Evolutionary Intelligence
https://doi.org/10.1007/s12065-024-00949-0

RESEARCH PAPER

# Deep learning applications in the Internet of Things: a review, tools, and future directions

**Parisa Raoufi[1] · Atefeh Hemmati[2] · Amir Masoud Rahmani[3]**

Received: 26 December 2023 / Revised: 20 April 2024 / Accepted: 27 May 2024


**Abstract**
The emergence of the Internet of Things (IoT) has enabled the proliferation of interconnected devices and sensors, generating vast amounts of often complex and unstructured data. Deep learning (DL), a subfield of machine learning (ML), has shown great promise in addressing the challenges of processing and analyzing such data. Considering the increasing importance of DL and data analysis, we decided to review the articles of the last few years in this field to pave the way for researchers. In this article, we used the systematic literature review (SLR) method, and in line with that, we selected and analyzed 56 articles published from 2019 to April 2024. We first discuss the DL models used in the IoT field and clarify their specific use cases. Secondly, we outline an analysis of research areas in DL-based IoT. In addition, our research extends to the tools and simulators used to evaluate studies in the DL-based IoT domain. We also examine the DL-based IoT research datasets. Finally, our review identifies future directions and open issues in DL-based IoT. We aim to contribute to an accurate understanding of the current state, challenges, and potential breakthroughs at the intersection of DL and the IoT.

**Keywords** Internet of Things · Artificial intelligence · Deep learning · Machine learning · Convolutional neural networks

## 1 Introduction

The world is changing with the development of the Internet of Things (IoT) and artificial intelligence (AI), particularly their combination. Thanks to the IoT, vast amounts of data are gathered from various sources. However, processing and analyzing the data generated by the numerous IoT devices takes a lot of work. Investing in new technologies is necessary to achieve the objectives and maximize the potential of IoT devices [1, 2].

The convergence of AI and the IoT may change the industry, business, and economic operations. AI uses the IoT to build intelligent machines that mimic intelligent behavior and assist in making decisions with little to no human involvement. Both professionals and laypeople can benefit from the combination of the two. IoT deals with the interaction of devices over the Internet, whereas AI enables devices to learn from their data and experiences. This article explains why combining IoT and AI is necessary [3, 4]. Deep learning (DL) is an essential and rapidly growing field of AI that can change our lives. DL can provide the best output to humans by examining the large and large data available on the Internet and help them a lot in life [5].

DL's significance stems from its capacity to resolve intricate problems beyond machines' capability in the past. Machine learning (ML) algorithms typically acquire the ability to react and behave in the situations presented to them. The significance of DL also lies in its scalability to handle huge datasets or big data. Since there is a lot of data in our daily lives mixed with technology, DL algorithms can examine and process a large amount of data and are considered a practical solution for all industries. DL also has the potential to revolutionize the way humans interact with technology. DL is also used in self-driving cars in most parts of the world [2].

Adding additional technologies, like ML, to an intelligent system boosts its efficiency and production. We can comprehend and enhance the data gathered from physical devices thanks to ML, DL, and AI because IoT devices aim to collect

✉ Amir Masoud Rahmani
rahmania@yuntech.edu.tw

[1] Department of Computer Engineering, Central Tehran Branch, Islamic Azad University, Tehran, Iran

[2] Department of Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

[3] Future Technology Research Center, National Yunlin University of Science and Technology, Yunlin 64002, Taiwan

Published online: 08 June 2024

and use data. To provide greater value to IoT, expert systems are employed to analyze data collected from connected devices. Software with machine intelligence features analyzes raw data collected and aggregated by a linked device network. After careful examination, the finished product includes insightful data[2, 5, 6].

Thankfully, improvements in ML paradigms have made it possible to use data analytics in IoT applications. These services are the foundation for IoT applications because DL models have demonstrated significant results in various fields, including image recognition, information retrieval, speech recognition, natural language processing, indoor localization, physiological and psychological state recognition, etc. [4, 7].

AI depends on the IoT for its future and worth. AI has shown a notable surge in recent years due to DL. To operate effectively, DL needs a lot of data. Numerous IoT nodes gather this quantity of data from the surrounding environment, and we see the daily growth of these nets and nodes. As a result, the IoT enhances AI's performance. Improved AI performance will also help show off the IoTs' potential and gain widespread acceptance. Both technologies will profit from this cycle in this way. IoT is, therefore, made more effective and beneficial by the application of AI.

Considering the daily progress of AI, especially DL, and also to keep the information updated to make it easier and smoother for researchers, we decided to write this systematic literature review (SLR) article. Some review studies done in the DL-based IoT field, for instance, Bhattacharya et al. [3], focused on the state-of-the-art applications of DL in smart cities. The authors surveyed various DL applications in smart city data, providing suggestions for future research areas. However, the review lacked discussion on future challenges and did not propose a taxonomy related to their work. Lakshman et al. [4] focused on the basics of IoT, data generation and processing, and DL techniques in the IoT context. The authors explored IoT fundamentals, DL techniques, and key reporting initiatives for DL in the IoT domain. They covered advantages, uses, and challenges associated with DL's application enabling IoT applications.

Bolhasani et al. [2] focused on DL in IoT applications within healthcare systems. The authors presented theoretical notions and technical taxonomy related to DL. They analyzed existing research to highlight key DL applications in healthcare and medical sciences. The study provides a comprehensive overview of DL applications in healthcare IoT, offering insights into theoretical frameworks, practical benefits, and areas for future investigation.

Previous survey papers focused solely on a single IoT application, such as smart cities or healthcare, neglecting other aspects. Alternatively, some surveys had shortcomings, which our SLR addresses. To mention previous papers' weaknesses, we can point out that some studies do not provide taxonomy [3–6]. In some other reviews, the review type needs to be clarified [5, 7]. In others, future work and open issues are not considered [3]. Some others still need to provide a comprehensive and sufficient review [3]. In this article, we tried to cover these shortcomings and provide a comprehensive and complete review so that researchers can quickly find an overview in this field and save time. The main contributions of our work are as follows:

- Presenting a new taxonomy in the field of DL-based IoT
- Outlining key areas where future research could improve the use of the DL technique in IoT
- Exploring tools and simulators for evaluation of DL-based IoT research results
- Discussing common evaluation criteria used in the DL-based IoT domain
- Exploring challenges and future work methods in the field of DL-based IoT

The remainder of this paper is organized as follows: Section 2 discusses some related concepts. Section 3 provides some related work. The methodology presented in Section 4 also illustrates the taxonomy. Section 5 offers the organization and categorization of research articles. Section 6 includes research questions, answers, and results. Finally, Section 7 concludes this article.

## 2 Overview of concepts and terminology

This section provides a basic understanding of the key concepts and terms necessary to understand DL-based IoT better. Also, to get familiar with these concepts and words used in the rest of the article.

### 2.1 Overview of concepts

Like how the human brain and body interact, AI and IoT are closely linked. The human body gathers sensory information from touch, hearing, and sight. The human brain processes this information, transforming light into recognized objects and sounds into speech that can be understood. Following a choice, the human brain communicates with the body to direct actions like selecting a thing or talking. The IoT's network of interconnected sensors provides raw data on what is happening in the outside world, much like the human body. AI is like a human brain; it understands data and decides what to do with it. AI is used in the IoT to understand the environment and make decisions. AI is divided into subbranches such as ML, DL, etc. DL is also a new branch of the artificial neural network(ANN) [3, 8, 9].

The IoT contains many highly complicated data, such as various sensors, radio frequency records, video and image

data, descriptive data, location information, environmental parameters, and sensor network data. These factors pose significant IoT management, analytical, and data mining issues. As a result, there is a critical need for data analysis systems to investigate and evaluate the broad and continuous flow of real-world data applications, including monitoring of temperature, air pollution tracking, stock market, and information security, among others. IoT uses AI to comprehend its surroundings and make decisions. ML, DL, and other sub-branches of AI are separated. Regarding ANN, DL is a new topic [3, 10].

AI and the IoT are pushing the boundaries of data processing and smart business, and this trend will continue in the coming years. The convergence of AI and the IoT has created enormous business potential worldwide. IoT sensors detect external information and replace it with a signal that humans and machines can distinguish from each other. Then, it is time for AI to help build smart machines that support this data to support the decision-making process with minimal or no human intervention [2, 4, 7].

DL is a branch of ML related to extensive layers of neural networks. A deep neural network (DNN) analyzes data in the way a person approaches a problem. DL is widely associated with ML and AI. Of course, this may increase the possibility of misunderstanding and misunderstanding. In other words, DL is an AI function that replicates how the human brain processes data, creates patterns, and applies those patterns to decision-making. In other terms, DL is a sub-branch of ML that processes auditory and visual sensory inputs using many layers of linear transformations [7, 11].

By breaking down each complex notion into simpler ones using this strategy, the computer eventually reaches basic concepts for which it can make judgments, negating the need for constant human oversight to provide the machine with the knowledge it needs. How information is delivered is a crucial DL concern. The machine should be given information in a way that ensures it obtains the crucial details it needs to make decisions in the shortest amount of time possible [3].

Due to Fig. 1, everyone is very excited about data, IoT, and AI. IoT refers to numerous things, including objects and devices in our environment connected to the Internet, and can be managed and controlled by applications on smartphones and tablets. In simple terms, the IoT connects sensors and devices with a network through which they can interact with each other and their users. This concept can be as simple as connecting a smartphone to a TV or as complex as monitoring urban infrastructure and traffic. This network includes many devices around us, from washing machines and refrigerators to our clothes [12].

In today's technology, ML means eliminating human intervention wherever possible. It means the data can find its patterns and make autonomous decisions without someone writing new code. Every time IoT sensors receive data, someone needs to categorize it, analyze it, and ensure it feeds back to the device for decision-making. However, managing the considerable data flow will not be accessible if we have many data. For example, self-driving cars must make multiple decisions at the moment, and it is entirely impossible to rely on humans. This is where the role of ML appears in the IoT. The IoT's huge and real potential emerges when combined with AI. ML is used in almost all industrial IoT platforms today. Soon, it will probably not be easy to find an IoT-based device that does not use AI [3, 8].

DL is a subset of ML distinguished from ML by executing learning independently of human supervision. In ML, a human must typically label the data for the computer to understand it. In contrast, DL algorithms can correctly infer

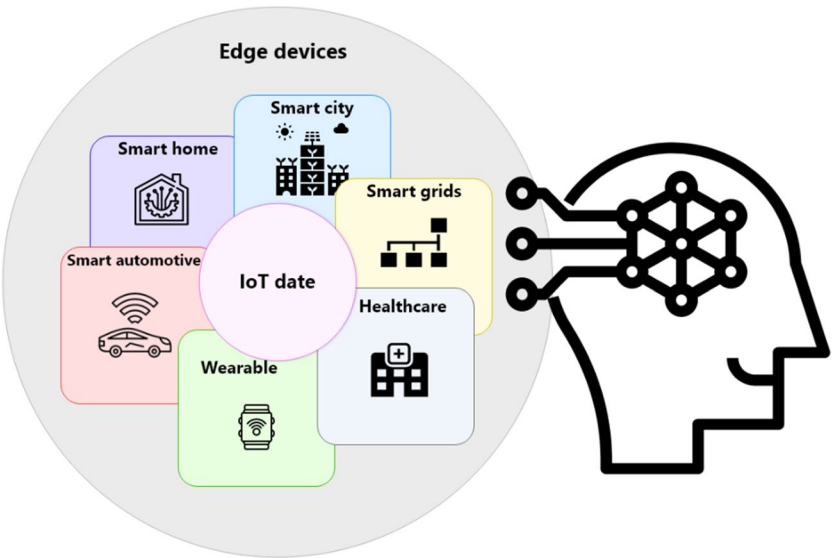



**Fig. 1** DL-based IoT

this data without humans' labeling requirements. DL algorithms can determine how several data collection points relate. These algorithms can automatically assign attributes to each data point rather than needing a human to sort through the data and label each. Due to inadequate hardware and processing power, DL has not been functioning. Modern ML algorithms were developed, examined, and improved over time [2].

DL has made significant advances in AI. It is also evident that integration of AI with IoT has ushered in a new era of intelligence and connectivity. The combination of DL and IoT has paved the way for transformative applications from smart homes to industrial automation. However, to fully and usefully utilize the potential of AI and IoT integration, we need to delve deeper into the domain of edge computing where large-scale language models (LLM) play a central role [13].

