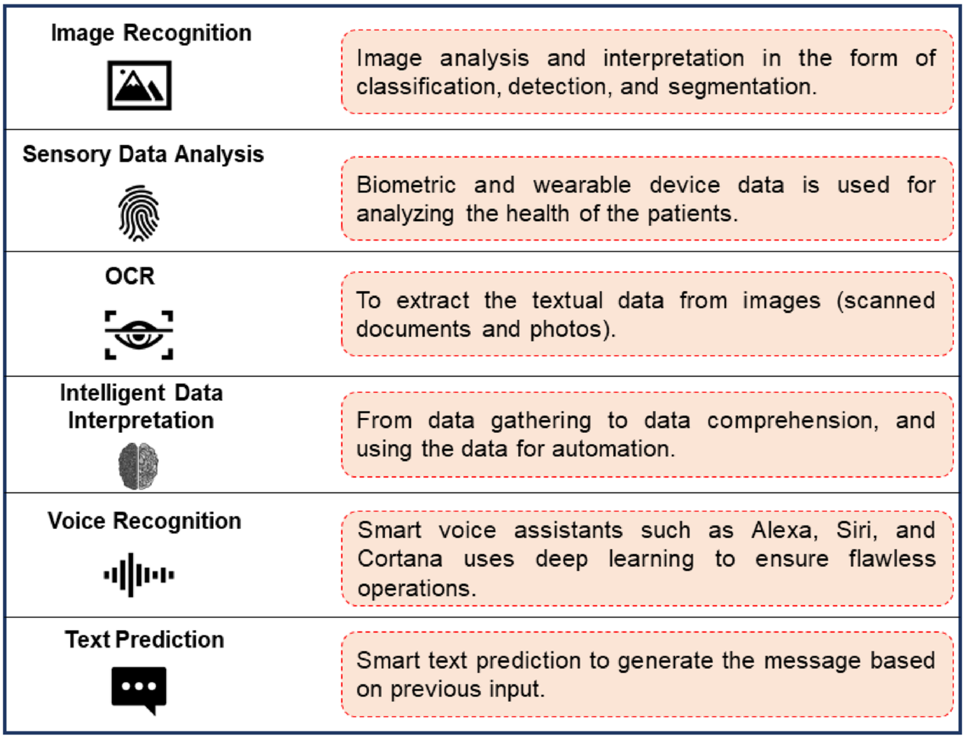


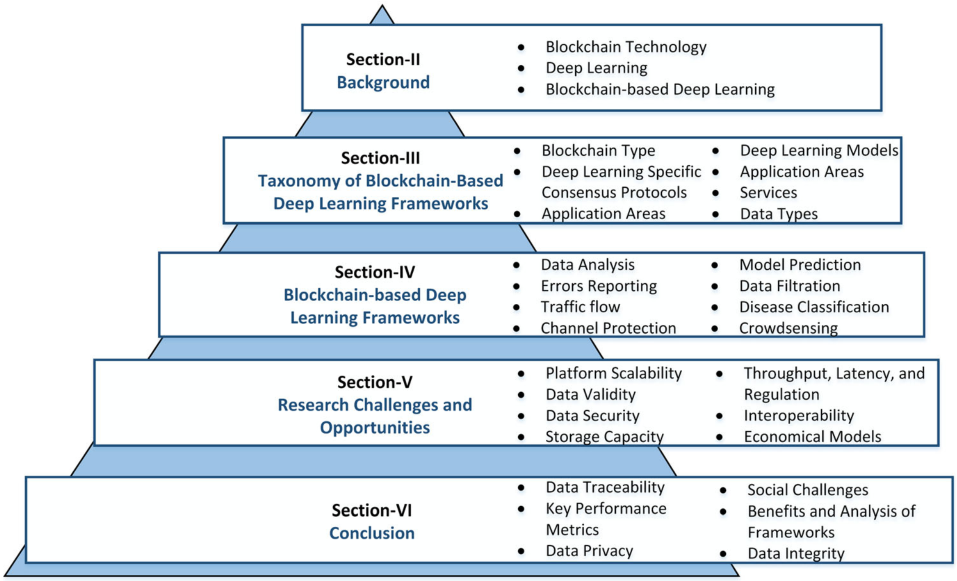
Image Recognition	Image analysis and interpretation in the form of classification, detection, and segmentation.
Sensory Data Analysis	Biometric and wearable device data is used for analyzing the health of the patients.
OCR	To extract the textual data from images (scanned documents and photos).
Intelligent Data Interpretation	From data gathering to data comprehension, and using the data for automation.
Voice Recognition	Smart voice assistants such as Alexa, Siri, and Cortana uses deep learning to ensure flawless operations.
Text Prediction	Smart text prediction to generate the message based on previous input.

based data storage and processing. Blockchain is a revolutionary technology that maintains a shared ledger of data among participants in a decentralized network. It ensures that all ledger copies maintained by the participants are verified and proven consistent [3]. The differentiating features of blockchain assist in enhancing the robustness of the deep learning models by protecting the data against several types of attacks by adversaries. By design, blockchain represents a tamper-proof and tamper-resilient technology that assists in tracking the data to ensure that it has not been tampered with since its creation [3, 17, 18]. The main advantages of novel blockchain technologies include data immutability, transparency, security, provenance, traceability, and operational visibility, and such benefits are envisioned by the decentralized and Peer-to-Peer (P2P) architecture of blockchain. The business processes on the blockchain platform can be automated using self-executing smart contracts that eliminate the role of third parties in executing the services [19]. The smart contract enables the development of a cheap, swift, and reliable system for deep learning applications. The consensus algorithms implemented by the existing blockchain platforms ensure data integrity [18, 20–22]. Existing research works have employed private, public, and consortium blockchains to propose systems for various types of deep learning-based applications. Among these three categories, public blockchain platforms are vulnerable to inference attacks as transactions, pseudonymous addresses, and other user data are publicly available [3]. However, private and consortium platforms preserve data privacy in a much better way than public platforms.

Machine learning models are created, used, and trained by different entities. Blockchain technology enables establishing the provenance of machine learning models, thus leading to trusted Artificial intelligence (AI) systems. Blockchain technology presents a robust system and can incentivize the participants who share their data (data trading) which is used to train machine learning models. Storing machine learning-related data on the blockchain network reduces the chances of errors in the model, as the blockchain network will not have duplicate, missing, or noisy data, which is a fundamental necessity for AI-based models. Combining the values of blockchain and machine learning techniques offers many benefits to the applications, such as data authenticity assurance, data usage management, audit trails, and introducing new values to the business processes through automation enabled by smart contracts. For instance, the machine learning models embedded in smart contracts can automatically recommend a request for authorities to recall expired medicines. Blockchain can capture the evolution of machine learning models as it progresses and records the various stages of the model during its creation, updating, or usage [23, 24]. More specifically, it assists in identifying the owner of the machine learning algorithms, datasets, the source of data, participants, the base model, and processes involved during model creation using the record of immutable transactions stored on the blockchain. The blockchain features such as consensus algorithm, data immutability, and cryptographic hash functions assure that attacks on AI models including data, model, and algorithm poisoning are not possible [24, 25]. Existing studies have thoroughly explored the role and applications of blockchain technology in AI-related domains. For instance, the studies discussed in [26] and [27] have briefly discussed the role of blockchain in AI and machine learning fields, respectively. To the best of our knowledge, there exists no survey/review article that has thoroughly explored the role of blockchain in the deep learning domain to date. To fill this research gap, we have thoroughly reviewed the literature to highlight the strengths and weaknesses in the existing frameworks. It has proposed a taxonomy to classify the literature based on a set of parameters. It has investigated the reviewed literature based on chosen parameters. Finally, open research challenges along with guidelines are presented. The key contributions of this paper are as follows:

- We devise a taxonomy to categorize and classify the existing literature related to blockchain-based deep learning frameworks based on seven important parameters.
- We present insights into the state-of-the-art blockchain-based deep learning frameworks by highlighting their strengths and weaknesses.
- We compare the blockchain-based deep learning frameworks based on important parameters.
- We discuss several research challenges that can affect the performance, accuracy, and prediction quality of existing blockchain-based deep learning frameworks.

Figure 2 presents an organogram to show the organization of the paper. Section 2 discusses the necessary background about blockchain and deep learning models. Section 3 presents and discusses a detailed taxonomy that classifies the existing literature into several categories. Section 4 provides a detailed insightful discussion about existing studies related to blockchain-assisted deep learning frameworks. Section 5 highlights and discusses several open challenges in this field of research. Finally, Sect. 6 concludes the paper.



Section-II
Background

- Blockchain Technology
- Deep Learning
- Blockchain-based Deep Learning

Section-III
Taxonomy of Blockchain-Based Deep Learning Frameworks

- Blockchain Type
- Deep Learning Specific Consensus Protocols
- Application Areas
- Deep Learning Models
- Application Areas
- Services
- Data Types

Section-IV
Blockchain-based Deep Learning Frameworks

- Data Analysis
- Errors Reporting
- Traffic flow
- Channel Protection
- Model Prediction
- Data Filtration
- Disease Classification
- Crowdsensing

Section-V
Research Challenges and Opportunities

- Platform Scalability
- Data Validity
- Data Security
- Storage Capacity
- Throughput, Latency, and Regulation
- Interoperability
- Economical Models

Section-VI
Conclusion

- Data Traceability
- Key Performance Metrics
- Data Privacy
- Social Challenges
- Benefits and Analysis of Frameworks
- Data Integrity

Fig. 2 Organogram of the paper

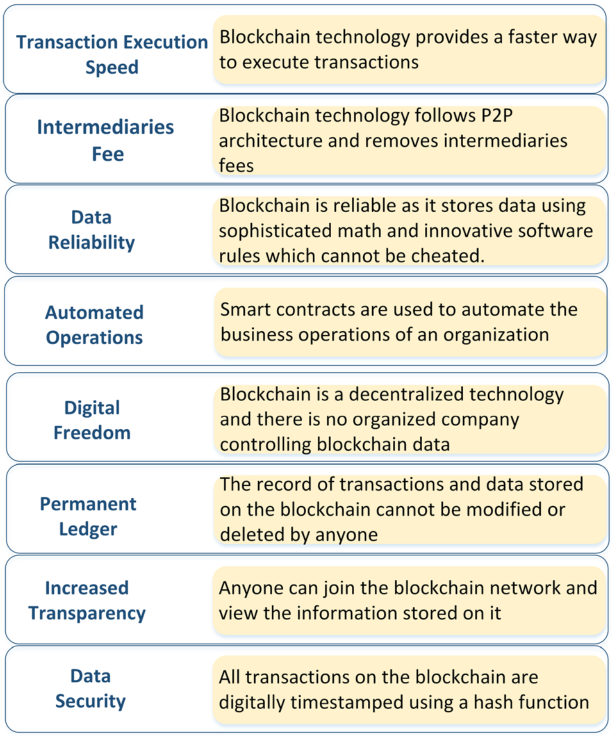
2 Background

This section briefly discusses the key features of blockchain and deep learning and the benefits of integrating them in terms of data security, automatic decision making, and enhanced robustness.

2.1 Blockchain technology

Blockchain stores information in a way that makes it extremely difficult for hackers to change, manipulate, or delete data. By design, it is a decentralized technology that is based on P2P architecture for storing and processing transactions and data. It consists of many nodes which verify and store the transactions in the form of blocks. Each block in the chain stores a set of transactions and it is ensured that the existing blocks are correctly linked to the newly created block to form the chain of blocks. After adding the block to its local chain by a miner, the newly added block is propagated to all participating nodes to ensure data consistency [28, 29]. The decentralized consensus protocol ensures that the blockchain transactions are validated and agreed upon by the miner nodes. For instance, the Proof-of-work and Proof-of-Stake are consensus protocols implemented by many blockchain platforms and they can keep the blockchain secure from any internal or external data hacking attack. Smart contracts, which represent an electronic program, are another important feature of blockchain technology and executes only when predetermined criteria are met [3, 17, 30]. Smart contracts are aimed at reducing the risk and cost of businesses. The key characteristics and features of blockchain technology such as data immutability, smart contracts, consensus algorithms, and decentralization assist in improving business effectiveness [31]. Figure 3 presents an overview of the blockchain benefits that make it different from the other technologies.

Blockchain technology can assist in protecting the Electronic health records (EHR) and Personal health records (PHR) of patients. EHR and PHR systems contain information related to diagnoses, medications, allergies,



Transaction Execution Speed	Blockchain technology provides a faster way to execute transactions
Intermediaries Fee	Blockchain technology follows P2P architecture and removes intermediaries fees
Data Reliability	Blockchain is reliable as it stores data using sophisticated math and innovative software rules which cannot be cheated.
Automated Operations	Smart contracts are used to automate the business operations of an organization
Digital Freedom	Blockchain is a decentralized technology and there is no organized company controlling blockchain data
Permanent Ledger	The record of transactions and data stored on the blockchain cannot be modified or deleted by anyone
Increased Transparency	Anyone can join the blockchain network and view the information stored on it
Data Security	All transactions on the blockchain are digitally timestamped using a hash function

Fig. 3 Advantages of Blockchain technology

medical disease history, test results, etc. EHR contains data about a patient from all clinicians, hospitals, test centers, laboratories, or other healthcare organizations. On the other hand, the PHR is a patient-centric system that is managed by patients in a controlled and private environment [17, 32]. It can ensure that the data is controlled and managed by patients, and it can be shared with other users in compliance with the patient's consent management policy. The consent management policies are implemented through self-executing smart contracts. However, blockchain becomes a costly technology if appropriate techniques are not implemented to manage large-sized healthcare data [17, 32]. To take full advantage of blockchain in many healthcare applications, the pointers and linkers can play a valuable role to minimize the data size. In addition, decentralized storage systems are also capable to securely store large-sized data and avoid single-point-of-failure-related problems. The examples of most widely used decentralized storage systems in healthcare sector include InterPlanetary File System (IPFS) [33], Cassandra [34], SWARM [35], Storj [36], OrbitDB, and Skeps, to name a few [33].

2.2 Deep learning and federated learning

Deep learning encompasses inside AI and machine learning. In existing deep learning approaches, a model learns the latent space representation of the most basic form of data, i.e, images, text, and speech signals. Figure 4 shows the relationship between deep learning, AI, and machine learning. Deep learning allows the hardware to execute many applications with human-like accuracy or even in some cases the better accuracy than humans. To highlight the widespread applications of deep learning in different fields, image classification [37], object detection [38], self-driving cars [39], disease prediction [40], and voice control [41] are well known use cases of deep learning techniques.

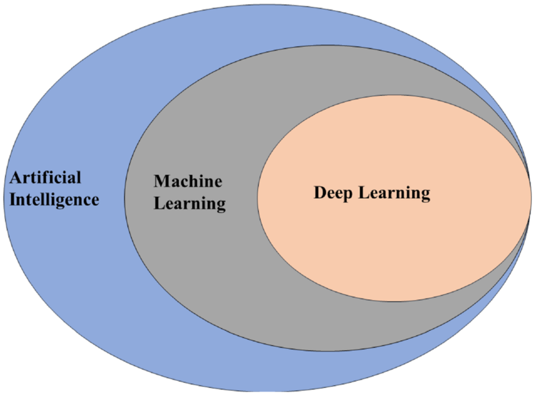


Fig. 4 Relationships between AI, machine learning, and deep learning [42]

The deep learning models are trained by using a large amount of labeled data.

The data representation techniques assist the machine learning algorithms in learning the globalized data sequences. Therefore, data quality plays a critical role, specifically in machine learning algorithms, to make the correct decisions using time series data. More specifically, if the data consists of inadequate descriptions, irrespective of how elaborate the algorithm is, the model will not perform well on such data. Hence, feature engineering is considered as it may assist the reconstruction of data by exploiting the sets of features from the raw data [43]. In the case of a deep learning algorithm, the models are intelligent and automatically extract the high-level latent space features from the basic form of the data. By design, the deep learning models consist of multiple layers. The lower-level layers are responsible for extracting lower-level features, while higher layers extract more abstract features from the input data. The number of layers in DL affects the accuracy and security.

Tuning the deep learning model's hyper-parameters plays a critical role in achieving optimal performance for the given application [44, 45]. In terms of scalability, the important parameter to consider is the choice of network topology on which the model is deployed. For example, a deep learning model, instantiated within the client-server architecture, provides a scalable solution because the model is trained only at the server end, and each node will procure the trained model instance to perform its desired task [46]. Here, if the need arises to re-train (or fine-tune) the DL model, it can be effectively performed at the server end, and this newly tuned instance can be broadcast to each of the nodes within the network (by analyzing the blockchain timestamps) to perform the underlying job. Furthermore, deploying the model on the server end also provides ease in tuning the model's intrinsic hyper-parameters (only once), which can then be generalized across the whole network. On the contrary, deploying the DL model in a peer-to-peer (P2P) network is not a very scalable solution because each node (within the network) will have to explicitly train/re-train its model instance to perform the desired task. This explicit re-training (or fine-tuning) process can be exhaustive (especially for performing a similar or inter-related task across each node)—also, it's highly dependent on the computational capacity of each node.

Deep learning model analyzes the data patterns in the Internet traffic and learns from them to train a model to recognize cybersecurity threats. There are many types of security threats over the Internet that cybersecurity teams are facing nowadays. For instance, malware, data breaches, social engineering, phishing, Denial-of-service (DOS), and insider attacks steal the users' private data for harmful purposes. Deep learning has the potential to combat the aforementioned cybersecurity threats. It assists in detecting the intrusion in the system and preventing it from occurring in the future. Deep learning models analyze the data traffic and verify the transaction signatures to detect the intrusion. It immediately notifies the user about the intrusion detected (through blockchain smart contracts). Furthermore, deep learning models can recognize suspicious activities within the system based on the data traffic analysis to identify bad actors or malware. Natural Language Processing (NLP) which is one of the subtypes of deep learning detects social engineering-related data theft threats [47, 48].

Federated learning methods aim to address the data confidentiality and security challenges in deep learning methods. Data has become the new oil in the recent era owing to ongoing advancements in deep learning. It is forecasted that smartphone users will grow to 4.3 billion in number by the year 2023, and they will generate a huge amount of data [49]. This data can be gathered from powerful sensors such as GPS, cameras, and microphones. On the other hand, deep learning models require a large amount of data to update the models and optimize the user experience at the expense of data leakage and privacy issues. Federated learning (FL) [43] represents a system that does not rely on a central authority to collect and process the users' data. Also, traditional systems require high bandwidth while uploading and downloading data to a central server for model generation. As a result, the central server suffers from a large communication overhead, thus necessitating the use of a decentralized server model to generate models in a resource-efficient way [50]. The unique characteristics of federated learning models have motivated several industries to employ federated learning-based systems for predictive text, voice recognition, cybersecurity threat identification, enhancement of the performance of devices in IoT applications, and disease diagnosis.

2.3 Blockchain-based deep learning

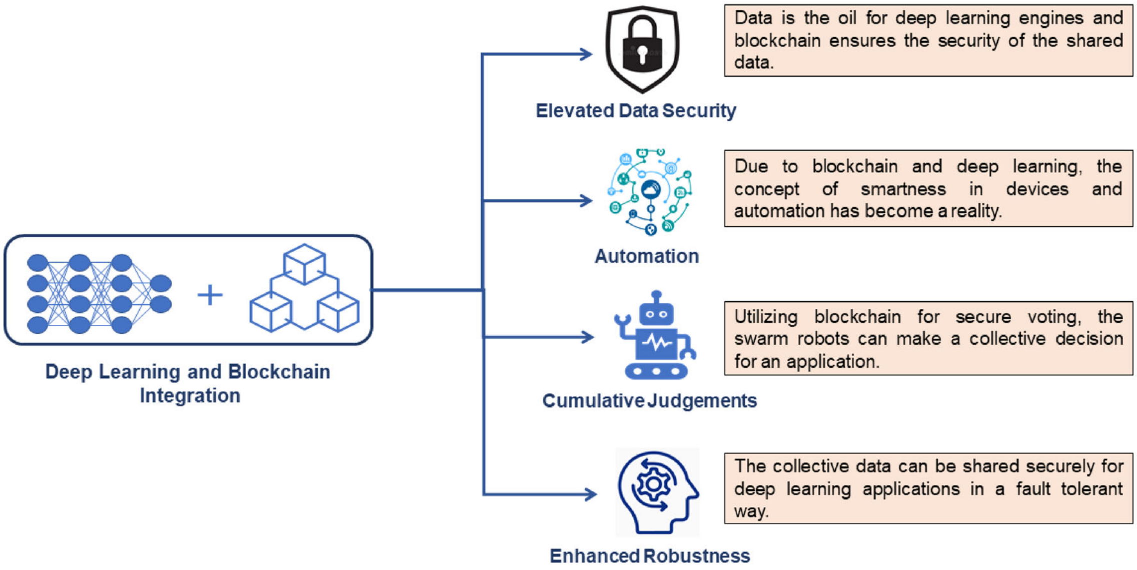
The re-usability and trusted sharing of deep learning models is an essential requirement that can be fulfilled by blockchain technology. Similarly, auditability, data verification, attestation of results, provenance, traceability of ownership, usage, and guarantee of fairness are the main motivations behind the integration of blockchain and deep learning [24]. Deep learning models are fed with a large data of diverse examples and such data is used by the models to learn the features and produce an output with probability vectors in place. Even though deep learning models perform exceptionally well on raw data, but the quality of the data still matters regarding the prediction for many real-world scenarios. The blockchain is a global database in which all network nodes can hold and exchange data in a manner that is decentralized and verifiable.

Table 1 summarizes some of the key features of blockchain and deep learning that assist in improving deep learning-based applications [26]. Figure 5 highlights the four key aspects, scenarios, and categories that take advantage of such integration. By design, blockchain is a fault-tolerant technology that secures the data, while deep learning focuses on utilizing such data to train the models and make accurate predictions. The data immutability feature of blockchain ensures that data leakage cannot happen, thus protecting the deep learning model or data against several types of attacks or data noise issues. As a result, the predictions made by such models are more trustworthy and accurate. The collaboration of blockchain and deep learning opens doors towards automation of several tasks where data requires careful handling and high security. Deep learning and blockchain together can provide a stable, permanent, and decentralized infrastructure for the critical data that deep learning-driven applications would acquire, process, and employ. The following is a summary of the benefits associated with the integration of blockchain technology with deep learning algorithms:

- Data Security: Being a decentralized technology, the information stored on the blockchain is highly secure. Private blockchain platforms are deployed for storing and processing private and confidential information. It is required to keep the private keys of the nodes secret as these keys can be the only way to access the blockchain data. Deep learning algorithms can run on stable data provided by the blockchain, thereby resulting in more trustworthy, accurate, and reliable decision-making [51].
- Automatic Decision Making: Blockchain is a well-known technology that processes transactions on a P2P basis. It makes it easy to verify the decisions made by the deep learning models through the traceability feature. It also ensures that the documents have not been tampered with during the human-assisted auditing phase [24, 52].

Table 1 A summary of the deep learning and blockchain features that assists in improving deep learning-based applications [26]

Blockchain	Deep learning	Potential outcomes
Immutable	Scalable	Flexibility in learning strategies
Transparent	Layered	Collaborative model udate
Integrity	Resource intensive	Enhanced scalability
Cybersecurity	Data intensive	Upgraded data security



Data is the oil for deep learning engines and blockchain ensures the security of the shared data.

Due to blockchain and deep learning, the concept of smartness in devices and automation has become a reality.

Utilizing blockchain for secure voting, the swarm robots can make a collective decision for an application.

The collective data can be shared securely for deep learning applications in a fault tolerant way.

Fig. 5 Benefits resulted from the integration of deep learning and blockchain technology

- Cumulative Judgements: In many cases, the autonomous digital agent makes decisions based on the data gathered related to a particular scenario. Deep reinforcement learning and swarm robotics are examples of such agent-based decision-making system [53–55]. The voting-based approach can assist the robots to make decisions based on the data collected by swarm robotics on the blockchain.
- Enhanced Robustness: In few cases, the accuracy of decisions made by the deep learning models surpasses human-level accuracy. Hence, the highly accurate deep learning model increases the trust of the stakeholders in the decisions. Also, being backed by decentralized technology, the robustness of the deep learning-based system can be ensured. The integration of deep learning with blockchain can be valuable in business forums where the parties can work in a trust-less and automatic environment [56].

3 Taxonomy of blockchain-based deep learning frameworks

Automated decision-making, data security, accurate forecasting, efficient data market management, and enhancement of the robustness of the system are the key benefits that can be achieved through the unionization of blockchain and deep learning techniques. This section presents a thematic taxonomy to classify the existing literature related to the unionization of blockchain and deep learning techniques based on a set of parameters. The identified parameters as shown in Fig. 6 highlight the commonalities and differences among the state-of-the-art blockchain-based deep learning frameworks. Given below is a brief introduction to the selected parameters along with their technological details.

3.1 Blockchain type

This parameter classifies the existing studies into three categories based on the blockchain platforms selected by the existing blockchain-based deep learning frameworks. Many of the services and applications enabled by the deep learning methods have time-related constraints; therefore, the time-specific modalities of the blockchain assists in improving such services. Based on the design, characteristics, and policies, blockchain platforms considered by state-of-the-art deep learning frameworks can be classified into public, private, and Consortium/Federated categories as discussed below.

3.1.1 Public blockchain

The public blockchain platform leveraged by the existing blockchain-assisted deep learning frameworks allows permissionless or unrestricted access to the distributed ledger by the users or machine learning devices. Users access the ledger copy that is distributed among all nodes within the public blockchain network and performs transactions. Public blockchain platforms maintain transaction anonymity because of decentralized data storage and processing. Furthermore, public blockchain platforms are secure against several types of attacks; therefore, they assist the deep learning models in coming up with the correct and trustworthy results [28, 30, 57].