We certainly know that deploying LLM on edge devices will bring both challenges and opportunities. Edge devices such as sensors and smartphones are inherently resource-constrained compared to cloud servers. This limitation creates significant challenges in terms of adapting to the computational needs of LLMs due to their size and complexity. However, advances in hardware, along with optimization techniques, are gradually overcoming these obstacles. Integrating LLMs at the edge brings many benefits and furthers the synergy between DL, IoT and AI [13–15]:

- Low Latency: By processing data locally on edge devices, the latency associated with transferring data to centralized servers for analysis is significantly reduced. This is particularly beneficial for applications that require real-time or near-real-time decision making, such as autonomous vehicles and industrial automation systems.
- Privacy and Security: Edge computing ensures that sensitive data remains local and is processed within the confines of the edge device. This reduces privacy and security concerns associated with moving data to the cloud, especially for applications that deal with personal or confidential information.
- Bandwidth Efficiency: Edge computing reduces the load on network bandwidth by reducing the amount of data that must be transferred to centralized servers. This not only reduces bandwidth usage, but also potentially lowers costs, especially in scenarios with limited or expensive connectivity.

### 2.2 Terminology

This section examines some essential words often repeated in DL-based IoT. In this regard, phrases, terms, meanings, and concepts are summarized in Table 1.

## 3 Related work

Several review studies related to our research area were examined and compared in this section to review related studies in the field of DL-based IoT and to find research gaps, shortcomings, and strengths of related review studies. Table 2 compares related studies in aspects of the year of publication, type of publication, main scope, publisher, taxonomy presentation, and future issues.

Bolhasani et al. [2] studied DL regarding IoT applications in healthcare systems. Initially, they put out theoretical notions and theories of technical taxonomy and DL. The analysis of relevant research was then used to illustrate the main DL applications for IoT in healthcare and medical sciences. After discussing each study's basic premise, benefits, drawbacks, and limits, recommendations for more research were made.

Bhattacharya et al. [3] surveyed many state-of-the-art applications of DL to data from smart cities. The article's conclusion included some suggestions for other research areas. Unfortunately, the authors did not discuss any future issues or challenges in this review article. Also, they did not recommend any taxonomy related to their work. Lakshman et al. [4] discussed the basics of the IoT, data generation, and data processing. They also talked about the different DL techniques. They investigated and compiled key reporting initiatives for DL in the IoT region on other datasets. The authors also covered the advantages, uses, and difficulties of DL's approach to enabling IoT applications.

Zikria et al. [5] reviewed the 9 articles in the DL for intelligent IoT issues to showcase the most recent developments in this field of research. They also offered suggestions for future research. However, they must still provide a taxonomy to wrap up their review. It was better that they review more than 9 articles to cover more research gaps. Kant Singh [6] thoroughly analyzed DL models of IoT-based healthcare systems utilizing various case studies. Also, they provided future research directions for DL in IoT-based healthcare systems and reviewed recent trends, contexts, possibilities, problems, and limits. They did not offer any taxonomy.

Saleem et al. [7] briefly introduced the DNN and its many topologies. They provided a comprehensive analysis of DL-driven IoT use cases. Also, a DL-based model for Human Activity Recognition (HAR) was developed in this article. They suggested a taxonomy due to DL-driven IoT. They also discussed future problems in the field of DL-driven IoT.

Chen et al. [9] conducted a thorough survey covering deep reinforcement learning (DRL) algorithms and discussed IoT applications that use DRL. They thoroughly examined the benefits and drawbacks of the cutting-edge

**Table 1** Phrases and terms in the DL-based IoT domain

| Phrases and terms | Meaning and concept |
|---|---|
| Supervised Learning Algorithm | This algorithm trains DL models in which labeled data containing target input and output should be used |
| Unsupervised learning algorithm | This algorithm implements specific techniques on the data, allowing one to recognize patterns and summarize and group the data |
| Semi-supervised learning algorithm | This algorithm is between the previous two modes and is the best choice for model building when data labeling is expensive |
| Reinforcement Learning Algorithm | This algorithm is trained to make decisions based on trial and error and examines how software actions are performed in an environment |
| Accuracy | The percentage of accurate classification predictions to all predictions |
| Precision | The model's ability to categorize positive information is measured by precision |
| Recall | It is also known as sensitivity. The recall is a metric for how "sensitive" a classifier is to finding positive examples |
| F-score | A system's accuracy and recall values are harmonically averaged to provide an F-score |
| ANNs | A model with a minimum of one hidden layer. A neural network with several hidden layers is referred to as a DNN |
| Bayesian neural network (BNN) | A neural network that takes into consideration weight and output unpredictability |
| Recurrent neural network (RNN) | It is a category of ANN where connections between nodes can lead to a cycle where the output of one node can influence the input of another node |
| DNN | A neural network with several hidden layers |
| Convolutional neural network (CNN) | It is a subclass of ANN most frequently used to assess visual data |
| Long Short Term Memory Network (LSTM) | It is an RNN designed for learning from and responding to time-related input, where the intervals between relevant events may be ill-defined or uncertain |
| Data analysis | Gaining knowledge of data by considering samples, measurements, and visualization |
| Dataset | A group of raw data |
| Decision tree | A method for supervised learning that consists of a hierarchy-based collection of criteria and leaves |
| Preprocessing | Processing data before a model is trained |
| Training | The process of choosing a model's optimal input parameters |
| Validation set | The percentage of the dataset used for the initial comparison with a trained model |
| Transfer learning | It is a technique in which a model is pre-trained and adapted to a new task and uses the knowledge gained from a previous task |
| Auto-encoder (AE) | It is a neural network designed for unsupervised learning that learns efficient data representations by encoding and decoding input data |
| Generative Adversarial Network (GAN) | A class of DL models in which two networks consisting of a generator and a detector are simultaneously trained to generate real data |

DRL algorithms after first providing a quick assessment of them. They then discussed using DRL algorithms in a wide range of IoT applications. Furthermore, they outlined future research paths and highlighted upcoming issues that would be crucial to the continued success of DRL in IoT applications.

Thakur et al. [16] explored the integration of DL and the IoT. They highlighted the transformative impact of technology in the twenty-first century. They also comprehensively covered various aspects of the synergy between DL-based IoT, emphasizing its benefits and applications across diverse industries. Unfortunately, they did not provide any taxonomies for classifying their article. Additionally, their survey lacked numerical and analytical charts to present the results they obtained.

Heidari et al. [17] discussed the importance of managing smart cities and societies to optimize resource utilization and enhance quality of life through technologies. They highlighted AI, ML, and DL as promising tools for automating activities in smart cities, presenting research issues and paths for exploration. Seven key emerging developments in IT, including IoT and cloud computing, are categorized, while CNN and LSTM are identified as common ML methods. Their analysis provides insights into current research trends, such as accuracy-focused optimization, offering a comprehensive overview and guiding future investigations.

Amiri et al. [18] explored the significance of ML, particularly DL, in addressing medical and bioinformatics challenges within the context of IoT. Various applications in this domain were discussed by the authors. They emphasized

**Table 2** Related work in the DL-based IoT domain

| Ref | Main scope | Year of publication | Type of publication | Publisher | Taxonomy presentation | Future challenges discussion |
|---|---|---|---|---|---|---|
| Bolhasani et al. [2] | DL applications for IoT in healthcare | 2021 | Systematic review | Elsevier | ✓ | ✓ |
| Bhattacharya et al. [3] | DL for Future Smart Cities | 2020 | Review | Wiley Online Library | ✗ | ✗ |
| Lakshman et al. [4] | DL Techniques for IoT Data | 2022 | Review | MDPI | ✗ | ✓ |
| Zikria et al. [5] | DL for intelligent IoT | 2020 | Unclear | Elsevier | ✗ | ✓ |
| Kant Singh [6] | DL models in IoT-based healthcare | 2021 | Book chapter | Taylor & Francis | ✗ | ✓ |
| Saleem et al. [7] | use-cases of DL for the IoT | 2020 | Unclear | Elsevier | ✓ | ✓ |
| Chen et al. [9] | Deep reinforcement learning for IoT | 2021 | Comprehensive survey | IEEE | ✓ | ✓ |
| Thakur et al. [16] | DL for IoT | 2023 | Review | Springer | ✗ | ✓ |
| Heidari et al. [17] | Applications of ML/DL in smart cities | 2022 | SLR | Elsevier | ✓ | ✓ |
| Amiri et al. [18] | DL applications in IoT-based bio- and medical informatics | 2024 | SLR | Springer | ✓ | ✓ |

the importance of addressing challenges in DL implementation for medical and bioinformatics research. Their findings underscore the evaluation metrics commonly used in assessing DL approaches for accuracy, sensitivity, specificity, and scalability, among others. This comprehensive evaluation aimed to stimulate further research in advancing medical and bioinformatics applications.

## 4 Methodology

This study collects, organizes, and evaluates publications in the DL-based IoT using a systematic literature review (SLR) methodology.

### 4.1 Research question selection

To fulfill the objectives of this study, the following seven Research Questions (RQs) are suggested:

- RQ1: What are the DL models used in the IoT applications? And what are their uses?
- RQ2: What research fields are there in the field of DL-based IoT?
- RQ3: What tools and simulators are used to evaluate the studies in the field of DL-based IoT?
- RQ4: Which datasets are used in the field of DL-based IoT?
- RQ5: What are the common evaluation criteria used in the DL-based IoT domain?
- RQ6: What are the future directions and open issues in DL-based IoT?

When the RQs have been established, the publishers Elsevier, IEEE, Springer, Wiley, and others are selected as article search engines. This preparation includes searching online scientific resources for papers published between 2019 and April 2024.

### 4.2 Article searching process

The following search term was used to locate articles that were related:

("Internet of Things" OR "IoT" OR "Smart City" OR "Healthcare" OR "Smart Home" OR "Smart Grid" OR "Industrial IoT" OR "Smart Agriculture " OR "Smart Transportation") AND ("Deep learning" OR "DL" OR "Deep Neural Network" OR "DNN" OR "Convolutional Neural Network" OR "CNN" OR "Long Short Term Memory Network" OR "LSTM" OR "Recurrent Neural Network" OR "RNN").

### 4.3 Article selection process

In our initial search, we found a total of 261 relevant articles. From the initial set of articles found, we applied a selection process to identify a final set of 56 articles for detailed review. Our selection criteria are designed to ensure we represent the broader landscape of DL-based IoT research. Specifically, the articles were selected based on the following criteria:

- Citation criteria: Highly cited articles published between 2019 and 2022, along with articles published in 2023 and April 2024, regardless of the number of citations, were selected.

- Relevance: Articles that had keywords related to DL-based IoT in their titles as well as articles published by reputable publishers were considered.
- Publication types: Both journal articles and conference papers were included to ensure comprehensive coverage of the literature.

Articles were excluded based on the following criteria:

- Articles published before 2019
- Articles in languages other than English
- Articles with less than 8 pages
- Articles that do not have keywords related to DL-based IoT in their titles

In Fig. 2 we illurestrate our criteria for selecting final 56 papers for analysing.

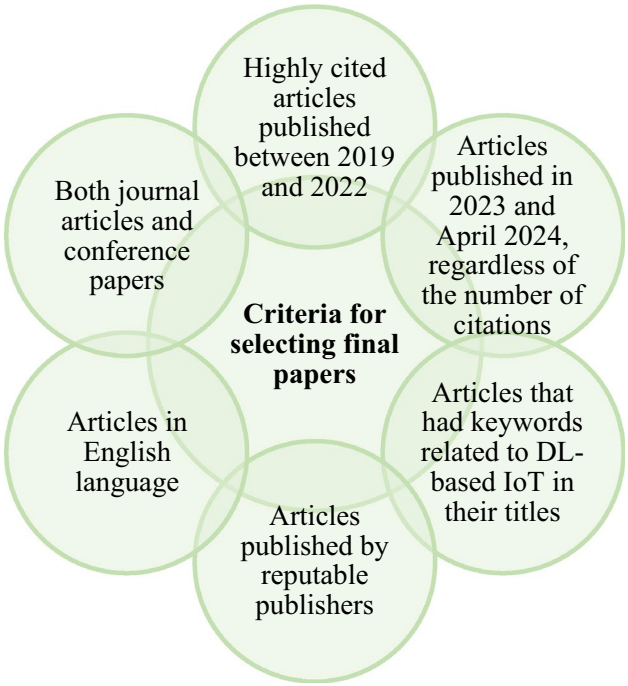



**Fig. 2** Criteria for selecting final papers

## 4.4 Data extraction and evaluation

Each of the selected publications underwent thorough evaluation to ensure their relevance and quality to the DL-based IoT topic. This process resulted in the identification of 56 high-quality articles for detailed analysis.

The selection of 56 articles for detailed review was driven by a clear rationale aimed at representing the breadth and depth of DL-based IoT research during the specified timeframe. These articles were chosen based on their citation impact, relevance to the research questions, and publication in reputable venues. By employing a systematic approach to article selection, we aimed to ensure the comprehensive coverage of DL-based IoT literature while maintaining high standards of quality and credibility. This selection process aligns with the study's objective of providing a holistic understanding of the current state and future directions of DL-based IoT research.