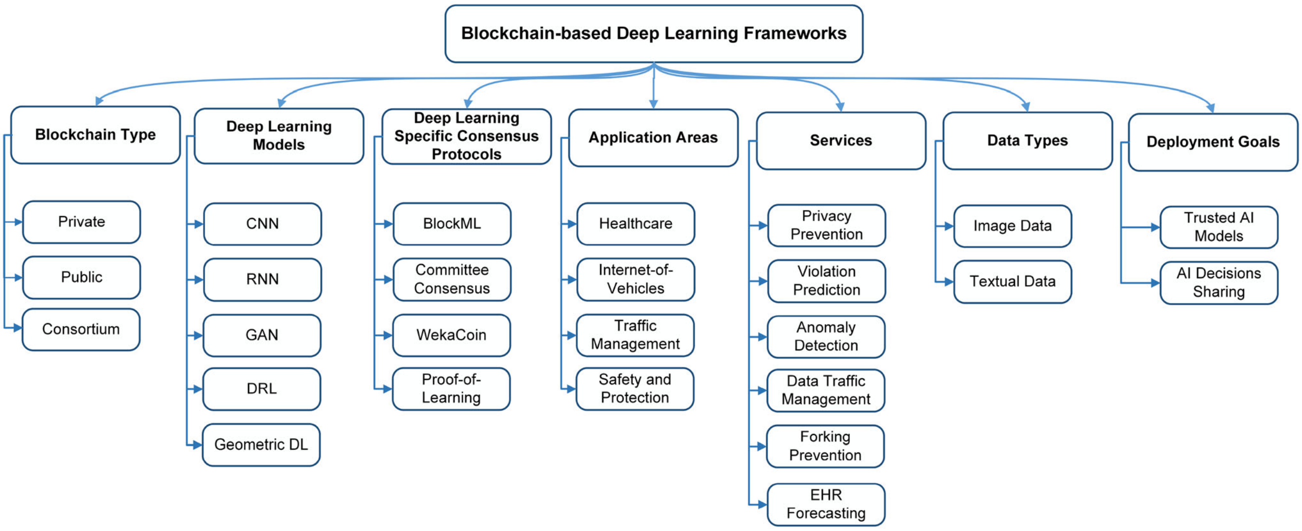


Fig. 6 A taxonomy of blockchain for deep learning frameworks

3.1.2 Private blockchain

The private blockchain platforms leveraged by the blockchain-assisted deep learning frameworks are controlled and managed by a single entity. The private platforms are permissioned where the authority lies within the controlling entity [58]. As the identities of the validators and nodes are known to the central authority, hence the private network requires relatively lesser complex mathematical calculations to verify the transactions. As a result, the private platform's transaction execution speed is higher than the public platform.

3.1.3 Consortium/federated blockchain

Consortium blockchain platforms leveraged by the existing blockchain-assisted deep learning frameworks hold the characteristics of both private and public blockchain platforms. A consortium blockchain functions as a permissioned network and, multiple heterogeneous groups can have the authorization role, unlike private networks where a single authority is responsible for controlling and managing the network [59]. In general, anyone on the blockchain network can access the content on the blockchain, but only a limited authorized group of users can append the data to the ledger. Also, the transaction validation rate of consortium platforms is faster than public blockchain platforms.

3.2 Deep learning models

A deep learning model processes the collected data and makes patterns useful in decision-making in various use cases. Based on the configuration of neural network layers, the deep learning models used for decision-making in several applications areas are categorized into five major categories. Below is a brief overview of deep learning models that have used blockchain-based data to generate patterns and make decisions.

3.2.1 Convolution neural network

Convolution Neural Network (CNN) which is also known as ConvNet processes an image to identify the objects, assign weights to the objects, and classify them according to the context. It also enables detecting object instances in the processed image [60]. The existing blockchain-based deep learning frameworks have employed CNN to classify images, detect objects, and segment the instances in various use cases. The advantage of selecting CNN in blockchain-based studies is the minimum preprocessing time required by the algorithm because of choosing adaptable filters to determine the characteristics of the image.

3.2.2 Recurrent neural network

A CNN model outperforms on images data used as input. However, Recurrent Neural Network (RNN) uses sequential or time-series data to generate patterns [61]. The famous applications of RNN for blockchain-based solutions include voice or speech recognition, speech-to-text conversion, voice search, and natural language processing (NLP). Also, the input data is independent of each other in CNN; whereas, the previous inputs are linked and influence the output in RNN models. Long Short-Term Memory (LSTM) [62] and Gated Recurrent Units (GRU) [63] are upgraded versions of RNN to address the shortcoming of RNN and are widely employed for accurate forecasting.

3.2.3 Generative adversarial networks (GAN)

The generative model learns the patterns in an unsupervised manner and is capable of generating unique data. More specifically, it is a form of generative modeling that employs deep learning techniques such as convolutional neural networks. By design, the GAN model consists of a generator and a discriminator network. The generator is responsible for producing new examples, whereas the discriminator learns to classify the data as real or fake [64].

3.2.4 Deep reinforcement learning (DRL)

DRL is inspired by theories of human behavior that basis on behavioral ecology and enables expert systems to understand the data more precisely. Intelligent agents take actions in an environment consisting of DRL models to learn. Furthermore, agents are implicitly validated or penalized depending on their behavior. Behaviors that lead to the desired result are rewarded, thus termed as reinforced learning-based model [65].

3.2.5 Geometric deep learning

It is a deep learning variant that focuses on developing neural networks that basis on non-euclidean data [66]. A graph is a specific example of non-euclidean data. The data modeling can be done with fewer efforts and resources while using graph-based data. The graphs are input to the geometric deep learning models rather than the data in the traditional form to generic neural networks. In a nutshell, geometric deep learning has the potential to extract more fine granular details from the data.

3.3 Deep learning specific consensus protocols

DL-specific consensus algorithms are aimed at minimizing the convergence time of the consensus process in validating the transactions and assisting the agents in multiagent-based systems to reach an agreement while using minimal resources. The traditional consensus algorithms such as PoW are slow and less energy efficient. The existing blockchain-based studies have proposed various DL-specific algorithms such as BlockML, Committee algorithm, WekaCoin, and Proof-of-learning consensus algorithms. These algorithms are briefly discussed below.

3.3.1 BlockML

The BlockML protocol aims at minimizing the compute-intensive operations in the neural network's training phase using a supervised approach [67]. Unlike the existing systems (e.g., [68]), BlockML forces the miners to solve a task using the same type of input data and generates a competitive model. The suppliers and training data are the vital components in the BlockML architecture. Suppliers are those entities that publish a machine learning task, reward, test data, hash, and training data to IPFS. Once the model training by the miners has been completed, miners immediately publish the results on the blockchain, and the supplier releases the test dataset on IPFS. Based on the model results generated for the test data, the miners fetch the competing solution and the highest-ranked solution is considered the winner. The block of the successful miner will be appended to the blockchain.

3.3.2 Committee consensus mechanism

The consensus algorithm in this category updates the global model using the federated learning-based approach. The Committee consensus mechanism (CCM) updates the model locally and then adds it to the blockchain [69]. The blockchain part of federated learning is divided into two parts. The first part of the blockchain stores the randomly initialized model at the time of system initialization. The second part of the blockchain is responsible to store the model updates generated by the participating nodes in the federated learning-based systems. These nodes essentially obtain the model present in the blockchain and undergo the training locally. Afterward, the local gradient is verified and added to the new block; hence, the blocks continue to grow in number.

3.3.3 WekaCoin

WekaCoin is a cryptocurrency that provides a decentralized database of machine learning models which are publicly accessible. The block of WekaCoin holds three items such as transactions, the hash of the previous block, and the information about the machine learning contest used to verify the block in the previous iteration. By registering a model transaction on the chain, the newly created model can be uploaded to the WekaCoin network [70]. The deep learning model trainers only store the hash value of their models at this point. Deep learning model trainers publish their models to IPFS once the test data is available. The advantage of this approach is the minimum energy requirement as only relevant tasks are solved by miners, the development of a library of machine learning algorithms, and datasets that are publicly available.

3.3.4 Proof-of-learning

The proof-of-learning algorithm practices a nonconvex optimization approach of potentially large-sized neural networks and exhibits NP-hardness characteristics because of computational complexity in solving the asymmetry challenge. In this system, a Secure mapping layer (SML) is introduced for tamper prevention [71]. In the Proof-of-learning algorithm, the consensus nodes provide the processing power and strive to train the model for the issued task. Once a model has been created that fulfills the minimal training exactness, the successful miner broadcasts the new block and declares its success.

3.4 Application areas

This parameter specifies the key application areas that are focussed on by the state-of-the-art blockchain-assisted machine learning frameworks to ensure data integrity. The main categories include healthcare, IoVs, traffic management, and safety and protection.

3.4.1 Healthcare

Healthcare is another field in which patient data is of extreme importance. The healthcare data that comprises clinicians reviewing images and scans is used for research purposes for training deep learning models for the prediction of communicable and non-communicable diseases such as COVID-19 and cancer, identifying new ways of diagnosing illness, improving the quality of healthcare services, and improving clinical trials [2]. The healthcare data consists of records related to the diseases, disease symptoms, and medical profile of the patient that assist the deep learning models in predicting the medical condition of a future patient. The high security and precision of such healthcare data is the utmost requirement of the healthcare industry as it directly affects human life. Storing such data on centralized systems makes them vulnerable to attacks. Hence, blockchain technology, in this case, can play a vital role in preserving all the data as it is inherently decentralized and provides a solution to data security threats. Blockchain assures that the data is safe from any planned or accidental loss.

3.4.2 Internet of vehicles

The rise in the popularity of the Internet of Things (IoT) as an instrumentation technology for vehicles has rapidly increased the networking capabilities of existing devices [72]. The vehicles in modern transportation systems are usually equipped with devices, sensors, and intelligent software to get connected and exchange data with each other. By introducing blockchain on the Internet of Vehicles (IoV) network, the data can be securely shared among the entities as highlighted in Fig. 7 [73]. The V2V communication mode of intelligent transportation systems enables a vehicle to share data about a vehicle's position, speed, congested route, heading angle, and road state with other vehicles. The vehicles make different decisions, such as optimal route finding based on such data. The V2I model enables vehicles to share the captured data with roadside units, such as RFID readers and cameras, lane markers, and streetlights, to improve public safety, efficiency, and convenience [74].

3.4.3 Traffic management

Road congestion is one of the biggest issues in urban areas and can result in large financial and ecological losses due to increased carbon dioxide emissions [75, 76]. One of the most widely used approaches for forecasting traffic intensity using real-time data is crowdsourcing. However, current crowdsourcing models are impractical due to centralized-based data storage and human safety concerns.

3.4.4 Safety and protection

Blockchain technology is well-known and practicable in various fields due to the security feature it provides in the terms of the immutability of the data. The differentiating features of blockchain technology and decentralized data storage characteristics can assist in securing various application areas. However, the 51% attack on the blockchain that can affect data integrity is still possible [77]. Data can be protected from such attacks using machine learning approaches. Signatures represent the malicious code's fingerprint and aid in detecting and identifying malware.

3.5 Services

This parameter highlights the primary service of blockchain-based deep learning solutions targeting various application areas such as healthcare, vehicular communication, and IoT. Existing approaches aim for many objectives including privacy preservation , traffic violation prediction, anomaly detection, cellular traffic management, and forking prevention.

3.5.1 Privacy preservation

Privacy protection assures that the data about an entity cannot be disrupted or viewed by unauthorized users. Blockchain technology has the potential to preserve the privacy of the data used in deep learning models [78]. Privacy protection can be ensured through proxy re-encryption and highly profound data encryption schemes.

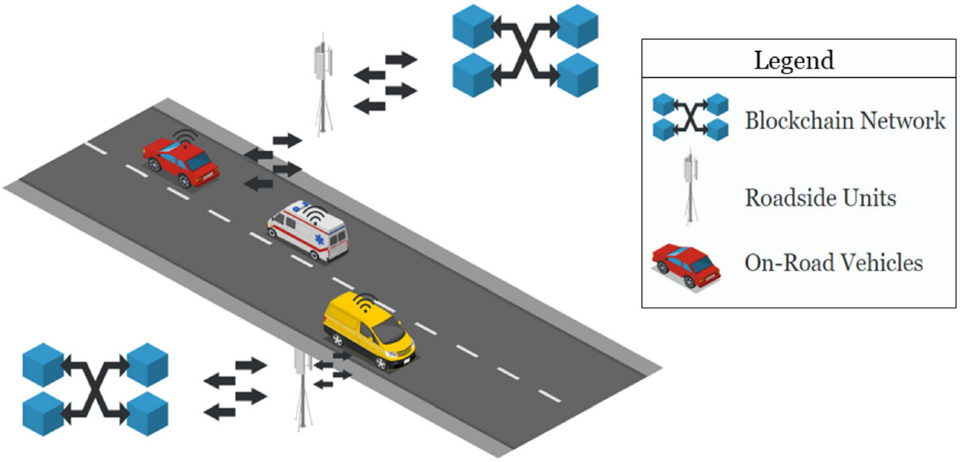


Fig. 7 Blockchain-based system for Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communication

3.5.2 Violation prediction

Traffic violations can be of many types including overspeeding, driving while intoxicated, illegal lane changes, or failure to stop at a red light. The blockchain and deep learning amalgam play a key role in predicting the violations made by the drivers. For instance, the deep learning-based classifier can use highly secure blockchain data to predict traffic incidents [79]. Such data can be immutably recorded on the blockchain and can be used by the roadways department for road network planning, insurance companies for damage assessment, and law enforcement agencies for lawmaking.

3.5.3 Anomaly detection

Anomaly detection deals with the identification of data that deviates from normal behavior. Based on blockchain-based data and transactions, the encoder-decoder network in deep learning can learn the normal behavior of a real-world phenomenon [80]. It can immediately detect an anomaly based on the analysis of data. Furthermore, such anomalies can be distinguished by the model as well.