Figure 3 shows the selection of publications by year and publisher in the DL-based IoT fields after 56 articles were chosen using inclusion/exclusion criteria. Figure 4 shows The distribution of the selected articles by the publisher in the DL-based IoT fields. Figure 5 shows the distribution of the selected articles by the year of publication in the DL-based IoT fields.

Figure 6 briefly shows the steps of selecting 56 research articles reviewed in this article.

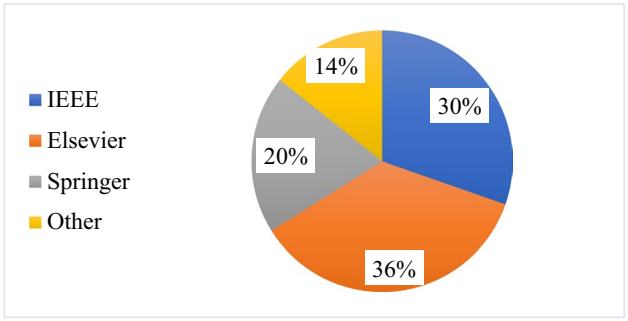
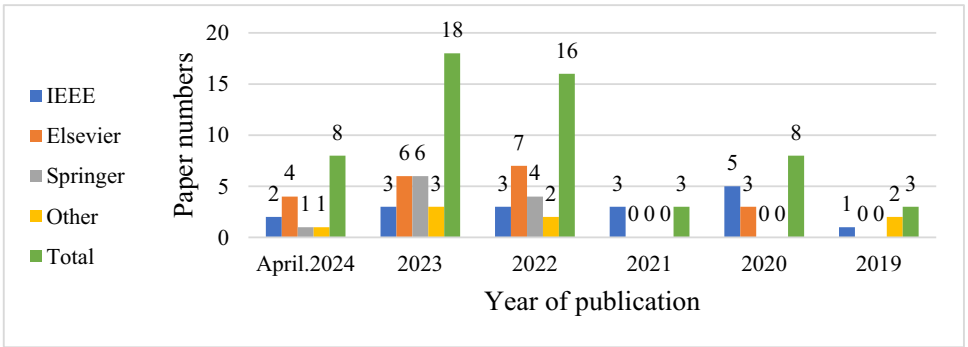



**Fig. 4** The distribution of the selected articles by the publishers

**Fig. 3** Distribution of research papers by publisher



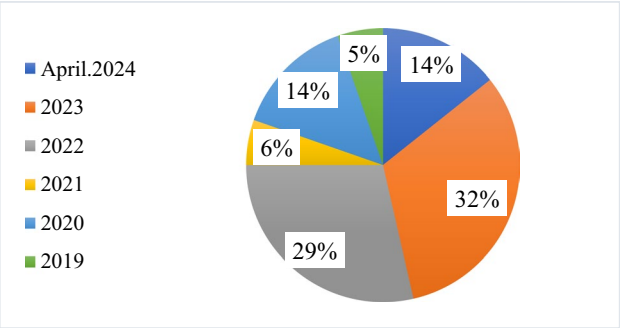



**Fig. 5** The distribution of the selected articles by the year of publication

According to the research findings, also considering the clarification of the domains of research in the DL-based IoT, which domains and areas they cover, Fig. 4 shows the proposed taxonomy of DL-based IoT, including security, edge and fog computing, energy management, healthcare, and smart city.

By observing Fig. 7, we can understand that in the leading group of security, there are subgroups such as attacks and threats, intrusion detection, and cyber security, which themselves include subgroups such as physical-based attacks, network-based attacks, endpoint attacks, software-based attacks, distributed denial of service, denial of service, malware attack, botnet attack, worm, virus, architecture-based attacks, authentication, authorization, adware and spyware, cyber-attacks, phishing attack, access control, password attacks, internal network risks, employee-generated risks, social engineering attacks, third party threats, and vulnerabilities.

On the other hand, in the other sub-group of applications, it can be mentioned that the IoT includes branches such as smart homes, smart wearables, smart grids, smart agriculture, healthcare and medical systems, and smart cities, which can use DL to solve their challenges. To separate the applications, we can refer to smart city subgroups, which include smart transportation, smart parking, traffic congestion management, obstacle detection, accident detection, smart infrastructure, and urban modeling. Healthcare and medical system subgroups include medical images, image segmentation, drug development and discovery, drug repurposing, antimicrobial drug identification, biomarker discovery, genome, disease prediction, health monitoring, and health records. The smart home branch includes smartphone sensors, energy demand forecasting, and human activity recognition. Smart wearables include smartphone sensors and smartwatch sensors. The extension of the smart grid provides power demand forecasting, electricity price forecasting, anomaly detection, and smart agriculture, including plant classification and plant disease prediction.

## 5 Organization

According to the 56 selected articles in the methodology section and our proposed taxonomy in Fig. 4, we reviewed the research articles in DL-based IoT. We put a summary of them in the subsections of this section. We divided the articles into five subgroups: security, edge and fog computing, energy management, healthcare, and smart city, which are the five main branches of our taxonomy, to make it easier to review and read this section and to organize articles with related contexts and in one direction in one category for researchers.

### 5.1 Security

In this subgroup, we categorize articles in the field of IoT security that have utilized DL to address their challenges. Table 3 summarizes these articles and compares them according to their challenge, domain, used dataset, DL method, evaluation environment, and obtained result.

Thamilarasu et al. [19] created an intelligent intrusion-detection system specifically designed for the IoT. To

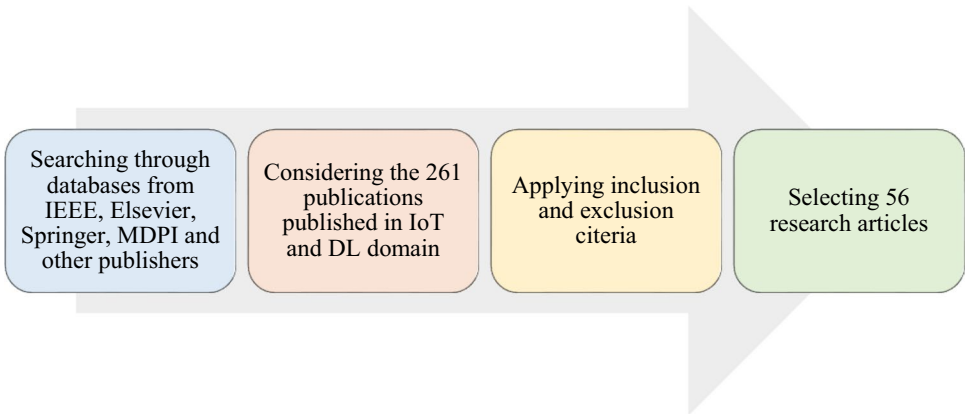



**Fig. 6** Steps of selecting research articles

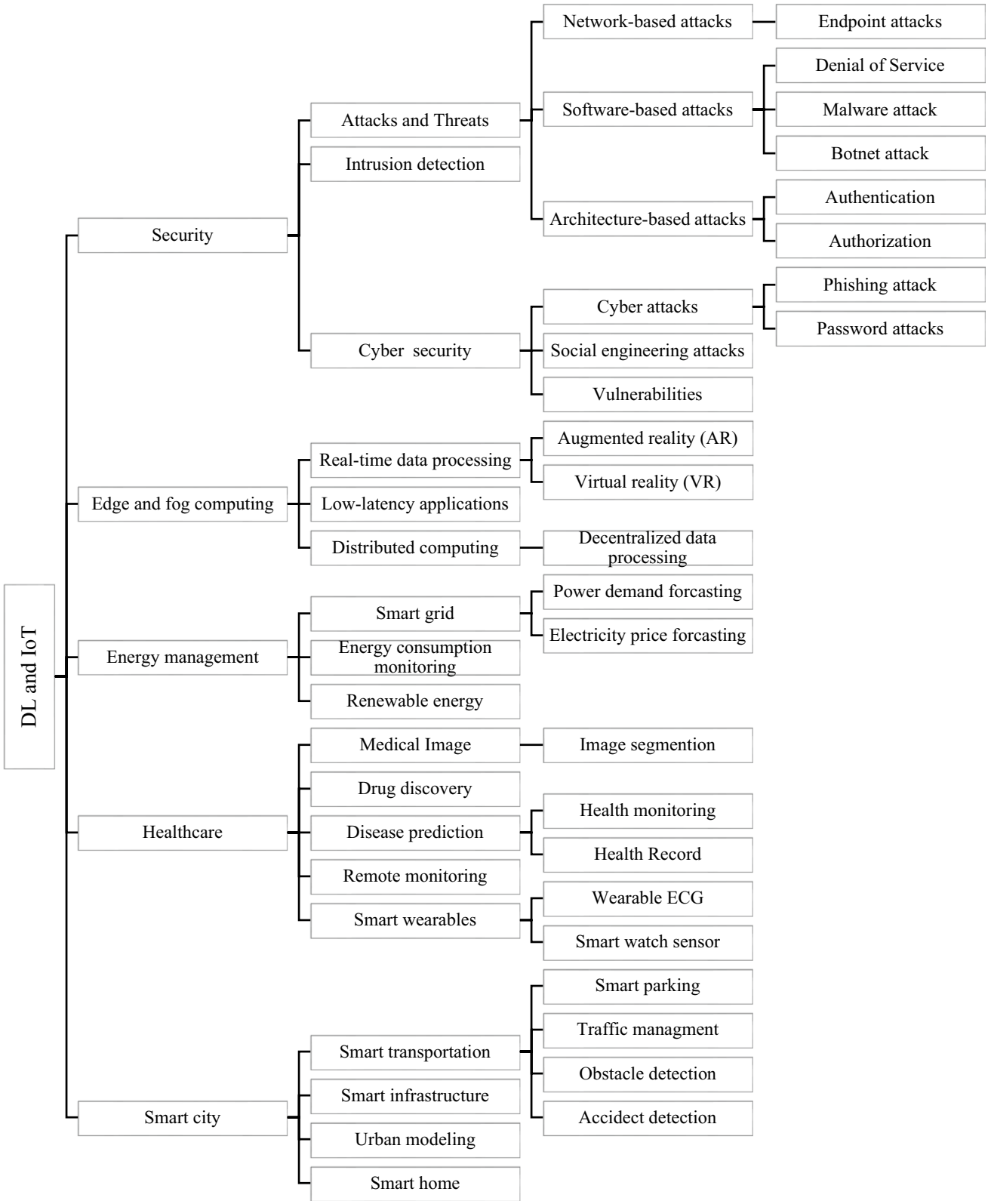



**Fig. 7** Taxonomy of DL-based IoT

develop adequate safety and defense, researchers considered the difficulty of growing IoT cyberattacks and spotting assaults on IoT systems in real time. They employed a DL system to identify fraudulent traffic in IoT networks. The detection solution supports compatibility between multiple protocols for networking used in IoT and offers security as a service. They tested the feasibility of the suggested detection system using simulation and real-network traces to demonstrate its adaptability. The findings showed the effectiveness of the proposed system in identifying actual invasions. Attack detection in their suggested technique could be more effective. It would be preferable to broaden their suggested IDS to recognize other IoT threats, such as position assaults.

Otoum et al. [20] presented a system for intrusion detection using DL to secure IoT. The system's most crucial components suggested in this work were data set management and appropriate learning features, which significantly impact the precision of attack detection. The proposed component combined the spider monkey optimization (SMO) method with the stacked-deep polynomial network (SDPN) to accomplish excellent detection identification. According to extensive investigation, the suggested approach performed better in accuracy, precision, recall, and F-score. They simulated their approach using Python Tensorflow. The authors want to test their proposed model in the future using classification techniques, such as Naive Bayes and Random Forest.

Albishari et al. [21] presented a unique DL-based early-stage detection method utilizing the IoT routing attack dataset. The IoT routing attack dataset was used to demonstrate their method and improve the capacity to identify routing assaults. The effectiveness of the proposed model's training was evaluated using binary classification techniques. The comparative findings showed that the suggested classifier has the best runtime rate and high training accuracy. The authors achieved good accuracy, precision, recall, and F-score.

Houda et al. [22] provided a revolutionary AI-powered framework to make it possible to analyze essential decisions made by ML/DL to identify intrusions and assaults in IoT networks. To identify and anticipate IoT assaults in real time, the researchers initially created an ML/DL-based intrusion detection system (IDS). Next, they build various AI models on top of our DNN framework to give cyber security professionals additional confidence in, access to, and clarify the ML/DL-based IDS's conclusions. The detailed results indicated the effectiveness and justification of the suggested approach.

Mohamed Ishaque et al. [23] created an IoT-enabled waste detection and categorization system named IoT-WDC-ODL. The IoTWDC-ODL approach enables IoT devices to gather data and transport it to the cloud for additional processing. In other words, the authors created an IoT-based garbage detection and categorization system utilizing the best DL approach. Also, the MobileNetv2 and Inception v2 models employ the single-shot detector for the item detection procedure. Moreover, the firefly approach is used to tune the hyperparameters of DL models. Lastly, the proper class labels for the rubbish items were chosen using the Gaussian Naive Bayes classifier. A wide range of simulations were performed on the benchmark dataset, and the outcomes were examined using several metrics. The results of the experiments demonstrated how the IoTWDC-ODL technology outperformed more contemporary methods.

Qazi et al. [24] presented an intelligent and effective DL-based network intrusion detection system (NIDS). This study proposed a non-symmetric deep AE to solve network intrusion detection challenges and showed its specific functions and results. DL-based approach, which uses the TensorFlow library and GPU framework, had a reasonable accuracy rate. The recommended method may be used to network protection research areas and DL-based identification and categorization systems.

Nasir et al. [25] developed a model to identify IoT traffic intrusions named DF-IDS with two phases. The first stage utilized SpiderMonkey to compare the top features from the feature matrix. The second phase involved training a DNN for intrusion detection using these characteristics and the provided labels. Compared to previous comparator models and earlier research, it improved accuracy and F1 score.

Sriram et al. [26] presented a botnet detection system that used network traffic flows and was DL-based. To identify assaults coming from infected IoT devices, network traffic flows were captured by the botnet detection framework, which then converted them into connection records and applied a DL model. The characteristics employed by the researchers were taken from network traffic patterns. Instead, it is possible to separate between regular activity and botnet assault activity by analyzing the payload data. Moreover, the CNN software may be utilized for multivariate analysis and learning to enhance the performance of current tasks. The network flow can also be displayed as graphics. The suggested DL model performed better than the traditional ML models.

Saharkhizan et al. [27] provided a method for detecting cyberattacks on IoT systems using cutting-edge DL. Their method specifically incorporated several LSTM modules into a group of detectors. They used a real-world dataset of Modbus network traffic to test the efficacy of the suggested technique, and they found that the approach was accurate in spotting cyberattacks on IoT devices. To offer a more obvious DL model to identify IoT cyberattacks, they should look at the explainability of LSTM models. Implement the suggested strategy for various IoT networks as well.

Ullah et al. [28] suggested a hybrid DL model for IoV cyber threat detection to help lessen unwanted attacks on

**Table 3** Security category research articles

| Ref | Challenge | Domain | Used DL method | Evaluation environment | Obtained result |
|---|---|---|---|---|---|
| Thamilarasu et al. [19] | Intrusion detection | IoT | DBN | Implementation using Keras and simulation using the Cooja network simulator and IoT network-traffic dataset | Scalability<br>Detection of real-world intrusions effectively |
| Otoum et al. [20] | Intrusion detection | IoT | SMO | Simulation using Python-Tensorflow and NSL-KDD dataset | Good accuracy, precision, recall, and F-score |
| Albishari et al. [21] | Routing attacks | IoT network | K-nearest neighbors, Naïve Bayes | Simulation | Good accuracy, precision, recall, and F-score |
| Houda et al. [22] | Intrusion detection | IoT | DNN | Implementation | Efficiency |
| Mohamed Ishaque et al. [23] | Waste detection | IoT | DBN | Simulation using Benchmark dataset | Good performance related to proposed models |
| Qazi et al. [24] | Intrusion detection | IoT | DBN | Implementation using TensorFlow using Benchmark dataset | Good accuracy |
| Nasir et al. [25] | Intrusion detection | Edge IoT | Spider Monkey | Implementation | Good accuracy and F-score |
| Sriram et al. [26] | Botnet attack | IoT | CNN | Implementation using TensorFlow3 with Keras using N-BaIoT | Good performance |
| Saharkhizan et al. [27] | Cyber attacks | IoT | LSTM | Simulation using Tensorflow and Modbus network traffic dataset | Effectiveness in comparison to other models |
| Ullah et al. [28] | Intrusion detection | Internet of Vehicles | LSTM | Simulation using Python 3.0 environment and Car-hacking and DDoS datasets | Good accuracy, precision, recall, and F-score |
| Abusitta et al. [29] | Anomaly detection | IoT | SVM, Logistic Regression | Implementation using Real-life IoT datasets | Effectiveness of the proposed framework in terms of enhancing the accuracy of detecting malicious data |
| Dina et al. [30] | Intrusion detection | IoT | CNN | Implementation using Bot-IoT, WUSTL-IIoT, and WUSTL-EHMS | Effectiveness of the proposed approach |
| Kumar et al. [31] | Security and privacy | Industrial IoT | CNN | Implementation using ToN-IoT and IoT-Botnet datasets | High efficiency and scalability |
| Naeem et al. [32] | Malware detection | Industrial IoT | CNN | Simulation using Python Tensorflow and Malware Fingerprints dataset | High predictive time and detection accuracy |
| Yazdinejad et al. [33] | Cyber threat detection in IIoT | IIoT | LSTM and AE | Implementation using GP and SWaT | Improved accuracy and efficiency |
| Wang et al. [34] | Privacy concerns and network intrusion in IoT | IoT | DNN | Implementation using Python and IoT-Botnet 2020 dataset | Improved Accuracy, reduced False Alarm Rate |
| Nazir et al. [35] | Detection of IoT threats | Cybersecurity | Hybrid CNN-LSTM | Implementation using IoT-23, N-BaIoT, CICIDS2017 datasets | Achieved heigh accuracy N-BaIoT and CICIDS2017 datasets |

**Table 3** (continued)

| Ref | Challenge | Domain | Used DL method | Evaluation environment | Obtained result |
|---|---|---|---|---|---|
| Sharma et al. [36] | Intrusion detection in IoT networks | Cybersecurity | DNN, and CNN | Implementation using NSL-KDD, UNSW-NB15 datasets | DL models outperformed traditional methods, achieving better accuracy. Utilized LIME and SHAP for explainability |
| Lilhore et al. [37] | Intrusion detection in IoT and 5G | Cybersecurity | MobileNetV3-SVM, Transfer learning | Implementation using CIC-IDS-2017, 2018, UNSW-NB15 datasets | Showed improvements in accuracy, precision, false positive rates, MCC, and AUC-ROC over existing approaches |
| Singh et al. [38] | Botnet-based IoT network traffic analysis | Cybersecurity | DL-based methods | Implementation using IoT traffic datasets | Conducted a comparative study on DL-based methods for IoT traffic analysis |

vehicular communications and identify accidents involving smart vehicles. The proposed model was based on recurrent gated units and LSTM. Precision, recall, and F1-score, together with the other performance metrics, attest to the suggested framework's enhanced functionality. They used a simulation method using a Python 3.0 environment to evaluate their idea.

Abusitta et al. [29] developed a DL-powered anomaly detection system for the IoT that can learn and collect reliable and practical characteristics that are not severely impacted by unstable situations. The classifier then used these traits to improve the precision with which it could identify fraudulent IoT data. Compared to the most advanced IoT-based anomaly detection methods, experimental findings based on real-world IoT datasets demonstrated the suggested framework's usefulness in improving the accuracy of detecting harmful data. They ought to strengthen and improve the proposed model's robustness so that it can recognize abnormalities.

Dina et al. [30] employed the customized loss function known as a focused loss to proactively down-weight simple instances and concentrate on the hard instances by constantly enabling scaled-gradient changes for training efficient ML models. They conducted rigorous experimental evaluations using three datasets from various IoT domains and contrasted our suggested method with cutting-edge intrusion detection algorithms. They discovered that the proposed method performed better in accuracy, precision, F1, and MCC scores.

Kumar et al. [31] introduced a new privacy-preserved threat intelligence framework (P2TIF) to safeguard private data and find cyber threats in IIoT contexts. The P2TIF framework that is presented consists of two key components. The first is a flexible blockchain module that allows secure IIoT data transfer and guards against data contamination threats. The second is a DL module that re-formats real data and shields it from inference assaults. According to security analysis and experimental findings, the suggested P2TIF architecture has excellent efficiency and scalability.

Naeem et al. [32] Provided a system that can recognize malware assaults on the Industrial IoT (MD-IIOT). An approach based on deep CNN and color picture display was suggested for thorough malware analysis. Their malware visualization method first transforms APK files into color pictures. A DCNN model is then chosen to capture the dynamic visual characteristics of the virus. Malware attack detection and malware image training are done using the DCNN model. Lastly, the strategy suggested for complex data sets is compared to the reliability of state-of-the-art detection techniques. The results of the proposed method are contrasted with those of earlier malware

detection techniques. Limited memory is the disadvantage of this work.

Yazdinejad et al. [33] presented a DL model to identify cyber threats in the IIoT environment, which addresses the challenges posed by large-scale, dynamic, and heterogeneous data generated by interconnected devices. They combined the strengths of LSTM and AE architectures to identify anomalous activities. They utilized LSTM to capture normal data patterns from past and present data while employing AE to identify essential features and reduce data dimensionality. Their evaluation was conducted on two real IIoT datasets, including Gas Pipeline (GP) and Secure Water Treatment (SWaT), which exhibit imbalances and long/short-term dependencies. The results indicated the performance of the proposed ensemble model, achieving high accuracy.

Wang et al. [34] addressed privacy concerns arising from zero-day hacks in the context of IoT devices transmitting sensitive information over the regular internet. They proposed a novel approach combining DNN, FL, and mutual information to enhance security for effective anomaly detection in IoT networks. The authors utilized decentralized on-device data to identify intrusions within the IoT network. Their evaluation was performed using the IoT-Botnet 2020 dataset. Results highlighted the efficiency of the DNN-based Network, improved accuracy, and a reduced False Alarm rate compared to other DL models.

Nazir et al. [35] introduced a Hybrid architecture using CNN and LSTM for IoT threat detection. They aimed to achieving high accuracy rates on various datasets such as IoT-23, N-BaIoT, and CICIDS2017. Their trengths include robust threat detection capabilities and adaptability to diverse IoT environments. Future work involves optimizing data processing with Principal Component Analysis (PCA) and deploying advanced techniques like model quantization and pruning.

Addressing IoT network security, Sharma et al. [36] proposed DL model for intrusion detection, achieving improved accuracy compared to traditional methods. Utilizing LIME and SHAP for explainability, the authors enhanced model understanding. Future research may explore further interpretability techniques. They used NSL-KDD, UNSW-NB15 datasets and DNN, CNN methods to provide their approach and evaluate it.

Lilhore et al. [37] introduced a hybrid model for intrusion detection in IoT and 5G networks, this study utilizes MobileNetV3-SVM and transfer learning to effectively analyze network activity patterns. The approach leverages the advantages of a multi-layered structure to distinguish between authentic and malicious behavior efficiently. The proposed framework is evaluated on various datasets (CICIDS-2017, 2018, UNSW-NB15), demonstrating improvements in accuracy, precision, and false positive rates over existing approaches.

Singh et al. [38] focused on botnet-based DDoS attacks in IoT networks. They conducted a comparative study on DL-based methods for IoT traffic analysis, identifying current challenges and future research directions. Future work could involve addressing scalability issues and enhancing model robustness. They used IoT traffic datasets.

### 5.2 Edge and fog computing

The integration of edge/fog computing and DL represents a development in the IoT. In this paradigm, data processing is decentralized, bringing computing tasks closer to data sources through edge/fog computing architectures. DL techniques enhance these decentralized systems' capabilities by enabling intelligent decision-making at the edge. Edge computing involves processing data close to the source, reducing latency, and increasing real-time responsiveness. This subgroup categorizes articles in edge and fog computing that have utilized DL to address their challenges. Table 4 summarizes these articles and compares them according to their challenge, domain, used dataset, DL method, evaluation environment, and obtained result.

Zhao et al. [39] examined mobile edge computing networks for the intelligent IoT, where many users receive assistance from several computational access points for portions of their computational activities (CAPs). In this approach, a neural network was taught to anticipate the offloading action using training data collected by the environmental system. A Deep Q-Network was utilized to extract knowledge of the offloading choice to maximize the system efficiency. They created the system using the deep reinforcement learning algorithm to propose the offloading approach intelligently. The suggested deep reinforcement learning-based technique may minimize the system cost of latency and energy consumption.

Wang et al. [40] increased the high-level dance moves' design impact and aided dancers in better mastering these motions. The body changes with sophisticated dance motions based on the DL algorithm and IoT technology were examined to develop the practical use of biological image visualization technology. DL was initially used for picture identification, and the technical architecture of the IoT technology was built. Eventually, the network structure of the dance generation model was established using DL-based IoT-edge computing.