3.5.4 Data traffic management

Predicting data traffic using deep learning methods in cellular networks can improve the reliability and efficiency of the network. The current 4G LTE systems can be configured to the blockchain-based cellular network by customizing the access points as blockchain nodes to manage the cellular traffic. Every newly created block can be transmitted to all nodes of the blockchain called Blockchain-radio access network (B-RAN) [81].

3.5.5 Forking prevention

Forking refers to the blockchain state where the ledger is divided into two potential paths or a situation where more than two blocks have the same block height. Blockchain and federated learning models can prevent forking by enabling the miners to verify the blocks correctly. For instance, by broadcasting an acknowledgment signal in blockchain-assisted deep learning models, the occurrence of forking can be prevented successfully [82]. On receiving the acknowledgment signal from all the miners, the generated block is considered valid.

3.5.6 EHR forecasting

EHR refers to the digital version of the patient-related data such as lab results, medical history, allergies, and prescribed medicines. Usually, patients use an open framework to regulate, manage, and share their medical data with family, friends, and physicians. Data integrity, data protection, and secrecy are all major concerns in such a setting. EHR forecasting using deep learning techniques assist in predicting health situations in a particular region and facilitating health care service in that region [83].

3.6 Data types

The attributes of this parameter define the type of data that can be accepted and processed by the deep learning models backed by innovative blockchain technology. The deep learning models supported by blockchain, in general, take data in the form of images and text.

3.6.1 Image data

CNN accepts the image as an input and extracts deep and shallow features to classify the data. Storing and protecting the images used by the CNN models for training purposes by employing blockchain can increase the trust of users on data [84]. To optimally utilize the capacity of the ledger, blockchain can store an immutable hash of the data stored on a third-party storage system such as IPFS. In this way, the data can be tracked via the immutable hash of the images stored on the blockchain.

3.6.2 Textual data

Textual data requires less storage capacity compared to image data. However, such data need to be optimally stored on the blockchain to ensure platform scalability. Most deep learning-based systems have considered textual data for traffic or disease prediction.

3.7 Deployment goal

Operational auditability, data verification, attestation of results, provenance, traceability of ownership, usage, and guarantee of fairness can be achieved by integrating blockchain and deep learning [24]. The deployment goal parameter highlights the main services that can be achieved by converging blockchain and deep learning technology. The main goals include model provenance and decision-making.

3.7.1 Trusted AI models

Blockchain technology assists in establishing the data provenance of the deep learning models used by various industries for data classification and prediction. The predicted results generated by the trusted AI models are highly reliable and correct. The model provenance enabled by blockchain technology can assist in verifying the trustworthiness of a deep learning model.

3.7.2 AI decisions sharing

Blockchain technology assists to securely share the data between the untrusted participants. The AI decision-sharing parameter describes the predictions made by the AI models based on time series data and secure sharing of the prediction among the potentially untrusted participants. For instance, in IoV, the AI model can identify a road accident and share this information with all the other vehicles registered to the network to avoid traffic congestion.

4 Blockchain-based deep learning frameworks

This section presents a detailed review of state-of-the-art blockchain-based deep learning solutions for various application areas. It compares existing blockchain-based deep learning framework schemes based on parameters selected from the literature.

4.1 Review of blockchain-based deep learning frameworks

This subsection presents a review of existing blockchain-based deep learning frameworks mainly targeting healthcare, vehicular networks, cellular traffic management, and blockchain safety and protection from adversarial attacks.

Data analysis The patient-centric data management and resource-friendly design of deep learning models has become an essential element particularly in pharmacogenomics research. The work discussed in [85] has proposed a decentralized system that provides pharmacogenomics data to deep learning models to predict ovarian cancer. The blockchain platform is used for sharing healthcare data, patient records, and ovarian cancer predictions made by the model with the participating organizations. The authors of the proposed research have validated their methods by comparing them with the CRYPTO++ standards. These standards are helpful to evaluate the schemes based on the amount of time it takes to encrypt and decrypt data. However, the proposed model has been tested using a small number of examples per class and makes predictions for unseen data belonging to the learned classes.

Provenance data for AI models An AI model can lead to erroneous or misleading results particularly when the source of the employed deep learning model is unverifiable or unknown. For instance, a deep learning model can be fed with poisoned data to compromise the malware detection functionality of the malware detection model. By tracking and tracing the data related to the history of an AI model ranging from its creation, the data used in training, the owner of the model, and the processes involved during model creation can assist in protecting the model from several data positioning attacks. Additionally, the availability of verifiable training data significantly increases the trust of the users in the AI models. The study presented in [24] has proposed a blockchain-based architecture for maintaining provenance data for AI models. The main classes implemented by the proposed system to generate highly trusted AI models include participants, datasets, models, operations, and compute pipelines or projects. The record of immutable transactions stored on the blockchain has assisted the users in successfully maintaining the data provenance of the generated AI models.

Model prediction In the healthcare field, cross-institutional health data sharing for research purposes is indispensable. It is invaluable to assure full compliance with the health data sharing rules defined by the regularities for data privacy and security preservation. Ensuring the security of the data is one of the main challenges faced by the healthcare industry. The study discussed in [68] has highlighted the importance of sharing deep learning models between the peers of the blockchain network instead of sharing data to preserve data privacy and confidentiality. The study has proposed a novel consensus algorithm called proof of information [68], which is well suited for healthcare model prediction. The proposed study has employed an incremental federated learning model for implementation purposes. The incremental learning approach trains the model using a single dataset in the first round. The second round considers a new dataset that belongs to a similar class but different unseen examples. The errors are calculated for each model trained by each member and the model with the least error is broadcasted to all participants. Finally, through blockchain, all participants receive this model as it exhibits least errors as the latest model.

Data filtration Deep learning methods are highly capable of accurately predicting the causes or sources of disease and ensures that high-quality data is used during the training process. In the research work proposed by Zheng et al. [86], a deep learning technique supported by immutable blockchain technology is employed to classify healthcare data into high-quality and low-quality classes and aligning such data with certified consumers. It was noticed that the majority of healthcare data comes from wearable devices at a continuous dynamic rate; such datasets can be highly valuable to build disease prediction models. The proposed study has mainly focused on dynamic data as this type of data is easily accessible via wearable devices. Sophisticated machine learning techniques are used to analyze the trend in the accumulated data to ensure the validity of data with high-performance validation patterns. The proposed study has assured that data is collected from a reliable device to build deep learning models.

Disease classification Deep learning models are famous for their feature extraction abilities and can play an important role in future medical research. Deep learning can analyze medical images like X-rays, MRI scans, CT scans, etc., and determine any health risks and flag anomalies in the CT scan images. Deep learning approaches are being utilized to predict various diseases such as pneumonia [87], multi-retinal disease, class imbalance in coronary heart disease, and cardiovascular disease [88–90]. Juneja et al. [91] have employed blockchain platform to train the deep learning models to classify arrhythmia disease. Arrhythmia is a disease related to irregularities in the heartbeat of the patient. In a study proposed in [91], two-layered stacked denoising auto-encoders (SDA) are used to extract the features and to detect anomalous heartbeats during the first phase. Furthermore, the SDA is retrained in the testing phase as well to control the false-positive rate in the output classification. The data gathered after retraining can be stored in the hardware and can be accessed by using the pointers that are stored in a time-stamped ledger.

Combined cooperative positioning A Global positioning system (GPS) is the most widely used system that assists in tracking and tracing vehicles. The issues in GPS-based vehicle position estimation include errors in range estimation, wrong installation, a low battery of GPS enabled devices, and impossible vehicle lane-level navigation. A cooperative positioning approach that aims at optimizing the GPS-based position estimation is useful to improve the lane-level vehicle positioning accuracy [92]. The cooperative positioning approach suggested in [92] involves a blockchain-based system for combining cooperative positioning (CP) with the IoV to increase GPS data validity and consistency. Through decentralized blockchain technology, data can be efficiently and securely shared among the vehicles that are part of the vehicular network. The cooperative vehicles in a vehicular network can share the trained deep learning models to estimate the error in vehicle estimated position. The vehicles can evaluate the relatively accurate location based on the traffic sign and share the updated model with other vehicles. The proposed study overlooked the cost and security analysis to highlight the feasibility of their proposed research.

Crowdsensing In a traditional vehicle-to-vehicle (V2V) communication network, the data privacy of the participating vehicles can be affected by the centralized-based data storage and processing. To ensure that the vehicle exchanges data in a trustworthy way, blockchain proves to be a viable solution. The work presented in [93] has proposed a blockchain-based system to handle the network latency in centralized systems. The proposed research has also solved data protection problems in the vehicular crowdsensing frameworks by employing 5G-enabled IoV [93]. A deep reinforcement learning-based algorithm has been used to deal with block mining in blockchain networks. The system consists of a registration authority that authenticates and authorizes the roadside units, which acts as the miner nodes within the blockchain network. The vehicles that are involved in data exchange are also registered to the blockchain networks. The participating entities collect the data and share it with the group head who is responsible for refining the data and transferring it to the server via base stations.

Errors reporting As discussed in [92], GPS-based techniques do not provide high lane-level positioning accuracy for vehicles. To handle this issue, Li et al. [94] proposed an edge and deep learning-based error sharing framework to improve the accuracy of position estimation of vehicles. The proposed study has also secured the data published by participants by employing a blockchain platform for vehicle position error reporting and sharing. In the proposed framework, the evolution of positioning error is evaluated using the deep learning-based prediction model. The role of blockchain was limited to a trusted bridge between the existing vehicles and edge computing nodes. Furthermore, the deep neural network-based error correction protocol was deployed on the edge server to exploit the low latency and higher computing capacity of edge servers while training the deep learning model. To reduce the block confirmation time, the proposed study has employed a delegated proof-of-stake (DPoS) consensus algorithm instead of PoW which is resource unfriendly by design.

Crowdsourcing The researchers in [95] have employed a crowdsourcing technique that uses a blockchain network to identify and report an unwanted situation that occurred on the road such as road jamming due to vehicle accidents. The proposed system has employed a smart contract that enables vehicles to register and leave the network freely and anytime. The proposed system has chosen a proof-of-authority consensus algorithm for data validation while considering the resource constraint nature of participating entities/devices. On encountering an unwanted incident on the road, the user can employ the DApp interface to report the event to the blockchain. Furthermore, to assure that the same participants do not transfer the same data to the blockchain for economic rewards, only the first participant who shares the information about the jammed route can receive tokens as a reward. It encourages users to exchange information quickly to claim the rewards. The proposed study has used LSTM to predict the likelihood of traffic jams at a specific moment, location, and point.

Traffic flow prediction: Federated learning is a decentralized machine learning approach that trains an algorithm on various devices that have local data samples. A Federated learning network supported by blockchain technology ensures the preservation of the privacy of the client data. Blockchain has a vital role to play in such systems to ensure that the malicious participant can not share the false update; thus, the efficacy of the global model cannot be compromised. Also, the federated learning framework supported by blockchain enhances the accuracy in traffic flow prediction and improves traffic management with enhanced user security and privacy. The researchers utilized this technology along with Gated Recurrent Unit (GRU) neural network to process live traffic data for improved traffic management. A consortium blockchain is used to keep the distributed ledger running by employing a small number of preselected miners [96].

Computations offloading Mobile edge-cloud computation offloading (MECCO) [101] presents a viable solution to overcome the challenges related to resource limitation of mobile devices by identifying and offloading resource-intensive tasks of an application to the mobile edge-servers. Although MECCO has significantly augmented mobile battery lifetime, preserving the privacy and security of the traffic data is overlooked [102]. As a result, the MEECO can be fully exposed to a variety of risks and attacks. Nguyen et al. [99] has proposed a blockchain and deep reinforcement learning-based solution to handle the security and privacy challenges in the MECCO system. The proposed system has employed blockchain-based smart contracts to increase trust among participants of the edge-cloud computing network.

Channel protection Singh et al. [98] has proposed a blockchain and deep learning-based approach that has implemented Zero Knowledge Proof (ZKP) for validating the registered machines. The proposed method registers the machines such as a drone and verifies them using ZKP before the transactions from such machines can be granted. The proposed study selects a miner node using a novel selection algorithm that involves a deep Boltzmann machine. The blockchain assures that data integrity will not be compromised during device-to-anything (D2X) communication. Also, ZKP-based validation assists in limiting the chances of network hacking by malicious drones. In the proposed study, at the time of machine registration, the role of each drone like the ordinary drone or miner drone is explicitly mentioned. In a nutshell, the proposed research has produced a model for safe device-to-device (D2D) and D2X communications.

Utility-based blockchain security The deep learning model can assist in safeguarding the blockchain from several types of attacks by adversaries. In decentralized technology, a double-spending problem that represents a flaw in existing cryptocurrency systems can occur, and it can enable the user to spend the same token two times. According to Rosenfeld's [103] model, the number of verification required to keep the attacker's effectiveness in double-spending below 10%, 1%, and 0.1% is 2, 4, and 6, respectively. In [100], a utility function-based approach has been proposed, in which utility function is calculated based on the value of the product or service being traded against the cryptocurrency [100]. The evaluation of such a utility function depends on the utilization of smart systems to infer the value of the product or service. Relying on the cost of the commodity/service, this utility feature will influence the attacker's decision regarding attacking the blockchain. This utility function once fed into machine learning-based classification algorithms can assist in determining whether or not an intrusion is likely to occur.