Chen et al. [41] addressed the importance of managing Online Public Sentiment (OPS) during public emergencies in the context of the rapid development of the Internet. They aimed to establish a safe and credible online environment, leveraging the IoT-native big data. They focused on the evolution of OPS during public emergencies and proposed an approach using the LSTM Neural Network model. Results indicated that the Adam-optimized LSTM NN model

**Table 4** Edge and fog computing category research articles

| Ref | Challenge | Domain | Used DL method | Evaluation environment | Obtained result |
|---|---|---|---|---|---|
| Zhao et al. [39] | Mobile edge computing | IoT | Deep Q-Network | Simulation | Good latency and energy consumption |
| Wang et al. [40] | High-level dancing moves assist dancers in mastering these moves | IoT | DRL | Implementation | Good performance |
| Chen et al. [41] | Managing Online Public Sentiment in Emergencies | Social Media | LSTM | Simulation | High prediction accuracy |
| Belmonte-Fernández et al. [42] | Addressing IoT challenges with fog/edge computing | IoT, Fog/Edge | Not specified | Implementation | Improved responsiveness, resilience, and elasticity |
| Lv et al. [43] | Enhancing storage security in edge-fog-cloud computing | Edge computing | Not specified | Implementation | Improved storage security, cost-effective, high availability |
| Chen et al. [44] | Effective IDS for dynamic IoT environment | Fog Computing | CNN | Implementation using AWID, CIC-IDS2107 datasets | Improved detection performance and robustness, low-latency and high-accuracy intrusion detection for IoT |
| Peláez-Rodríguez et al. [45] | Prediction of low-visibility events due to fog | Fog/Edge | RNN, LSTM GRU | Implementation using real-world datasets | Good performance |
| Xu et al. [46] | Green communications for IIoT | Industrial IoT | Deep neural networks | Simulation | Proposed algorithms improved system performance without exhaustive search |
| Wang et al. [47] | Body condition scoring of dairy cows | Agriculture | YOLOv7, EfficientID, EfficientBCS | Implementation | Reached heigh accuracy |

achieves high prediction accuracy in forecasting the hotness of OPS during dynamic evolution.

Belmonte-Fernández et al. [42] proposed an architecture that leveraged the collaboration of fog/edge computing and cloud computing in developing DL solutions for IoT. In their approach, time-consuming and hardware-intensive DL models were constructed in the cloud using data from the fog/edge. Subsequently, these models were deployed back to the fog/edge for practical use. Their architecture was designed based on the principles of reactive systems, emphasizing responsiveness, resilience, and elasticity through asynchronous message passing between components.

Lv et al. [43] focused on enhancing the storage security of edge-fog-cloud computing to improve cloud storage security. The researchers utilized data from an intelligent manufacturing industrial machine, which undergoes pre-processing. The researchers integrated digital twins technology with the perception data of machine manufacturing to construct a digital twin, simulating the online data-driven behavior of the machine manufacturing process. The digital twins model was developed and saved using 3Dmax. Experimental results highlighted the effectiveness of digital twins technology, and the proposed two-layer cloud database model demonstrates cost efficiency compared to other models.

To overcome the challenge of time-consuming parameter tuning in DL-based IDSs, Chen et al. [44] proposed a multi-objective evolutionary CNN (MECNN) designed to run on fog nodes within Fog computing on IoT. The proposed MECNN employed a CNN classifier for intrusion detection and a modified multi-objective evolutionary algorithm based on decomposition (MOEA/D) to evolve the CNN model. Experimental studies on AWID and CIC-IDS2107 datasets demonstrated that the MECNN model improves detection performance and robustness, providing enhanced protection for IoT compared to other state-of-the-art IDSs.

Peláez-Rodríguez et al. [45] addressed the challenge of predicting low-visibility events due to fog using various DL-based ensemble algorithms. They considered seven different DL architectures, each generating multiple individual learners. Their models underwent a random selection of hyperparameters, covering aspects of data preprocessing, model architecture, and training procedures within predefined discrete ranges. The performance of the proposed methodology

was tested in predicting low-visibility events caused by orographical and radiation fog in the north of Spain.

Xu et al. [46] focused on green communications for IIoT, this research introduces algorithms to improve energy efficiency using RIS and UAVs. Future research may involve practical testing and optimization for diverse IIoT environments. They simulated their idea using DNN method. Their approach employs deep neural networks to jointly optimize transmit power, channel allocation parameters, and RIS reflection coefficients, improving system performance while ensuring quality of service for IIoT scenarios.

Wang et al. [47] addressed dairy cow health, the authors introduced an intelligent Edge-IoT platform for estimating Body Condition Score (BCS) using DL. Their approach integrates edge computing with DL models for cow detection, identification, and BCS estimation, achieving high accuracy rates within low latency. Future work may involve field testing and scalability assessment. YOLOv7, EfficientID, EfficientBCS methods used to achieve their approach.

### 5.3 Energy management

Energy management in the context of DL-based IoT involves the integration of smart technologies to optimize energy consumption, increase efficiency, and contribute to sustainable practices. IoT devices, often with limited resources, benefit from intelligent energy management systems that use DL algorithms for data analysis and decision-making. DL models can be used to analyze complex patterns in energy consumption data, predict consumption trends, and optimize operating parameters to minimize energy waste. This subgroup categorizes articles in the energy field that have utilized DL to address their challenges. Table 5 summarizes these articles and compares them according to their challenge, domain, used dataset, DL method, evaluation environment, and obtained result.

Xin et al. [48] described a DL architecture for managing electricity (DLA-PM). It estimates future electricity demand quickly and enables positive collaboration among power distributors and consumers. In the suggested approach, mobile devices are linked to a global IoT cloud server connected to smart grids to maintain uninterrupted power supply and usage. It employs various preprocessing techniques to handle multiple electrical data, uses a successful short-prediction decision-making procedure, and executes it utilizing resources.

Teng et al. [49] suggested a DL structure for building innovative symmetric presentations of asymmetric datasets. It is included in a model built specifically for DL-based attack detection in SGCS environments. The proposed attack detection model makes use of DNN and Decision Tree classifiers. The cyber-attack has been identified and countered in the suggested strategy, which employs ensemble DL methods specially adapted for SGCS. The results showed that the recommended strategy outperforms established methods like Random Forest and Support Vector Machine. The suggested approach was more effective than traditional schemes.

Li et al. [50] presented the SI-LFC (swarm intelligence load frequency control) technique. The units in each region were regarded as autonomous actors in their suggested methodology. A centralized offline learning policy based on swarm intelligence was also implemented to balance interests among many operators. They also presented an evolutionary multiagent deep meta-actor–critic (EMA-DMAC) method that combined evolutionary learning and meta-reinforcement learning to facilitate quick collaborative

**Table 5** Energy management category research articles

| Ref | Challenge | Domain | Used DL method | Evaluation environment | Obtained result |
|---|---|---|---|---|---|
| Xin et al. [48] | Power management | Smart city | CNN | Simulation | High accuracy and low error in predicting power consumption |
| Teng et al. [49] | Cyber-attacks | Smart grid | DNN and Decision Tree classifiers | Implementation using smart grid control systems (SGCSs) dataset | The suggested approach is more effective than traditional schemes |
| Li et al. [50] | Efficient load frequency control in multiarea microgrids | Energy, microgrids | Swarm Intelligence | Simulation of four-area LFC model for Sansha Island in CSG | Good performance |
| Tomazzoli et al. [51] | Scalable and autonomous energy management | IoT, energy management | ML | Implementation | Autonomous energy efficiency |
| Puri et al. [52] | Generating electrical energy from IoT-based renewable sources | IoT, renewable energy | ANN, ANFIS | Implementation | Predicting total power generated from renewable sources |

learning of swarm agents. Simulation was used to show the efficiency of the suggested strategy.

Tomazzoli et al. [51] addressed the increasing demand for scalable and autonomous management systems in smart environments, focusing on energy efficiency for home and business applications. They proposed a system architecture aimed to achieve autonomy in energy efficiency, allowing for optimal energy configuration without human intervention, even in the face of topology changes. The application scenarios include smart industries, where energy managers can monitor and optimize numerous sparse divisions effortlessly, and smart homes, where an autonomous system can help impaired individuals avoid energy waste.

Puri et al. [52] addressed the global energy consumption challenge by focusing on generating electrical energy from multiple renewable energy sources using an IoT-based system. Two different AI models, namely ANN and Adaptive Network-based Fuzzy Inference System (ANFIS), were employed to predict the total power generated from renewable energy resources. The models were evaluated using statistical parameters such as Root Mean Square Error (RMSE) and the R2 correlation coefficient. The results indicated that the ANN model outperforms the ANFIS model in terms of predictive accuracy and performance, showcasing the potential of AI techniques in optimizing electrical energy generation from diverse renewable sources.

### 5.4 Healthcare

DL contributes to significant advancements in diagnostics, treatment planning, and patient care in healthcare. DL models can analyze medical images, such as X-rays and MRIs, for more accurate and timely diagnoses. Moreover, they can assist in predicting disease outcomes and recommending personalized treatment plans based on patient data. This subgroup categorizes articles in healthcare and IoMT that have utilized DL to address their challenges. Table 6 summarizes these articles and compares them according to their challenge, domain, used dataset, DL method, evaluation environment, and obtained result.

Raj et al. [53] created a feature selection model for medical picture categorization. A method called Opposition-based Crow Search (OCS) was suggested to improve the performance of the DL classifier. After the OCS algorithm identified the best features from properly established photos, the gray level was chosen for study in the proposed multi-texture feature model. The suggested outcomes were put into practice in MATLAB. Future studies will employ segmentation systems, and specific autonomous classification approaches to locate the tumor portion in clinical pictures. A feature reduction approach should also be taken into account.

Zhou et al. [54] suggested a human activity detection method based on an enhanced Bayesian Complexity Network. In other words, limited IoT systems require a top-notch network for reliable analysis of patient information due to their restricted computational capability. In light of this, the authors of this work developed a novel human activity detection that enables any expert system to receive data via conventional radio frequency communications or cloud-assisted low-power scattering communications. Moreover, the Bayesian network uses cutting-edge DL architecture and an efficient evacuation technique to address security challenges. According to experimental findings, IoT wearable sensor data was susceptible to various observational and epistemic sources of uncertainty, including noise and dependability.

Vaiyapuri et al. [55] provided a system for intelligent home healthcare that uses an optimized deep CNN (IMEFD-ODCNN) to identify geriatric falls—the suggested design aimed to provide clever DL techniques for smartphones and recognition. The input video recorded by IoT systems is principally normalized in the proposed model. After that, the most appropriate feature vectors for fall detection were selected. The classification of fall and non-fall occurrences was completed using the Sparrow Search Optimization Algorithm (SSOA) with variable AE. The proposed model needs improvement in scalability.

Guan et al. [56] suggested a genetic algorithm (GA) with LSTM based on IoT and smart wearables. Notably, the LSTM compensates for GA's subpar local search ability. The network model was trained and simulated using LSTM-GA in this study.

Sahu et al. [57] developed a DL-based continuous authentication solution. The recommended security mechanism continually confirmed that authorized users are present during a session. The system uses a DL-based LSTM categorization algorithm to accept data from users and authenticate them. They Implemented their approach using TensorFlow in the Google Colab environment.

Sheela A et al. [58] proposed a method to recognize time series data that combines learned or surface-level properties using DL. The suggested method addresses the common issues that arise in DL models when network computing is required. The hybrid system structure in this proposed model solves the drawbacks of the conventional DL paradigm when computational nodes are included. The suggested technique for real-time computing in the IoT was accomplished with spectral domain preparation before the data was transmitted to the neural network architecture. Studies showed that the proposed strategy outperformed the methods that were already used.

Qiu et al. [59] improved transformation-based compression standards like JPEG by sending substantially less data than one picture at the sender's end. To recover the original

**Table 6** Healthcare category research articles

| Ref | Challenge | Domain | Used DL method | Evaluation environment | Obtained result |
|---|---|---|---|---|---|
| Raj et al. [53] | Prediction and early diagnosis of critical diseases | Internet of Medical Things | DBN | Implementation using MATLAB and Brain, lung, breast cancer, and Alzheimer's disease datasets | Improving sensitivity, validity, and accuracy in the detection of medical pictures |
| Zhou et al. [54] | Human Activity Recognition | Wearable IoT Device | Bayesian Convolution | Implementation using two public datasets | Effectiveness |
| Vaiyapuri et al. [55] | Elderly Fall Detection | smart Homecare | CNN | Simulation and implementation using UR fall detection dataset | Improving the accuracy |
| Guan et al. [56] | Football in Colleges' Effectiveness | College football training | LSTM | Implementation and simulation | Making college football training more effective |
| Sahu et al. [57] | Security threats | IoT-enabled healthcare service | LSTM | Implementation using TensorFlow in Google Colab environment | Improving the security |
| Sheela A et al. [58] | Resource limitations | Wearable sensors | CNN | Implementation using Human activity datasets | Good performance in comparison to other models |
| Qiu et al. [59] | JPEG Compression | IoT | CNN | Implementation | Good image quality |
| Khanna et al. [60] | Cardiovascular disease detection | Healthcare | LSTM, DNN | Implementation using Biomedical ECG Signals | Improved accuracy |
| Obayya et al. [61] | Skin cancer detection with DL-based IoT | Medical imaging, IoT | CNN | Simulation using Python | Outperforms other methods in accurately identifying skin lesions |
| Wang et al. [62] | Efficient music retrieval and user experience | Music information retrieval | Naive Bayes Classifier | Simulation | Highest accuracy in music emotion classification |
| Kumar et al. [63] | Security in IoT-enabled healthcare systems | Healthcare | AE, LSTM | Implementation using CICIDS-2017, ToN-IoT datasets | High accuracy |
| Srivastava et al. [64] | Data analysis in IoMT for healthcare | Healthcare, IoMT | Hybrid DL | Implementation using PhysioNet biosignal repository | High accuracy rates |

data at the receiver's end, they suggest a two-step strategy that combines the most recent signal processing-based recovery technique with a deep residual learning model. The proposed approach may increase the effectiveness of multimodal big data transfer in the IoT.

Khanna et al. [60] introduced a DL-based IoT-enabled healthcare disease diagnosis (IoTDL-HDD) model for the detection of cardiovascular diseases (CVDs) using biomedical electrocardiogram (ECG) signals. Their proposed IoTDL-HDD model utilized DL models for CVD detection, employing a Bidirectional LSTM (BiLSTM) feature extraction technique to capture relevant features from ECG signals. They improved the efficiency of the BiLSTM technique using the artificial flora optimization (AFO) algorithm as a hyperparameter optimizer. Furthermore, they employed a fuzzy deep neural network (FDNN) classifier to assign class labels to the ECG signals. The performance of the IoTDL-HDD was evaluated on biomedical ECG signals, and the experimental outcomes demonstrate the superiority of the proposed model, achieving high accuracy.

Obayya et al. [61] introduced an approach named "ODL-SCDC," which combines DL with IoT technology to detect skin cancer using connected devices and sensors. They utilized IoT to collect skin images with abnormalities, employing high-resolution cameras and specific sensors in wearable devices. Their proposed ODL-SCDC model incorporated advanced techniques to enhance skin cancer classification, including hyperparameter selection and feature extraction. The results demonstrated the

superiority of the ODL-SCDC model over other methods, showcasing its ability to identify skin lesions accurately.

Wang et al. [62] intended to let users and the emotional representation of music engage in a shared emotional relationship. Their method includes identifying music's emotional representation through multi-modal technologies and IoT sensors. As part of their procedure, they classified the emotions of musical compositions and consumers' sentiments about the music using a Naive Bayes Classifier based on DL. The researchers then analyzed users' emotional experiences and music's emotional representation. The findings showed that different emotional categories in music had different recognition rates under Naive Bayes categorization.

Kumar et al. [63] proposed the Blockchain-orchestrated DL approach for Secure Data Transmission (BDSDT). They aimed to enhance data integrity and security in IoT-enabled healthcare systems. Their approach began with introducing a scalable blockchain architecture, leveraging Zero Knowledge Proof (ZKP) for data integrity and secure transmission. BDSDT integrated with off-chain storage using the InterPlanetary File System (IPFS) to tackle data storage cost issues and employs an Ethereum smart contract to address data security concerns. The authenticated data was then utilized in a DL architecture for intrusion detection in the healthcare system network. Experimental results on two public datasets (CICIDS-2017 and ToN-IoT) demonstrated that BDSDT achieved high accuracy.

Srivastava et al. [64] focused on the role of data analysis in the Internet of Medical Things (IoMT) to enhance the efficiency and automation of healthcare systems. Specifically, they employed ML techniques to use electrocardiogram (ECG) and electroencephalogram (EEG) signals for health condition analysis. DL is highlighted as a successful ML approach, efficient with time-varying signals. In the experiment, both EEG and ECG signal datasets sourced from the PhysioNet biosignal repository were used. They introduced a hybrid DL architecture to analyze these signals. The proposed method demonstrated notable accuracy, achieving a high accuracy rate for classifying ECG signals and a high accuracy rate for identifying bodily activity from EEG signals.

### 5.5 Smart city

The convergence of smart city and DL technologies represents a transformative synergy with great promise for enhancing urban living. In the context of smart cities, DL can be applied to optimize various aspects, such as traffic management, energy consumption, and waste management. In this subgroup, articles in smart cities that have utilized DL to address their challenges are categorized. Table 7 summarizes these articles and compares them according to their challenge, domain, used dataset, DL method, evaluation environment, and obtained result.

Mohamed Shakeel et al. [65] offered a novel behavior recognition and computation technique (BRCS) to anticipate cow behavior while considering contemporary agricultural and animal cultures. The suggested technique employed a recurrent DL model to identify the patterns repeatedly. The pattern was separated into idle and non-idle data to increase the accuracy of the predictions. To categorize anomalies, distinct data structures were traced for the sequential time and analyzing data. The performance of the suggested approach was confirmed using metrics for metric precision, accuracy, calculation time, and average error.

Yang et al. [66] proposed Modified Proximal Policy Optimization (Modified PPO). This algorithm was perfectly suitable for the SD-IoT traffic control system. They iteratively alter the clip hyperparameter to constrain the maximum range between the existing and future policies. Moreover, based on the SD-IoT data acquired to improve urban traffic control efficiency, the suggested algorithm manages traffic signals and cars globally. Experimental findings for various vehicle counts indicated that the presented technique is more efficient and reliable than the original methodology. Researchers should modify the incentive function to improve intelligent agents' policy. They should also consider additional elements like the space between cars and junctions to broaden the reward function and bring it closer to reality.

Liu et al. [67] worked on a water quality prediction model that requires reliable information. With the help of comprehensive DL theories and using the LSTM DNN's good efficiency in estimated time, a drinking-water quality model was developed and built to anticipate large water quality data. However, the suggested approach only considers one aspect, although more intricate data sets with many parameters are available for sequencing in water quality monitoring. They want to improve the model by fusing several water quality metrics and the set of To increase model accuracy, forecast target variables using multidimensional input data.

Contreras-Castillo et al. [68] suggested SAgric-IoT, an IoT and CNN framework for precision farming, to track environmental and physical factors, diagnose early disease, and autonomously manage fertilization and watering in greenhouses. According to the findings, SAgric-IoT is a dependable IoT platform with low packet loss levels, significantly reducing energy usage. It also offers a procedure for accurately identifying diseases and classifying them. The suggested model must be implemented in real-time manufacturing monitoring after being trained using all the plant photos. In this approach, the health of the harvests may be evaluated using samples of their leaves, and the security camera system's photos can be used to recognize the plants precisely.

**Table 7** Smart city category research articles

| Ref | Challenge | Domain | Used DL method | Evaluation environment | Obtained result |
|---|---|---|---|---|---|
| Mohamed Shakeel et al. [65] | Farming and animal husbandry applications | Smart farming | K-nearest neighbor | Implementation | Accuracy, precision, computation time, and mean error measures are used to verify the performance of the suggested system |
| Yang et al. [66] | Urban traffic congestion | Transportation | DBN and Restricted Boltzmann Machines | Simulation | The proposed method is more competitive and stable than the original algorithm |
| Liu et al. [67] | Prediction of water quality | IoT Environment | LSTM | Implementation | The proposed model can predict the quality of drinking water |
| Contreras-Castillo et al. [68] | Greenhouse monitoring | Agriculture and IoT | CNN | Implementation using the Public PlantVillage dataset | The suggested paradigm is a dependable IoT platform with little packet loss, significantly lowering energy usage |
| Singh et al. [69] | Centralized power, transmission delay, security and privacy, and flexibility | Smart city | LSTM | Simulation using Corda v 3.0 and CordaDApp node and smart factory dataset | Improving the performance |
| Vinayakumar et al. [70] | Persistent threats of botnets | Smart city | CNN and LSTM | Implementation using Scikit-learn, TensorFlow, and Keras | Improvement in F1-score, speed of detection, and false alarm rate |
| Simla et al. [71] | Tween people and animals in agriculture | IoT and image processing | CNN | Implementation using CIFAR-10 dataset | Improved early warning system for animal intrusion |
| Rezaee et al. [72] | Optimizing emergency medical vehicle routes | Smart Cities, IoT, ITS | Hybrid Cascade-ResNet | Implementation | Improvement in route optimization over similar methods |
| Arepalli et al. [73] | Early hypoxia detection in aqua water | Aquaculture | Spatially Shared Attention-LSTM | Implementation | Attained heigh accuracy in predicting hypoxia conditions |
| Arepalli et al. [74] | Water contamination analysis in aquaculture | Aquaculture | AODEGRU | Implementation | Achieved heigh accuracy on a public dataset and on a real-time dataset |

Singh et al. [69] suggested a distributed ledger DL-based IoT-oriented architecture in which Software-Defined Networking (SDN) defines the protocols for data transit in the network, and the blockchain creates a distributed environment during the CPS communication phase. The DL-based cloud at the application layer was used in the suggested design to address communication latency, centralization, and scalability difficulties. According to the examination of the results of our installation, performance increased. They simulated the suggested approach using Corda v 3.0 and the CordaDApp node.

Vinayakumar et al. [70] suggested a botnet detection system based on a two-level DL framework for semantically differentiating botnets and acceptable activities. The framework's prospective architecture for evaluating DNS data makes it highly scalable on a server with commodity hardware. At the initial level of the system, siamese networks are used to calculate the similarity measures of DNS requests based on a preset threshold for choosing the most often occurring DNS information across Ethernet connections. A domain-creation technique based on DL architectures is recommended for the second framework level to classify regular and abnormal domain names. Using two datasets, the suggested framework was assessed and contrasted with current DL models. The experimental outcomes showed considerable gains in F1-score, detection speed, and false alarm rate.

Simla et al. [71] addressed conflicts between humans and animals, posing threats to crops and human life and resulting in resource loss. The authors proposed a solution using wireless sensors equipped with an animal intrusion warning system. These sensors, strategically placed around the farm, detect movement and trigger a camera to capture an image in the case of unauthorized intrusion. The assessment included a comparison of lightweight communication protocols, evaluating factors such as response time, packet delivery ratio, energy consumption, latency, bandwidth requirements, and security.

Rezaee et al. [72] focused on leveraging unmanned aerial vehicles (UAVs) as edge devices for the IoT in smart cities and intelligent transportation systems (ITSs). Their primary objective was to optimize the routes of emergency medical vehicles based on population behavior, overcrowding, and abnormal activity patterns. They explored video surveillance through the Internet of Multimedia Things to identify the most efficient routes for emergency medical vehicles during abnormal or overcrowded situations. The proposed approach involved using a hybrid Cascade-ResNet to detect street overcrowding, utilizing multiple data points for congestion detection. They claimed an improvement over similar methods, highlighting the proposed approach's effectiveness, flexibility, and accuracy.

Arepalli et al. [73] presented a lightweight SSA-LSTM model for early hypoxia detection in aquaculture ponds. The authors goaled to achieving remarkable accuracy in detection. Their approaches strengths lie in its spatial and temporal dependency capture, outperforming traditional LSTM models. They used Aquaculture ponds dataset. Their future work could focus on scalability and real-world deployment challenges.

Arepalli et al. [74] addressed water contamination in aquaculture, the authors proposed a framework that integrates IoT-based data collection with an attention-based Ordinary Differential Equation Gated Recurrent Unit (AODEGRU) model for accurate classification of water contamination levels. The approach utilizes real-time data to evaluate water quality indicators and calculate a contamination index, achieving high accuracy rates on both public and real-time datasets. They lacked scalability and real-world deployment considerations.

## 6 RQs and discussion of results

We thoroughly reviewed and analyzed related research articles on DL-based IoT in the previous section. In this section, according to the analysis done in the 56 articles mentioned in the previous section, we will discuss the results and answers to our six RQs.

This section is organized into separate subsections, each dedicated to specific aspects of our research. Through a systematic exploration, we aim to provide comprehensive insight into the complexities of DL-based IoT applications. The following subsections elaborate on the different dimensions of our research and clarify DL models, research contexts, evaluation tools, datasets, evaluation criteria, and future directions in the DL-based IoT field.

### 6.1 DL models within IoT applications

This subsection presents a comprehensive survey of DL models used in IoT applications, according to the articles reviewed in the previous section. To answer RQ1, we discuss DL models and reveal their specific applications in the context of the IoT. In this section, we identify the models and clarify their distinct use cases, training paradigms, strengths, and weaknesses.