As discussed in this section, it can be comprehended that the majority of the deep learning systems utilize blockchains for secure storage of data and easy accessibility. Contrarily, some of the methods discussed the employment of deep learning models to protect the blockchain from foreign attacks.

Figure 8 discusses a federated learning-based system [24, 25] that securely shares the models constructed by the devices based on the local data. Federated learning is a decentralized type of machine learning. In a traditional machine learning approach, the data is captured at a centralized device and used to learn and train the models. However, storing data on centralized devices affects the users' data privacy. Federated learning is a decentralized approach that allows users to train their models locally using private data, thus preserving data privacy. After training the model locally, the training results (not data) are shared by each participating device with the centralized coordinator to update the global model. The centralized coordinator will now update its central machine learning model based on the training results shared by each device. This central machine learning model is represented as a global updated model in figure 8. This global model can be accessed by the participating devices. The role of blockchain in this scenario is to coordinate between participants to share the local and global models. Figure 9 shows a high-level design of a blockchain, deep learning, and IPFS-based system that can automate healthcare, telecommunication, and transportation management services.

4.2 Comparison of existing frameworks

This subsection compares existing blockchain-based deep learning frameworks based on several parameters. The main parameters that are considered for the analysis include solution category, blockchain type, consensus approach, deep learning methods, data set used by the deep learning model, study strengths, and limitations. Table 2 compares the research works discussed above in different categories where deep learning and blockchain are playing their part.

The majority of the literature related to blockchain-assisted machine learning frameworks have considered healthcare, the IoV, cellular traffic management, and blockchain safety and protection fields. The frameworks discussed in [68, 85], and [86] focused on creating machine learning-based models for disease prediction and data filtration in the healthcare industry. However, the frameworks discussed in [92, 93], and [94] have focused on employing machine learning-based models and blockchain technology for position estimation and error reporting for vehicles that are part of the established IoV. The works discussed in [95] and [96] have focused on employing

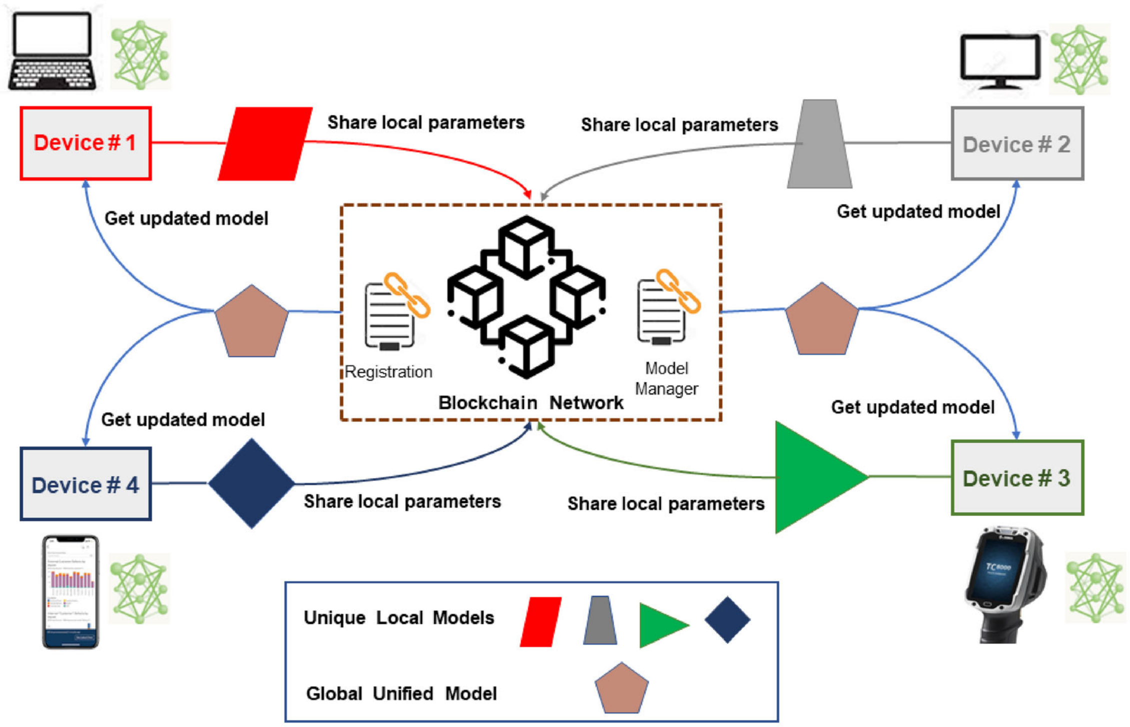


Fig. 8 Blockchain-based federated learning

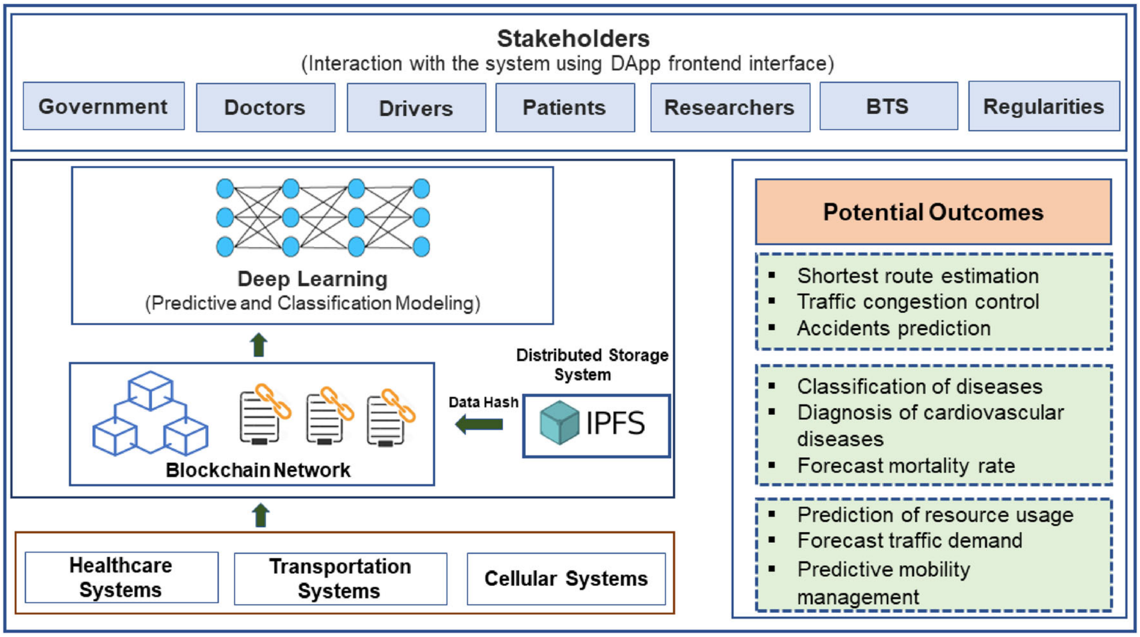

Stakeholders
(Interaction with the system using DApp frontend interface)
Government
Doctors
Drivers
Patients
Researchers
BTS
Regularities
Deep Learning
(Predictive and Classification Modeling)
Distributed Storage System
Data Hash
IPFS
Blockchain Network
Healthcare Systems
Transportation Systems
Cellular Systems
Potential Outcomes
Shortest route estimation
Traffic congestion control
Accidents prediction
Classification of diseases
Diagnosis of cardiovascular diseases
Forecast mortality rate
Prediction of resource usage
Forecast traffic demand
Predictive mobility management

Fig. 9 Highlighting the system components and participants and describing the outcomes resulting from the integration of deep learning with blockchain technology in various fields

machine learning-based models and blockchain technology for cellular traffic management.

Among [68, 85], and [86] frameworks, the study discussed in [86] has considered LSTM deep learning model; whereas [85] and [68] has considered one-shot learning with Siamese neural network and incremental learning, respectively. Moreover, the frameworks discussed in [85] and [86] have selected the human protein atlas and EHR based data set for model training purposes, respectively. Among [92, 93], and [94] studies, the work discussed in [92] has considered Deep Boltzmann Machine deep learning models. On the other hand, [92] has considered DNN and Reinforcement learning deep learning model. The MIT-BIH database dataset is considered for model training in [92]. The work discussed in [95] and [96] has considered custom datasets for model training purposes. Also, [95] and [96] has followed CNN and DNN deep learning models for data classification.

The existing frameworks have implemented various services related to privacy prevention, violation prediction, anomaly detection, forking prevention, and EHR forecasting in healthcare, cellular data traffic, and vehicular communication networks [68, 85, 86, 92, 94, 95]. In a vehicular communication network, blockchain can be useful to securely share the decision made by an AI model such as road accident detection or traffic jamming with the authorized participating entities. Such data is useful to efficiently manage road traffic by evenly distributing the traffic on the roads [104]. In comparison to the works presented in [68, 85, 86, 92, 94], and [95], the work presented in [24] has discussed how blockchain platforms can be fully adopted to verify the genuineness of an AI model based on the model provenance data. The data about the AI models stored on the blockchain in [24] includes training datasets, training results, model owner credentials, the source of data, and participants.

5 Research challenges and opportunities

This section briefly discusses several research challenges and opportunities regarding the integration with deep learning and blockchain technology. The major issues that hinder the successful integration of these technologies include platform scalability, assurance of secure data exchange, transaction execution latency, platform interoperability support, the large volume of data collected from sensors and wearable devices, secure economical models, and computationally expensive consensus protocols. In the following section, the aforementioned challenges and opportunities are discussed.

5.1 Platform scalability

A major issue faced by architects designing deep learning applications is the availability of multiple variants and configurations of blockchain technology. A scalable blockchain platform can efficiently handle the large volume and velocity of transactions generated by diverse users. A sizeable blockchain network would necessitate a

Table 2 A comparison and analysis of state-of-the-art blockchain-based deep learning frameworks

Category	Blockchain type	Consensus protocol	Deep learning method	Dataset used	Study strengths	Study limitations
Ovarian cancer prediction [85]	N/A	N/A	One-shot Learning	Human Protein Atlas	Requires less training time as only one example per class is required	Performance is not as good as DNN
Data exchange [68]	Private	Proof-of-Information	Incremental Learning	N/A	The considered model can learn new classes on a pre-trained network	Possibility of happening of catastrophic forgetting
EHR prediction [97]	N/A	N/A	LSTM	EHR-based dataset	Superior performance on sequential data	Longer training time and excessive memory requirement
Arrhythmia classification [91]	N/A	N/A	SDA+ Sigmoid	MIT-BIH Dataset	Deals well with noise and random variations in the data	Requires large amount of data for better results
Miner Node Selection [98]	N/A	Zero-Knowledge Proof	Deep Boltzmann Machine	N/A	Data labelling is not required	Expensive in terms of memory and CPU cycles
Communication security [99]	Private	N/A	DNN+ Reinforcement Learning	N/A	Implemented method can solve complex problems that conventional methods cannot	Not suitable for simple problems
Securing blockchain [100]	N/A	Proof-of-Work	N/A	Game Theory-based Utility Function	Data is fully secured	This approach is not practical
Traffic jam prediction [95]	Public	Proof-of-Authority	ANN+LSTM	Historic Traffic Data + Custom Dataset	ANN are fault tolerant as information is distributed over all the nodes	Black box behavior makes it impossible to develop a relation between dependent and indepen- dent variables
Traffic flow prediction [96]	Consortium	Delegated PBFT	GRU	N/A	Lesser expensive in terms of memory requirement as compared to LSTM	Lesser learning ability compared to LSTM
Incident prediction [81]	N/A	N/A	CNN	Custom Dataset	Excellent feature extraction cap- ability and computationally efficient solution	Large dataset is required for training and noise in data can cause misclassification
GPS correction [94]	Public	Delegated PoS	DNN	Custom Dataset	Reduced need of feature engineering	Not prone to data redundancy and inconsistency

comparable amount of accounts to implement deep learning-based services targeting healthcare [68], road traffic jam [95], and traffic management in cellular networks. The deployment of blockchain on such a massive scale will lead to several problems mainly related to demand of users for internet connectivity, data velocity, speed, and volume of transactions generated by participants. Considering the storage and computational requirements of the ever-growing blockchain ledger, the number of blocks and transactions to be added to the blockchain must be reduced considerably to satisfy the anticipated requests by the users. To carefully handle the scalability challenges of existing blockchain platforms, the compression algorithms having lightweight design, high compression ratio, and resource inexpensive nature should be integrated into the existing blockchain-assisted deep learning solutions [105]. However, many of the current compression algorithms are unable to fully provide the appropriate ratio necessary to bring down the economic cost of the large-scale deployment to provide deep learning-based services using blockchain.