#### 6.1.1 RQ1: What are the DL models used in the IoT applications? And what are their uses?

DL models play an important role in IoT applications. DL models provide the ability to process large amounts of data generated by IoT devices for various tasks such as classification, prediction, anomaly detection, and more. Table 8

**Table 8** Comparative analysis of DL models in IoT applications

| DL model | Training paradigm | Strength | Weakness | Application in IoT |
|---|---|---|---|---|
| CNNs | Supervised | ✓ Excellent for image-based IoT applications<br>✓ Can automatically learn hierarchical features<br>✓ Robust to spatial variations | ✓ Limited to grid-like data<br>✓ May not be suitable for non-grid IoT data | ✓ Medical data analysis<br>✓ Image analysis<br>✓ Object recognition<br>✓ Video analysis<br>✓ Surveillance in IoT environments<br>✓ Recommendation systems |
| RNNs | Supervised | ✓ Suitable for sequential data in time-series IoT applications<br>✓ Captures temporal dependencies | ✓ Gradient vanishing/exploding problems<br>✓ May struggle with long-term dependencies | ✓ Image classification<br>✓ Speech recognition<br>✓ Time-series prediction<br>✓ Natural language processing in IoT devices<br>✓ Handling sequential sensor data |
| LSTM | Supervised | ✓ Overcomes the vanishing/exploding gradient problem in RNNs<br>✓ Well-suited for time-series data in IoT | ✓ More complex than basic RNNs<br>✓ May require more computational resources | ✓ Predicting missing data in healthcare scenarios<br>✓ Detecting congestive heart failure<br>✓ Modeling traffic flow in urban road networks |
| GRU | Supervised | ✓ Similar to LSTMs but computationally more efficient<br>✓ Suitable for more straightforward IoT applications with sequential data | May not capture long-term dependencies as effectively as LSTMs | Suitable for simpler IoT applications with sequential data |
| Forward Neural Networks (FNNs) | Supervised | ✓ Simple architecture<br>✓ Good for basic tasks | ✓ Limited capacity to capture complex patterns<br>✓ May not perform well in highly dynamic IoT environments | ✓ Sensor data classification<br>✓ Basic decision-making in IoT systems |
| RBM | Unsupervised, Supervised | ✓ It is helpful in classification, feature extraction, and dimensionality reduction<br>✓ Extract complex features from data<br>✓ Can be trained in a supervised manner<br>✓ Can perform classification tasks | ✓ Complex training procedure<br>✓ Limited applicability to sequential data<br>✓ RBMs may struggle to capture long-term dependencies in data | ✓ Collaborative filtering<br>✓ Intrusion detection<br>✓ Topic modeling<br>✓ Medical data analysis |
| AE | Unsupervised | ✓ Useful for unsupervised learning<br>✓ Anomaly detection in IoT data | ✓ May require large amounts of labeled data for pre-training<br>✓ Might not perform well in scenarios with highly variable data | ✓ Image generation<br>✓ Recommendation systems<br>✓ Unsupervised learning and anomaly detection in IoT data<br>✓ Detecting unusual patterns or outliers in sensor data |
| GANs | Semi-supervised | ✓ Can generate synthetic data for training<br>✓ Useful for scenarios with limited labeled data | ✓ Training GANs can be challenging, and the generated data may not perfectly represent the real data distribution | ✓ Enhancement of medical data analysis<br>✓ Synthetic data generation for training models<br>✓ Image generation<br>✓ Quality improvement<br>✓ De-noising in IoT applications |

**Table 8** (continued)

| DL model | Training paradigm | Strength | Weakness | Application in IoT |
|---|---|---|---|---|
| DBN | Unsupervised, Supervised | ✓ Trained in a greedy way<br>✓ Hierarchical feature learning<br>✓ Capability for both unsupervised and supervised learning:<br>✓ Good performance in sequential data | ✓ Dependency on pre-training<br>✓ Computational complexity<br>✓ Require large amounts of labeled data for supervised training | ✓ Image recognition<br>✓ Drug discovery<br>✓ Medical image analysis<br>✓ Text classification<br>✓ Natural language processing<br>✓ Speech recognition<br>✓ Intrusion detection |
| Transfer learning | Supervised | ✓ Enables the use of pre-trained models on similar tasks<br>✓ Useful for scenarios with limited labeled data | ✓ May not generalize well to diverse IoT datasets<br>✓ Computational demands may be high for certain IoT applications | ✓ Adapting pre-trained models to new tasks in IoT scenarios with limited labeled data<br>✓ Improving performance in situations with challenges in obtaining large labeled datasets |

highlights a comparative analysis of different DL models in IoT applications.

Figure 8 shows the percentage of AI methods used by the articles in the previous section. We can claim that most researchers used LSTM, CNN, and DBN in the DL-based IoT domain.

## 6.2 DL-based IoT domains

Because the core of our research is the interaction between DL and the IoT, the answer to RQ2 is the focus and directs our attention to the broad field of research in this area. In this section, we systematically outline the various research fields that contribute to the synergy of DL-based IoT according to the articles we analyzed and categorized in the previous section.

### 6.2.1 RQ2: What research fields are there in the field of DL-based IoT?

According to the 56 selected articles, we can say that the articles deal with various research fields that intersect DL-based IoT. In the domain of DL-based IoT, we found vital research areas in the analyzed articles, each representing a specialized area where these technologies have yielded valuable results. The chart in Fig. 9 shows the percentages related to fundamental research areas:

Security, with 36% as a critical aspect in this context, encompasses a range of challenges, including network-based attacks, software vulnerabilities, and intrusion detection, emphasizing the need for robust cybersecurity measures. Edge and fog computing, with 16%, introduce real-time data processing, low-latency applications, and distributed computing that provide a foundation for increased responsiveness and decentralized data management.

Energy management invests 9% in smart grid technologies, forecasting electricity demand and prices, and monitoring energy consumption, all of which contribute to sustainable practices.

Healthcare applications (21%), including medical imaging, drug discovery, and disease prediction, demonstrate the profound impact of DL-based IoT in revolutionizing the healthcare landscape. Finally, smart city initiatives with 18%, including smart transportation, infrastructure, urban modeling, and smart homes, emphasize the transformative potential of integrating DL-based IoT technologies to create smart and connected urban environments.

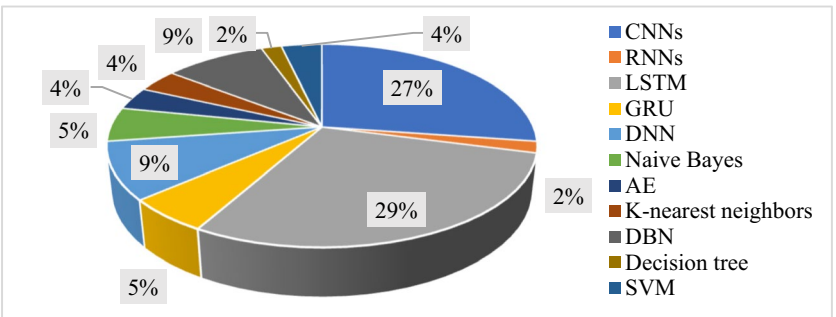



**Fig. 8** Percentage of main AI methods in DL-based IoT area

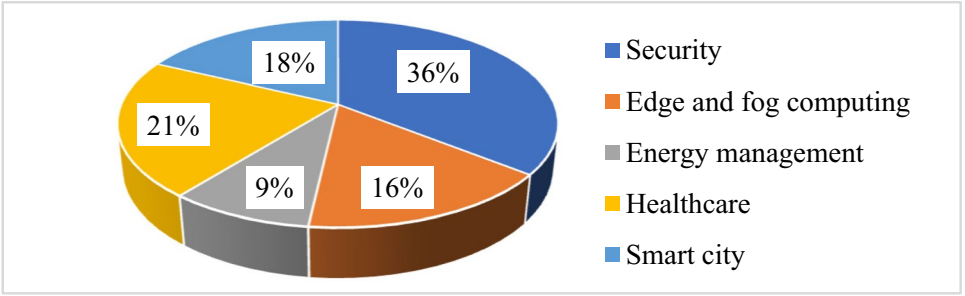



**Fig. 9** Percentage of main categories in DL-based IoT area

Considering these percentages, we can claim that they significantly emphasize security considerations, followed by a substantial contribution to edge and fog computing, healthcare applications, energy management, and advancements in smart city designs.

## 6.3 Assessment tools and simulators for evaluating DL-based IoT

In this section, we answer RQ3, which guides our exploration of the methods and frameworks employed to evaluate studies in the dynamic context of DL-based IoT. We look closer at the tools and simulators central to assessing the effectiveness and implications of different approaches in this field.

### 6.3.1 RQ3: What tools and simulators are used to evaluate the studies in the field of DL-based IoT?

By reviewing the research articles in the DL-based IoT field, we found that most % of the articles, 70% of them used implementation to evaluate their innovation. On the other hand, 27% used the simulation method to evaluate their idea. Also, 3% of the articles used both methods, i.e., simulation and implementation. Figure 10 shows the percentage of different methods used to evaluate the proposed idea in DL-based IoT according to the 56 articles analyzed in Section 5.

A software tool or a simulator is certainly needed for simulating or implementing. Most reviewed articles used Python, especially the TensorFlow framework, for implementation or simulation. The names of some simulators and tools are shown in Fig. 11, including MATLAB, Corda, Cooja, WILL API, Keras-Python, Scikit-Learn, Python, and TensorFlow-Python.

In Table 9, we provide a critical evaluation of the tools and simulators used for evaluating DL-based IoT studies. We coonsider discussing the advantages, limitations, and suitability of each tool for different types of IoT applications and DL models.

## 6.4 Datasets in the DL-based IoT landscape

The use of datasets is an essential aspect of research in DL-based IoT. Addressing RQ4, this section presents a comprehensive exploration of the datasets that underlie the studies on DL-based IoT that we analyzed and mentioned in the previous section.

### 6.4.1 RQ4: Which datasets are used in the field of DL-based IoT?

According to the review of the articles in Section 5 and Fig. 12, we can claim that 45% of the articles used datasets to evaluate their innovation, and 55% did not use any datasets. The most common datasets are listed below:

- NSL-KDD dataset [20]
- IoT network-traffic dataset [19]
- Benchmark dataset [23, 24]
- 01Modbus network traffic dataset [27]
- Car-hacking and DDoS datasets [28]
- IoT-Botnet datasets [31, 34]
- Malware fingerprints dataset [32]
- GP dataset [33]

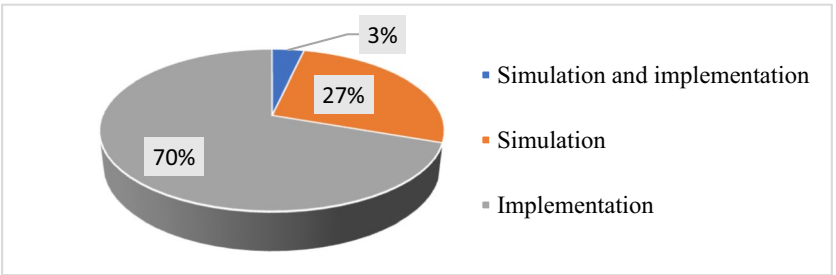



**Fig. 10** Percentage of evaluation environment in DL-based IoT domain

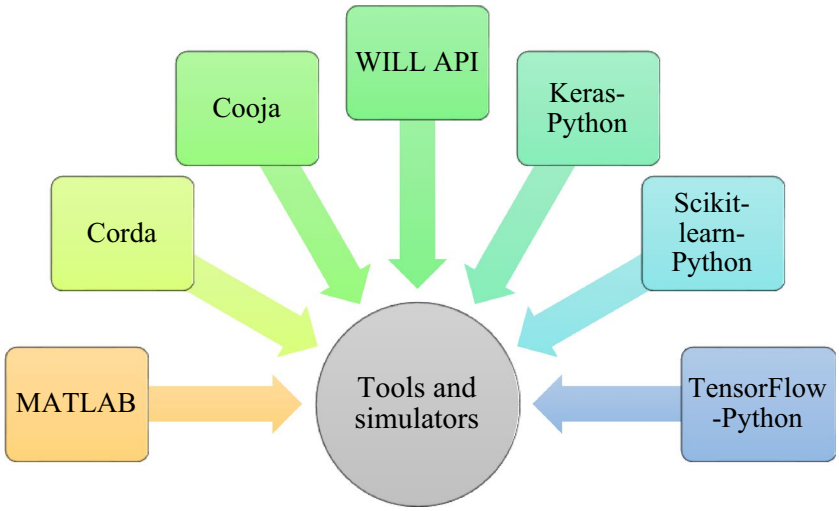



**Fig. 11** Evaluation tools and simulators in the DL-based IoT domain

- SWaT dataset [33]
- AWID dataset [44]
- CIC-IDS2107 dataset [44]
- SGCSs dataset [49]
- Brain, lung, breast cancer, and Alzheimer's disease datasets [53]
- UR fall detection dataset [55]
- Human activity datasets [58]
- CICIDS-2017 dataset [63]
- ToN-IoT dataset [31, 63]
- Public PlantVillage dataset [68]
- CIFAR-10 dataset [71]
- N-BaIoT [26]

In the following we offer a thorough examination of the datasets used in DL-based IoT research. We discuss the characteristics of these datasets, including size, diversity, and relevance to real-world IoT scenarios. Ans also we address any potential biases or limitations associated with the datasets.