5.2 Data validity and secure sharing

Data sharing between the participants of the healthcare industry such as doctors, patients, and nurses can be kept anonymous owing to state-of-the-art encryption and decryption techniques implemented by private blockchain platforms. With the exponential rise of IoT and wearable devices in healthcare and vehicular networks, consumers are greatly concerned about data privacy, security, and confidentiality. A multi-layer blockchain architecture that supports data fusion and allows advanced analytical authentication for user groups can assist in securely sharing data between the participants. Moreover, during the recent pandemic caused by COVID-19, the verification and secure sharing of immunity passports [2, 106, 107] between the authorized users through a blockchain platform has become a must to meet requirements in the healthcare industry for developing a viable, secure, and fault-tolerant solution. More specifically, the potential of blockchain in validating and archiving immutable data in real-time opens up the possibility to ensure the authenticity of the data. The blockchain platform serves as the basis for many research projects. Particularly, it allows companies and organizations, along with technologists and experts in sharing knowledge and validating the data in new systems in a trusted and reliable way. However, the IoT devices used to collect vehicular data can be faulty or inappropriately deployed. As a result, the data generated by such devices is often incorrect, misleading, and unreliable. Metadata binding [108] is a technique that requires the IoT devices to calculate and send their accuracy and precision level along with collected data to the blockchain. Such metadata can be analyzed by the deployed smart contracts to verify the validity of the data before storing it on the blockchain. Also, in a few cases, a reputation-aware system can be advantageous in calculating the trust level of participating users based on their behaviors. Smart contracts can be deployed to ensure that data from only highly reputed entities is recorded on the blockchain.

5.3 Structural enhancement and storage capacity

Deep learning is regarded as a resource expensive and data-intensive approach that assists in solving various real-world problems. Blockchain technology aims to ensure that the data stored on it is highly secure, immutable, verifiable, transparent, and visible to the authorized stakeholders. The deep learning-based approach which can systematically evaluate the performance of blockchain platforms is economically desirable. Such an approach can provide a means to analyze the current structure of the blockchain for possible structural enhancement. Deep learning algorithms not only solve the real-world object detection and classification problems in various domains, but these models can also provide a way to compress the data with a higher compression ratio [105, 109, 110]. However, blockchain represents an ever-growing ledger as new blocks keep on adding to the existing network. As the size of the blockchain increases, the efficiency of the network is profoundly affected. Deep learning approaches can be employed for compressing the data as well as assisting in minimizing the redundant data. Since data stored on the blockchain is permanent, hence, the growing size of the ledger is a big concern and needs to be addressed properly.

5.4 Platform throughput and latency

Blockchain latency and throughput are important performance metrics that have a significant impact on the tasks' quality of service (QoS) [111]. In the blockchain, latency is the time taken by the network to verify and execute a transaction in order to store it on the ledger. In the deep learning context, latency refers to the amount of time required by the model to process a single data unit. High transaction execution latency is one of the major research challenges that should be addressed properly as it can affect the performance of processes that require quick decision-making. For example, in the vehicular network, the decision-making of vehicles requires fast transaction processing to minimize traffic jams or accidents. Forking [112] is another problem that occurs frequently as a result of long block propagation delays between miners. Blockchain throughput refers to the number of transactions executed in a unit of time. Every blockchain platform offers different transaction execution throughput and latency. For instance, the Ethereum blockchain platform offers a throughput of 16.5 transactions per second [17, 30]. On the other hand, private blockchain platforms such as Hyperledger Fabric can execute several thousand transactions in one second. Blockchain platforms with high throughput are more suitable for deep learning-based applications. A private blockchain network is more centralized and comprises a limited number of nodes that are distributed locally, so private platforms are much faster than public platforms. The consensus protocols implemented by the public platforms are one of the main reasons for the low transaction execution speed on the public blockchain platforms.

5.5 Cryptocurrencies, deep learning-based consensus protocols, and regulations

The cryptocurrency represents a virtual currency that cannot be double-spent or counterfeited due to blockchain network decentralization and data security that is ensured using encryption and decryption techniques which are based on one-way hashing functions. In comparison to conventional physical currencies, cryptocurrency is a decentralized token that is not regulated by any government agency. Due to the unlimited advantages of cryptocurrency, it has become an investment option similar to the stock exchange. Supervised deep learning models can be used to predict the future of cryptocurrencies. A few studies [113–115] have employed machine learning and deep reinforcement learning techniques to predict the price of crypto. However, due to the technical incompetence of deep learning models in text classification and prediction, the majority of the deep learning models remain imperfect and do not provide fruitful results. Hence, there is a potential research gap in the field of crypto price prediction that can be fulfilled using deep learning-based models. Moreover, many deep learning applications require the models to make quick predictions as the lag in the outcome may cause unwanted consequences. Currently, the generic blockchain consensus algorithms require time on the scale of seconds [116]. Deep learning-specific consensus protocols can be designed based on the proofs related to the quality of the data, optimization techniques, nature of the learning models, and total convergence time of the model. Lack of standards and regulatory frameworks is another challenge to the existing blockchain technology and it affects blockchain adaptability to deep learning frameworks [2, 3, 3]. Extensive research is needed in this direction to propose standards and regulatory frameworks for blockchain technology.

5.6 High-speed computing/storage devices and platform interoperability

Every node in a blockchain network maintains an up-to-date copy of the ledger, thereby enabling fault tolerance in the system. The escalating storage requirements of blockchain nodes can affect the scalability of the blockchain network. The leveraging of centralized-based high-speed data storage and processing devices can lead to several security concerns. The potential of such high-speed storage devices can be exploited when they are employed as blockchain nodes. However, this method is not economically feasible due to high configuration cost of nodes. Also, the failure of centralized storage devices can lead to the breakdown of the deep learning services. The spike in management costs and complicity is another downside of using external high-speed storage devices. Distributed file system such as IPFS [33] or Swarm [35] provides a decentralized storage system, but the sensitive data can be compromised as data on such storage systems are available to the public. However, the encryption-decryption methodology can be used for storing the perceptive information, which comes at the cost of enhanced delays. The interoperable blockchain platforms enable the participants of healthcare and vehicular communication networks to share data and information uninterruptably, securely, quickly, and seamlessly [2, 32]. The platform interoperability is affected by many factors such as the choice of blockchain-supported languages, consensus protocols, cryptographic hashing algorithms, and the type of data being used by the participants.

5.7 Secure economical models

The emerging deep learning models have shown unlimited opportunities and potential in various fields due to technological advancements in existing computing technologies and storage systems. More specifically, in the healthcare sector, deep learning models are used to classify normal patients from unhealthy ones or predict the spread of a disease in a particular community or region. On the other hand, blockchain technology is employed for ensuring data security and integrity. In a network of vehicles, the smart vehicles which are equipped with deep learning models can share the data with the neighboring vehicles using blockchain to avoid traffic-jamming-related issues [117]. However, such models have shown limitations in terms of high cost and less resource efficiency. Currently, existing deep learning models require high-performing computing devices for training purposes [118]; and, the blockchain is an expensive storage medium. More research is needed to propose cost-efficient, resource-friendly, fast, and high-performance-based blockchain-assisted deep learning frameworks.

6 Conclusion

In this paper, we have reviewed the state-of-the-art blockchain-based deep learning frameworks. We presented the key features of blockchain and deep learning along with a detailed discussion on the benefits resulted from their integration. The successful integration of deep learning with blockchain can facilitate in terms of data security and privacy to the existing systems and enhance the QoS in several applications mainly related to healthcare, blockchain security, data traffic management, and vehicular communication in urban areas. We devised a taxonomy to categorize the reported literature in several categories based on seven parameters such as blockchain type, deep learning models, deep learning specific consensus protocols, services, application areas, deployment goals, and data types. The critical aspects of existing blockchain-based deep learning frameworks are analyzed through a comprehensive analysis of the reported frameworks. Finally, we identified and discussed several technological and social challenges and barriers that require further research to unlock the full potential of blockchain in deep learning-based systems. Our concluding remarks along with the key recommendations include:

- Data traceability, immutability, and integrity features of blockchain technology can assist in identifying the volume and type of data collected to train deep learning models. However, the existing blockchain-based systems are incapable of efficiently handling data quality problems, particularly in the healthcare and transportation industries.
- The key performance metrics such as system throughput, execution latency, and block propagation time, data volume, conflicting interests of participants, and smart contract vulnerabilities can significantly impact the effectiveness of existing blockchain-based deep learning systems.
- Private blockchain platforms ensure data privacy through private channels and access control policies. However, public blockchain platforms are prone to data privacy leakage problems because they have a zero-access control policy. Public blockchain platforms can successfully capture the evolution of deep learning models as it progresses and records the model state during its creation, updating, or usage stages.
- The efficiency of blockchain-based applications is highly affected by increasing the size of the blockchain network. Deep learning approaches can be employed for compressing the data as well as minimizing the redundant data.

Acknowledgements This work was supported by the Khalifa University of Science and Technology under Award RCII-2019-002-Research Center for Digital Supply Chain and Operations Management and CIRA-2019-001.

Data availability Not applicable.

Declarations

Conflict of interest The authors have not disclosed any competing interests.

References

1. Ayyoubzadeh, S.M., Ayyoubzadeh, S.M., Zahedi, H., Ahmadi, M., Kalhori, S.R.N.: Predicting COVID-19 incidence through analysis of google trends data in Iran: data mining and deep learning pilot study. JMIR Public Health Surv. **6**(2), e18828 (2020)
2. Ahmad, R.W., Salah, K., Jayaraman, R., Yaqoob, I., Ellahham, S., Omar, M.: The role of blockchain technology in telehealth and telemedicine. Int. J. Med. Inf. **148**, 104399 (2021)
3. Ahmad, R.W., Hasan, H., Yaqoob, I., Salah, K., Jayaraman, R., Omar, M.: Blockchain for aerospace and defense: opportunities and open research challenges. Comput. Ind. Eng. **151**, 106982 (2021)
4. Oh, Y., Park, S., Ye, J.C.: Deep learning COVID-19 features on CXR using limited training data sets. IEEE Trans. Med. imaging **39**(8), 2688–2700 (2020)
5. Shuja, J., Alanazi, E., Alasmary, W., Alashaikh, A.: COVID-19 open source data sets: a comprehensive survey. Appl. Intell. **51**(3), 1296–1325 (2021)
6. Shafay, M., Hassan, T., Velayudhan, D., Damiani, E., Werghi, N.: Deep fusion driven semantic segmentation for the automatic recognition of concealed contraband items. In: SoCPaR, 2020, pp. 550–559
7. Hassan, T., Shafay, M., Akçay, S., Khan, S., Bennamoun, M., Damiani, E., Werghi, N.: Meta-transfer learning driven tensor-shot detector for the autonomous localization and recognition of concealed baggage threats, Nov 2020. https://www.mdpi.com/1424-8220/20/22/6450
8. Berman, D.S., Buczak, A.L., Chavis, J.S., Corbett, C.L.: A survey of Deep learning methods for cyber security. Information **10**(4), 122 (2019)
9. Lawrence, S., Giles, C.L.: Overfitting and Neural networks: conjugate gradient and backpropagation. In: Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millenniu, vol. 1. IEEE, 2000, pp. 114–119
10. Al Ridhawi, I., Aloqaily, M., Jararweh, Y.: An incentive-based mechanism for volunteer computing using blockchain. ACM Trans. Internet Technol. **21**(4), 1–22 (2021)
11. Tan, L., Xiao, H., Yu, K., Aloqaily, M., Jararweh, Y.: A blockchain-empowered crowdsourcing system for 5g-enabled smart cities. Comput. Stand. Interfaces **76**, 103517 (2021)
12. Narayan, S., Tagliarini, G.: An analysis of underfitting in MLP networks. In: Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005., vol. 2. IEEE, 2005, pp. 984–988
13. Shuja, J., Bilal, K., Alasmary, W., Sinky, H., Alanazi, E.: Applying machine learning techniques for caching in edge networks: a comprehensive survey, arXiv preprint arXiv:2006.16864, 2020
14. Brown, T.B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell A., et al.: Language models are few-shot learners. arXiv preprint arXiv:2005.14165, 2020
15. Wang, L., Von Laszewski, G., Younge, A., He, X., Kunze, M., Tao, J., Fu, C.: Cloud computing: a perspective study. New Gen. Comput. **28**(2), 137–146 (2010)
16. Shiraz, M., Gani, A., Ahmad, R.W., Shah, S.A.A., Karim, A., Rahman, Z.A.: A lightweight distributed framework for computational offloading in mobile cloud computing. PLoS ONE **9**(8), e102270-9 (2014)
17. Ahmad, R.W., Hasan, H., Jayaraman, R., Salah, K., Omar, M.: Blockchain applications and architectures for port operations and logistics management. Res. Transp. Business Manag. (2021). https://doi.org/10.1016/j.rtbm.2021.100620
18. Bach, L.M., Mihaljevic, B., Zagar, M.: Comparative analysis of blockchain consensus algorithms. In: 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). IEEE 2018, pp. 1545–1550 (2018)

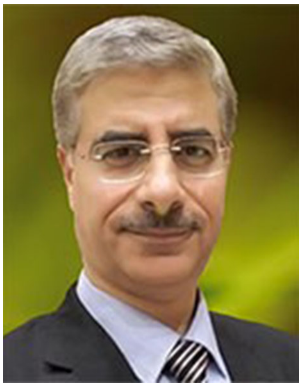
19. Zheng, Z., Xie, S., Dai, H.-N., Chen, W., Chen, X., Weng, J., Imran, M.: An overview on smart contracts: Challenges, advances and platforms. Fut. Gen. Comput. Syst. **105**, 475–491 (2020)
20. Ren, W., Hu, J., Zhu, T., Ren, Y., Choo, K.-K.R.: A flexible method to defend against computationally resourceful miners in blockchain Proof-of-work. Inf. Sci. **507**, 161–171 (2020)
21. Vashchuk, O., Shuwar, R.: Pros and cons of consensus algorithm Proof-of-stake: difference in the network safety in Proof-of-work and proof-of-stake. Electron. Inf. Technol. **9**(9), 106–112 (2018)
22. Singh, P.K., Singh, R., Nandi, S.K., Nandi, S.: Managing smart home appliances with Proof-of-authority and blockchain. In: International Conference on Innovations for Community Services. Springer, Berlin, 2019, pp. 221–232
23. Jerry Cuomo, P.: How blockchain adds trust to AI and IoT, Aug 2020. https://www.ibm.com/blogs/blockchain/2020/08/how-blockchain-adds-trust-to-ai-and-iot/
24. Sarpatwar, K., Vaculin, R., Min, H., Su, G., Heath, T., Ganapavarapu, G., Dillenberger, D.: Towards enabling trusted artificial intelligence via blockchain. In: Policy-based autonomic data governance. Springer, Berlin, 2019, pp. 137–153
25. Shinde, R., Patil, S., Kotecha, K., Ruikar, K.: Blockchain for securing ai applications and open innovations. J. Open Innov. **7**(3), 189 (2021)
26. Salah, K., Rehman, M.H.U., Nizamuddin, N., Al-Fuqaha, A.: Blockchain for AI: review and open research challenges. IEEE Access **7**, 127–149 (2019)
27. Tanwar, S., Bhatia, Q., Patel, P., Kumari, A., Singh, P.K., Hong, W.-C.: Machine learning adoption in blockchain-based smart applications: the challenges, and a way forward. IEEE Access **8**, 474–488 (2020)
28. Nakamoto, S.: Bitcoin: a peer-to-peer electronic cash system, Decentralized Business Review, p. 21260, 2008
29. Yeow, K., Gani, A., Ahmad, R.W., Rodrigues, J.J., Ko, K.: Decentralized consensus for edge-centric internet of things: a review, taxonomy, and research issues. IEEE Access **6**, 1513–1524 (2017)
30. Wood, D.D.: Ethereum: a secure decentralised generalised transaction ledger, 2014
31. Sookhak, M., Jabbarpour, M.R., Safa, N.S., Yu, F.R.: Blockchain and smart contract for access control in healthcare: a survey, issues and challenges, and open issues. J. Netw. Comput. Appl. **178**, 102950 (2021)
32. Ahmad, R.W., Salah, K., Jayaraman, R., Yaqoob, I., Ellahham, S., Omar, M.: Blockchain and COVID-19 pandemic: applications and challenges, IEEE TechRxiv, 2020
33. Benet, J.: IPFS - Content Addressed, versioned, P2P file system, 2014
34. Lakshman, A., Malik, P.: Cassandra: a decentralized structured storage system. SIGOPS Oper. Syst. Rev. **44**(2), 35–40 (2010). https://doi.org/10.1145/1773912.1773922
35. Hartman, J., Murdock, I., Spalink, T.: The Swarm scalable storage system. In: Proceedings of 19th IEEE International Conference on Distributed Computing Systems (Cat. No.99CB37003), 1999, pp. 74–81
36. Wilkinson, S., Boshevski, T., Brandoff, J., Buterin, V.: Storj: a peer-to-peer cloud storage network, 2014
37. Chan, T.-H., Jia, K., Gao, S., Lu, J., Zeng, Z., Ma, Y.: PCANet: a simple deep learning baseline for image classification? IEEE Trans. Image Process. **24**(12), 5017–5032 (2015)
38. Zhao, Z.-Q., Zheng, P., Xu, S.-T., Wu, X.: Object detection with deep learning: a review. IEEE Trans. Neural Netw. Learn. Syst. **30**(11), 3212–3232 (2019)
39. Daily, M., Medasani, S., Behringer, R., Trivedi, M.: Self-driving cars. Computer **50**, 18–23 (2017)
40. Hassan, T., Hassan, B., El-Baz, A., Werghi, N.: A dilated residual hierarchically fashioned segmentation framework for extracting Gleason tissues and grading prostate cancer from whole slide images, 2021
41. Sheridan, T.B.: Human-robot interaction: status and challenges. Hum Factors **58**(4), 525–532 (2016)
42. Mark, R.: Publish Date: May 27, The difference between Artificial intelligence, Machine learning and Deep learning. https://www.intel.la/content/www/xl/es/artificial-intelligence/posts/difference-between-ai-machine-learning-deep-learning.html
43. McMahan, B., Moore, E., Ramage, D., Hampson, S., y Arcas, B.A.: Communication-efficient learning of deep networks from decentralized data. In: Artificial intelligence and statistics. PMLR, 2017, pp. 1273–1282
44. Hassan, T., Akcay, S., Bennamoun, M., Khan, S., Werghi, N.: A novel incremental learning driven instance segmentation framework to recognize highly cluttered instances of the contraband items. IEEE Transactions on Systems, Man, and Cybernetics: Systems, To be published, 2022
45. Hassan, T., Hassan, B., Akram, M.U., Hashmi, S., Taguri, A.H., Werghi, N.: Incremental cross-domain adaptation for robust retinopathy screening via Bayesian deep learning. IEEE Trans. Instrum. Measur. **70**, 1–14 (2021)
46. Hassan, T., Aslam, S., Jang, J.W.: Fully automated multi-resolution channels and multithreaded spectrum allocation protocol for iot based sensor nets. IEEE Access **6**, 545–556 (2018)
47. Mahdavifar, S., Ghorbani, A.A.: Application of deep learning to cybersecurity: a survey. Neurocomputing **347**, 149–176 (2019)
48. Salloum, S.A., Alshurideh, M., Elnagar, A., Shaalan, K.: Machine learning and deep learning techniques for cybersecurity: a review. AICV **2020**, 50–57 (2020)
49. P. by S. O'Dea and A. 6, Smartphone users 2026, Aug 2021. https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/
50. He, C., Tan, C., Tang, H., Qiu, S., Liu, J.: Central server free federated learning over single-sided trust social networks, arXiv preprint arXiv:1910.04956, 2019
51. Marr, B.: Artificial intelligence and blockchain: 3 major benefits of combining these two mega-trends, Mar 2018. https://www.forbes.com/sites/bernardmarr/2018/03/02/artificial-intelligence-and-blockchain-3-major-benefits-of-combining-these-two-mega-trends/?sh=604fcaa04b44
52. Campbell, D.: Combining AI and blockchain to push frontiers in healthcare, Nov 2018. https://www.macadamian.com/learn/combining-ai-and-blockchain-in-healthcare/
53. Castelló Ferrer, E.: The blockchain: a new framework for robotic Swarm systems, Advances in Intelligent Systems and Computing, p. 1037–1058, Oct 2018. http://dx.doi.org/10.1007/978-3-030-02683-7_77
54. Janson, S., Merkle, D., Middendorf, M.: A decentralization approach for swarm intelligence algorithms in networks applied to multi Swarm PSO. Int. J. Intell. Comput. Cybern. **1**(1), 25–45 (2008). https://doi.org/10.1108/17563780810857112
55. Hassan, K., Tahir, F., Rehan, M., Ahn, C.K., Chadli, M.: On relative-output feedback approach for group consensus of clusters of multiagent systems. IEEE Trans. Cybern. **1–12**, 2021 (2021)
56. Magazzeni, D., McBurney, P., Nash, W.: Validation and verification of smart contracts: a research agenda. Computer **50**(9), 50–57 (2017)
57. Open source P2P digital currency. https://litecoin.org/
58. Dinh, T.T.A., Wang, J., Chen, G., Liu, R., Ooi, B.C., Tan, K.-L.: BLOCKBENCH: A Framework for Analyzing Private Blockchains, 2017

59. Li, Z., Kang, J., Yu, R., Ye, D., Deng, Q., Zhang, Y.: Consortium blockchain for secure energy trading in industrial internet-of-things. IEEE Trans. Ind. Inf. **14**(8), 3690–3700 (2018)
60. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. Commun. ACM **60**(6), 84–90 (2017). https://doi.org/10.1145/3065386
61. Rumelhart, D.E., McClelland, J.L.: Learning Internal Representations by Error Propagation, pp. 318–362 (1987)
62. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**(8), 1735–1780 (1997). https://doi.org/10.1162/neco.1997.9.8.1735
63. Chung, J., Gulcehre, C., Cho, K., Bengio, Y.: Empirical evaluation of gated recurrent neural networks on sequence modeling. In: NIPS 2014 Workshop on Deep Learning, December 2014
64. Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative Adversarial Nets, Ser NIPS'14, pp. 2672–2680. MIT Press, Cambridge (2014)
65. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., et al.: Human-level control through deep reinforcement learning. Nature **518**(7540), 529–533 (2015). https://doi.org/10.1038/nature14236
66. Bronstein, M.M., Bruna, J., LeCun, Y., Szlam, A., Vandergheynst, P.: Geometric deep learning: going beyond Euclidean data. IEEE Signal Process. Mag. **34**(4), 18–42 (2017)
67. Merlina, A.: BlockML: a useful proof of work system based on machine learning tasks. In: Proceedings of the 20th International Middleware Conference Doctoral Symposium, 2019
68. Kuo, T.-T., Ohno-Machado, L.: Modelchain: decentralized privacy-preserving healthcare predictive modeling framework on private blockchain networks, arXiv preprint arXiv:1802.01746, 2018
69. Li, Y., Chen, C., Liu, N., Huang, H., Zheng, Z., Yan, Q.: A blockchain-based decentralized federated learning framework with committee consensus. IEEE Netw. **35**(1), 234–241 (2020)
70. Bravo-Marquez, F., Reeves, S., Ugarte, M.: Proof-of-learning: a blockchain consensus mechanism based on machine learning competitions. In: 2019 IEEE International Conference on Decentralized Applications and Infrastructures (DAPPCON). IEEE, 2019, pp. 119–124
71. Lan, Y., Liu, Y., Li, B.: Proof of Learning (PoLe): Empowering machine learning with consensus building on blockchains, 2020
72. Salah, K., Alfalasi, A., Alfalasi, M., Alharmoudi, M., Alzaabi, M., Alzyeodi, A., Ahmad, R.W.:'Iot-enabled shipping container with environmental monitoring and location tracking. In: IEEE 17th Annual Consumer Communications & Networking Conference (CCNC). IEEE 2020, pp. 1–6 (2020)
73. Wang, C., Cheng, X., Li, J., He, Y., Xiao, K.: A survey: applications of blockchain in the internet of vehicles. EURASIP J. Wirel. Commun. Netw. **2021**(1), 1–16 (2021)
74. Dimitrakopoulos, G., Demestichas, P.: Intelligent transportation systems. IEEE Vehicular Technol. Mag. **5**(1), 77–84 (2010)
75. Ni, J., Lin, X., Zhang, K., Shen, X.: Privacy-preserving real-time navigation system using vehicular crowdsourcing. In: IEEE 84th Vehicular Technology Conference (VTC-Fall). IEEE 2016, pp. 1–5 (2016)
76. Xie, B., Chen, Y., Xu, M.: Evaluating urban traffic jam based on a urban cell transmission model (UCTM). In: 2012 12th International Conference on ITS Telecommunications. IEEE, 2012, pp. 211–215
77. Siddiqui, S.T., Ahmad, R., Shuaib, M., Alam, S.: Blockchain security threats, attacks and countermeasures. Adv. Intell. Syst. Comput. **1097**, 51–62 (2020)
78. Awan, S., Li, F., Luo, B., Liu, M.: Poster: a reliable and accountable privacy-preserving federated learning framework using the blockchain. In: Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, 2019, pp. 2561–2563
79. Philip, A.O., Saravanaguru, R.K.: Secure incident & evidence management framework (SIEMF) for internet of vehicles using deep learning and blockchain. Open Comput. Sci. **10**(1), 408–421 (2020)
80. Scicchitano, F., Liguori, A., Guarascio, M., Ritacco, E., Manco, G.: A deep learning approach for detecting security attacks on blockchain. In: ITASEC, 2020, pp. 212–222
81. Kurri, V., Raja, V., Prakasam, P.: Cellular traffic prediction on blockchain-based mobile networks using LSTM model in 4G LTE network. Peer-to-Peer Netw. Appl. **14**(3), 1088–1105 (2021)
82. Kim, H., Park, J., Bennis, M., Kim, S.-L.: Blockchained on-device federated learning. IEEE Commun. Lett. **24**(6), 1279–1283 (2019)
83. Vora, J., Nayyar, A., Tanwar, S., Tyagi, S., Kumar, N., Obaidat, M.S., Rodrigues, J.J.: BHEEM: a blockchain-based framework for securing electronic health records. In: IEEE Globecom Workshops (GC Wkshps). IEEE 2018, pp. 1–6 (2018)
84. Omaar, J., Schwerin, S., McMullen, G.: Forever isn't free: The cost of storage on a blockchain database. With assistance of Simon Schwerin, McMullen Greg. https://medium.com/ipdb-blog/forever-isnt-freethe-cost-of-storage-on-a-blockchain-database-59003f63e01, vol. 2, no. 28, p. 2019, 2017
85. Abraham, M., Vyshnavi, A., Srinivasan, C., Namboori, P.: Healthcare security using blockchain for pharmacogenomics. J. Int. Pharm. Res. **46**, 529–533 (2019)
86. Zheng, X., Mukkamala, R.R., Vatrapu, R., Ordieres-Mere, J.: Blockchain-based personal health data sharing system using cloud storage. In: 2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom). IEEE, 2018, pp. 1–6
87. Liang, G., Zheng, L.: A transfer learning method with deep residual network for pediatric pneumonia diagnosis. Comput. Methods Progr. Bomed. **187**, 104964 (2020)
88. Dutta, A., Batabyal, T., Basu, M., Acton, S.T.: An efficient convolutional neural network for coronary heart disease prediction. Expert Syst. Appl. **159**, 113408 (2020)
89. Hasan, N.I., Bhattacharjee, A.: Deep learning approach to cardiovascular disease classification employing modified ECG signal from empirical mode decomposition. Biomed. Signal Process. Control **52**, 128–140 (2019)
90. Arunkumar, R., Karthigaikumar, P.: Multi-retinal disease classification by reduced deep learning features. Neural Compu. Appl. **28**(2), 329–334 (2017)
91. Juneja, A., Marefat, M.: Leveraging blockchain for retraining deep learning architecture in patient-specific arrhythmia classification. In: 2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI). IEEE, 2018, pp. 393–397
92. Rohani, M., Gingras, D., Gruyer, D.: A novel approach for improved vehicular positioning using cooperative map matching and dynamic base station DGPS concept. IEEE Trans. Intell. Transp. Syst. **17**(1), 230–239 (2015)
93. Wang, S., Sun, S., Wang, X., Ning, Z., Rodrigues, J.J.: Secure crowdsensing in 5G internet of vehicles: When deep reinforcement learning meets blockchain. In: IEEE Consumer Electronics Magazine, 2020
94. Li, C., Fu, Y., Yu, F.R., Luan, T.H., Zhang, Y.: Vehicle position correction: a vehicular blockchain networks-based GPS error sharing framework. IEEE Trans. Intell. Transp. Syst. **22**(2), 898–912 (2020)
95. Hassija, V., Gupta, V., Garg, S., Chamola, V.: Traffic jam probability estimation based on blockchain and deep neural networks. IEEE Trans. Intell. Transp. Syst. **22**, 7 (2020)

96. Liu, Y., James, J., Kang, J., Niyato, D., Zhang, S.: Privacy-preserving traffic flow prediction: a federated learning approach. IEEE Internet Things J. **7**(8), 7751–7763 (2020)
97. Bhattacharya, P., Tanwar, S., Bodke, U., Tyagi, S., Kumar, N.: Bindaas: blockchain-based deep-learning as-a-service in healthcare 4.0 applications. IEEE Trans. Netw. Sci. Eng. **8**, 2 (2019)
98. Singh, M., Aujla, G.S., Bali, R.S.: A deep learning-based blockchain mechanism for secure Internet-of-Drones environment. IEEE Trans. Intell. Transp. Syst. **22**(7), 4404–4413 (2021)
99. Nguyen, D., Pathirana, P., Ding, M., Seneviratne, A.: Secure computation offloading in blockchain based IoT networks with deep reinforcement learning. IEEE Trans. Netw. Sci. Eng. (2021). https://doi.org/10.48550/arXiv.1908.07466
100. Dey, S.: Securing majority-attack in blockchain using machine learning and algorithmic game theory: a proof of work. In: 10th Computer Science and Electronic Engineering (CEEC). IEEE 2018, pp. 7–10 (2018)
101. Rimal, B.P., Van, D.P., Maier, M.: Mobile-edge computing vs. centralized cloud computing in fiber-wireless access networks. In: 2016 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). IEEE, 2016, pp. 991–996
102. Zhang, J., Chen, B., Zhao, Y., Cheng, X., Hu, F.: Data security and privacy-preserving in Edge computing paradigm: survey and open issues. IEEE Access **6**, 209–237 (2018)
103. Rosenfeld, M.: Analysis of Hashrate-based double spending. arXiv preprint arXiv:1402.2009, 2014
104. Rehman, A., Rathore, M.M., Paul, A., Saeed, F., Ahmad, R.W.: Vehicular traffic optimisation and even distribution using ant colony in smart city environment. IET Intell. Transp. Syst. **12**(7), 594–601 (2018)
105. Ni, F., Zhang, J., Noori, M.N.: Deep learning for data anomaly detection and data compression of a long-span suspension bridge. Comput. Aid. Civ. Infrastruct. Eng. **35**(7), 685–700 (2020)
106. Ouellette, M., Shaw, M.L.: For a decentralized vaccine passport, Health Policy, 2021
107. Khalid, T., Abbasi, M.A.K., Zuraiz, M., Khan, A.N., Ali, M., Ahmad, R.W., Rodrigues, J.J., Aslam, M.: A survey on privacy and access control schemes in fog computing. Int. J. Commun. Syst. **34**(2), e4181 (2021)
108. Wrona, K., Jarosz, M.: Use of blockchains for secure binding of metadata in military applications of iot. In: IEEE 5th World Forum on Internet of Things (WF-IoT). IEEE 2019, pp. 213–218 (2019)
109. Abbasi, G.A., Tiew, L.Y., Tang, J., Goh, Y.-N., Thurasamy, R.: The adoption of cryptocurrency as a disruptive force: Deep learning-based dual stage structural equation modelling and artificial neural network analysis. PLoS ONE **16**(3), e0247582 (2021)
110. Demertzis, K., Iliadis, L., Tziritas, N., Kikiras, P.: Anomaly detection via blockchained deep learning smart contracts in industry 4.0. Neural Comput. Appl. **32**(23), 361–378 (2020)
111. Irshad, T., Shan, R.-U., Ahmad, R.W., Khalid, A., Ab Hamid, S.H.: Multi-rat based adaptive quality of service (QOS) management in WBAN. Malays. J. Comput. Sci. **33**(4), 252–269 (2020)
112. da Silva, F.J.C., Damsgaard, S.B., Sorensen, M.A.M., Marty, F., Altariqi, B., Chatzigianni, E., Madsen, T.K., Schwefel, H.P.: Analysis of blockchain forking on an ethereum network. In: European Wireless,: 25th European Wireless Conference. VDE 2019, pp. 1–6 (2019)
113. Patel, M.M., Tanwar, S., Gupta, R., Kumar, N.: A deep learning-based cryptocurrency price prediction scheme for financial institutions. J. Inf. Secur. Appl. **55**, 102583 (2020)
114. Jiang, Z., Liang, J.: Cryptocurrency portfolio management with deep reinforcement learning. In: Intelligent Systems Conference (IntelliSys). IEEE 2017, 905–913 (2017)
115. Lahmiri, S., Bekiros, S.: Cryptocurrency forecasting with deep learning chaotic neural networks. Chaos Solitons Fractals **118**, 35–40 (2019)
116. Sankar, L.S., Sindhu, M., Sethumadhavan, M.: Survey of consensus protocols on blockchain applications. In :2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS). IEEE, 2017, pp. 1–5
117. Yang, F., Wang, S., Li, J., Liu, Z., Sun, Q.: An overview of internet of vehicles. China Commun. **11**(10), 1–15 (2014)
118. Strubell, E., Ganesh, A., McCallum, A.: Energy and policy considerations for deep learning in nlp. arXiv preprint arXiv:1906.02243, 2019

Muhammad Shafay received his undergraduate degree in Electrical Engineering from PIEAS, Pakistan. He is currently pursuing a Master's degree in electrical engineering and computer science with Khalifa University, UAE. His research interests include deep learning, blockchain, computer vision, and image processing.

Raja Wasim Ahmad received his Ph.D. in Computer Science from the Center for Mobile Cloud Computing Research (C4MCCR), Faculty of Computer Science and Information Technology (FSKTM), University of Malaya, Malaysia. He completed his masters in computer science from COMSATS University Islamabad (CUI), Abbottabad campus, in 2011. Currently, he is working as a Postdoctoral research fellow at Khalifa University, Abu Dhabi, UAE. He has published several research papers in leading journals and conferences. Most of his research contributions are published in top-cited journals such as IEEE Transactions on Emerging Topics in Computing, IEEE Systems Journal, IEEE Access, Journal of Networks and Computer Applications, Journal of Grid Computing, Cluster Computing, International Journal of Communication Systems, International Journal of Information Management, Journal of Systems and Software, Journal of Supercomputing, IET Intelligent Transport Systems, and Renewable and Sustainable Energy Reviews. Dr. Raja is also serving as a reviewer for several journals and conferences. His research work is well acknowledged in national and international conferences. One of his research articles has received an award of the best paper in SPECTS symposium held in 2018 in France.



Khaled Salah is a full Professor at the Department of Electrical and Computer Engineering, Khalifa University, UAE. He received the B.S. degree in Computer Engineering with a minor in Computer Science from Iowa State University, USA, in 1990, the M.S. degree in Computer Systems Engineering from Illinois Institute of Technology, USA, in 1994, and the Ph.D. degree in Computer Science from the same institution in 2000. Khaled has over 220 publications and 3 US patents, has been giving a number of international keynote speeches, invited talks, tutorials, and research seminars on the subjects of Blockchain, IoT, Fog and Cloud Computing, and Cybersecurity. Khaled served as the chair of the Track Chair of IEEE Globecom 2018 on Cloud Computing. Khaled is an Associate Editor of IEEE Blockchain Tech Briefs, and a member of IEEE Blockchain Education Commitee. Khaled is now leading a number of projects on how to leverage blockchain for Healthacare, 5G Networks, Combating Deepfake Videos, Supply Chain Management, and AI.

Ibrar Yaqoob (S'16-M'18-SM'19) is a research scientist at the Khalifa University, where he had joined as a postdoctoral fellow in October 2019. Previously, he worked as a research professor at the Department of Computer Science and Engineering, Kyung Hee University, South Korea, where he completed his postdoctoral fellowship. Prior to that, he received his Ph.D. (Computer Science) from the University of Malaya, Malaysia, in 2017. He worked as a research assistant at the Centre for Mobile Cloud Computing Research (C4MCCR), University of Malaya. His numerous research articles are very famous and among the most downloaded in top journals. He has been selected as a highly cited researcher (HCR) worldwide for the year 2021 by Clarivate (Web of Science). He is currently serving as an editor for various IEEE, Elsevier, and Springer journals. He has been involved in a number of conferences and workshops in various capacities. His research interests include blockchain, mobile edge/cloud computing, the Internet of Things, computer networks, and big data.

Raja Jayaraman is an Associate Professor in the Department of Industrial & Systems Engineering at Khalifa University, Abu Dhabi, UAE. He received his Ph.D. in Industrial Engineering from Texas Tech University, a Master of Science degree in Industrial Engineering from New Mexico State University, a Master and Bachelors in Mathematics from India. His expertise is in multi-criteria optimization techniques applied to diverse applications including supply chain and logistics, healthcare, energy, environment and sustainability. Raja's research interests are primarily focused in using blockchain technology, systems engineering and process optimization techniques to characterize, model and analyze complex systems with applications to supply chains, maintenance operations planning and healthcare delivery. His post doctorial research was centered on technology adoption and implementation of innovative practices in the healthcare supply chains and service delivery. He has led several successful research projects and pilot implementations in the area of supply chain data standards adoption in the US healthcare system. His research has appeared in top rated journals including: Annals of Operations Research, IISE Transactions, Energy Policy, Applied Energy, Knowledge Based Systems, IEEE Access, Journal of Theoretical Biology, Engineering Management Journal and others

Mohammed Omar is currently a full Professor and the Founding Chair of the Department of Engineering Systems and Management (currently renamed Industrial and Systems Engineering). Prior to joining the Masdar Institute/KUST, he was an Associate Professor and a Graduate Coordinator with Clemson University, Clemson, SC, USA. He was a part of the Founding Faculty Cohort of Clemson University research park in Greenville, SC, USA. He has over 100 publications in the areas of product lifecycle management, knowledge-based manufacturing, and automated testing systems, in addition to authoring several books and book chapters. He holds four U.S. and international patents. He was named a Tennessee Valley Authority Fellow of two consecutive years during the Ph.D. degree, in addition to being a Toyota Manufacturing Fellow. His professional career includes a Postdoctoral service at the Center for Robotics and Manufacturing Systems CRMS, and a Visiting Scholar at the Toyota Instrumentation and Engineering Division, Toyota Motor Company, Japan. His group graduated seven Ph.D. dissertations and over 35 M.Sc. theses. Four Ph.D. students are currently on academic ranks in U.S. universities. His work has been recognized by the U.S. Society of manufacturing engineers SME through the Richard L. Kegg Award. He has also received the SAE Foundation Award for Manufacturing Leadership. In addition, he has received the Murray Stokely Award from the College of Engineering, Clemson University. He has also led an NSF I/UCRC Center and a part of the DoE GATE Center of Excellence in Sustainable Mobility Systems. His current research interests include capabilities in composite fabrication and manufacturing analytics at a laboratory Masdar City

Campus. His current research group supported two Postdoctoral Scholar's Career Planning to become an Assistant Professor at the Texas A&M (TAMUQ), in 2013, and the University of Sharjah, in 2015. He currently serves as an Editor-in-Chief for the Journal of Material Science Research (Part of the Canadian Research Center), and as an Associate Editor for the Journal of Soft Computing (a Springer), handling the areas of decision science, knowledge-based systems, in addition to his membership on several editorial boards and conference organizations. Furthermore, he serves on the Advisory Board of the Strata PJSC (part of Mubadala Aerospace).