- NSL-KDD dataset
  - Characteristics: A subset of the KDD Cup 1999 dataset, it contains network traffic data generated in a military network environment. It includes both normal and various types of attacks.
  - Size: Moderate.
  - Diversity: Contains a variety of attack types, making it suitable for intrusion detection system (IDS) research.
  - Relevance: Relevant to cybersecurity applications in IoT networks.
  - Limitations: Being a subset, it may not fully represent the complexities of real-world IoT environments. Also, it's relatively outdated, potentially lacking recent IoT-related attacks.
- UNSWNB15 dataset
  - Characteristics: Contains network traffic data with nine different attack scenarios, collected from a real-world IoT testbed.
  - Size: Moderate to large.
  - Diversity: Covers various attack scenarios, providing a more realistic representation of IoT network traffic.
  - Relevance: Highly relevant as it captures real-world IoT network behaviors and attacks.
  - Limitations: Limited to the specific attack scenarios included in the dataset, potentially lacking in variety compared to continuously evolving real-world threats.
- SWaT dataset
  - Characteristics: Captures operational data from a real water treatment plant.
  - Size: Large.
  - Diversity: Represents real-world industrial processes in IIoT settings.
  - Relevance: Highly relevant for research on IIoT security and anomaly detection.
  - Limitations: Limited to the specific industrial process of a water treatment plant, may not generalize to other IIoT environments.

**Table 9** Evaluation tools and simulators advantages, limitations, and suitability

| Evaluation tools and simulators | Advantages | Limitations | Suitability |
|---|---|---|---|
| MATLAB [53] | MATLAB provides a comprehensive platform for prototyping and implementing DL models for developers and researchers. MATLAB offers a wide range of functions and toolboxes specifically designed for signal processing, image processing, and machine learning | MATLAB's performance may not scale well for large datasets or complex DL models | MATLAB is suitable for IoT applications where rapid prototyping and simulation are essential, such as predictive maintenance, anomaly detection, and smart sensor networks |
| Corda[69] | Corda is a distributed ledger platform designed for secure and efficient transactions. It can be useful in IoT applications involving secure data exchange and decentralized control | Corda's focus is on financial transactions and contract execution, so its direct applicability to DL-based IoT applications may be limited | Corda can be suitable for IoT applications where secure and transparent data exchange between devices or entities is critical, such as supply chain management or smart contracts in industrial IoT |
| Cooja[19] | Cooja is a network simulator specifically designed for IoT applications. It allows developers to simulate large-scale IoT networks and evaluate the performance of DL models in a virtual environment | Cooja's simulations may not always accurately reflect real-world conditions, and it may lack some advanced features present in other simulators | Cooja is suitable for testing and optimizing DL models in IoT applications involving sensor networks, energy management, and network protocols |
| Keras-Python [26, 70] | Keras is a high-level DL library that offers a user-friendly interface and seamless integration with TensorFlow. It enables rapid prototyping of DL models and supports both CPU and GPU acceleration | Keras may not be as customizable or suitable for low-level optimizations compared to other DL frameworks | Keras is suitable for IoT applications where simplicity, ease of use, and quick experimentation are priorities, such as image recognition, time series forecasting, or sensor data classification |
| Scikit-learn-Python[70] | Scikit-learn is a versatile machine learning library that provides a wide range of algorithms and tools for data preprocessing, model selection, and evaluation. It is easy to use and well-suited for small to medium-sized datasets | Scikit-learn focuses on traditional machine learning algorithms and may not directly support DL models without additional integration with DL frameworks | Scikit-learn is suitable for IoT applications where traditional machine learning techniques suffice, such as classification, regression, clustering, or anomaly detection with moderate-sized datasets |
| TensorFlow-Python[20, 26] | TensorFlow is a DL framework with extensive support for building and deploying DL models at scale. It offers flexibility, performance optimizations, and integration with various hardware accelerators | TensorFlow's learning curve can be steep for beginners, and its lower-level APIs may require more coding effort compared to higher-level libraries like Keras | TensorFlow is suitable for IoT applications where scalability, performance, and customizability are crucial, such as deep sensor fusion, large-scale image recognition, or natural language processing in IoT devices |

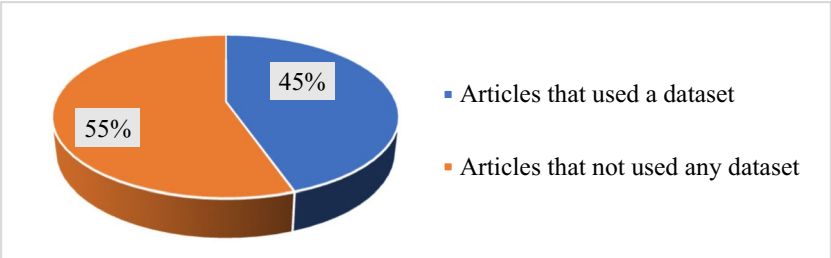



**Fig. 12** Percentage of using datasets in DL-based IoT domain

- AWID dataset
  - Characteristics: Contains RFID data collected in a controlled environment.
  - Size: Large.
  - Diversity: Covers various RFID tag readings.
  - Relevance: Relevant for RFID-based IoT applications and security research.
  - Limitations: Limited to RFID data, may not represent other types of IoT devices or communication protocols.
- CIC-IDS2107 dataset
  - Characteristics: Includes network traffic data collected in a simulated IoT environment with various attacks.
  - Size: Moderate to large.
  - Diversity: Covers multiple attack scenarios in an IoT context.
  - Relevance: Relevant for evaluating IDS in IoT environments.
  - Limitations: Simulated environment may not fully reflect the complexities of real-world IoT networks.
- CICIDS-2017 dataset
  - Characteristics: Includes network traffic data with benign and malicious activities.
  - Size: Large.
  - Diversity: Covers various types of attacks.
  - Relevance: Relevant for evaluating IDS and anomaly detection systems in IoT environments.
  - Limitations: May not capture the intricacies of IoT-specific attacks comprehensively.
- ToN-IoT dataset
  - Characteristics: Captures IoT network traffic in a controlled environment.
  - Size: Moderate to large.
  - Diversity: Includes benign and malicious traffic.
  - Relevance: Relevant for evaluating security mechanisms in IoT networks.
  - Limitations: Controlled environment may not fully represent real-world IoT scenarios.
- Public plantvillage dataset
  - Characteristics: Contains images of diseased and healthy plant leaves.
  - Size: Large.
  - Diversity: Covers various plant species and diseases.
  - Relevance: Relevant for IoT applications in agriculture, such as plant disease detection.
  - Limitations: Limited to plant-related applications, may not be directly applicable to other IoT domains.
- CIFAR-10 dataset
  - Characteristics: Contains images of objects categorized into ten classes.
  - Size: Large.
  - Diversity: Covers a wide range of object categories.
  - Relevance: While not specifically IoT-related, it can be used for image-based IoT applications like object recognition.
  - Limitations: General-purpose dataset, may not capture IoT-specific features or contexts.

## 6.5 Evaluation criteria in the realm of DL-based IoT

The answer to RQ5 specifies the evaluation method in the DL-based IoT field. In this section, we detail the common evaluation criteria used to measure the effectiveness and robustness of the proposed DL approaches in the previous section on IoT.

### 6.5.1 RQ5: What are the common evaluation criteria used in the DL-based IoT domain?

According to the articles analyzed in the previous section, we can claim that in the field of IoT, performance evaluation of DL models includes using various evaluation criteria to evaluate how capable the models are in performing

**Table 10** Common evaluation metrics that are used in the DL-based IoT domain

| Evaluation metric | Definition | Application in IoT |
|---|---|---|
| Accuracy [41, 44, 61] | Accuracy is a measure that is used to verify the overall accuracy of model predictions | • Critical infrastructure Monitoring<br>• Crop health monitoring and farm management in smart agriculture<br>• Health care<br>• Accurate and reliable control of home automation devices |
| Precision [42, 68] | Precision is a measure that measures the ratio of true positive predictions to the total predicted positives | • Use in intrusion detection systems for IoT networks<br>• Use in traffic management in smart cities<br>• Use of smart transportation systems in smart cities<br>• Energy efficiency<br>• Accurate health monitoring by ensuring accurate detection of health-related parameters for |
| Recall (Sensitivity) [21] | Recall is a measure that measures the ratio of true positive predictions to total true positives | • Application in health care monitoring devices<br>• Application in environmental sensors<br>• Application in strengthening power distribution systems by responding to faults in real time |
| F1-Score [28, 70] | F1-Score is the average of precision and recall and balances the two | • Optimizing the use of water resources in agriculture and urban environments<br>• Increasing efficiency and safety in transportation systems<br>• Collection and analysis of environmental data |
| Mean Squared Error (MSE) / Mean Absolute Error (MAE) [52] | MSE measures the mean squared difference between predicted and actual values, but MAE measures the mean absolute difference | • Forecasting energy consumption<br>• To evaluate the accuracy of predictions for various sensor data, such as temperature, humidity, or pressure<br>• Assessing the accuracy of predictive models of energy consumption in smart buildings |
| Latency [44, 71] | Latency is used to measure the time delay between the input and the prediction of the model | • Application in self-driving vehicles<br>• Application in decision-making processes in IoT systems<br>• Application in energy efficiency<br>• Measuring latency in decision-making processes to ensure real-time responsiveness and safety in autonomous driving<br>• Application in monitoring the time delay in energy-related decisions to optimize energy consumption |
| Energy efficiency [50, 75, 76] | The purpose is to evaluate the energy consumption of DL models during inference | • Optimizing energy consumption in IoT devices<br>• Assess the energy efficiency of control models of existing systems in smart buildings |
| Scalability [51] | It is used to evaluate the model's ability to scale with increasing data or devices | • Assessing scalability of models in augmented data management in healthcare IoT applications<br>• Assessing the scalability of IoT solutions in managing the growing number of connected devices for urban infrastructure |
| Cost-effectiveness [63, 77, 78] | It evaluates the model's efficiency regarding hardware, communication, and computing costs | • Optimizing the use of resources for cost-effective implementation of the IoT<br>• Assessing the cost-effectiveness of the model in terms of hardware, communication, and computing costs for efficient IoT deployment |

their tasks. Table 10 lists some of the common performance measures used in the literature.

### 6.6 Future directions and open issues

This section focuses on the future directions and challenges shaping the DL and the IoT path. According to the reviewed articles and the existing challenges and research gaps, in this section, we address the open issues of the primary challenges in the field of DL-based IoT.

#### 6.6.1 RQ6: What are the future directions and open issues in DL-based IoT?

Data sources are essential to the effectiveness of DL techniques. Due to the absence of substantial datasets, it is challenging to use DL in the IoT; additional data is required to improve DL's accuracy. The generation of raw data in DL models in an appropriate manner is another challenge faced by IoT applications. To provide more accurate findings, several DL approaches call for data pretreatment.

Preprocessing is more difficult for IoT applications since the system must deal with data from several sources with varying distributions and formats while representing missing data. Implementing data-collecting systems is indeed a crucial study area. The number of active sensors and their placement impact the data's quality. You must develop a data collecting module for the entire IoT system layout, even if the model architecture is adequately constructed. It needs to be a more dependable, cost-effective, and trustworthy model. Because we get data from several sources, security is the most significant concern in the IoT context. Due to the huge amount of IoT data distribution over the Internet for inspection and accessibility on a global scale, preserving data safety and secrecy is a top priority in many IoT applications. Anonymization is helpful for many purposes, but it may also be misused to re-identify data as anonymous.

DL models may benefit from any corrupt data stream since they are learning the features of the original data. DL models must be updated using specific techniques to locate irregular or false data. To address the demands of DNN management in resource-constrained devices, DL design is a significant problem for IoT system designers. This is anticipated to rise when new algorithms are added to IoT DL systems and dataset dimensions increase daily. DL also has several drawbacks. Complex and expensive IoT sensor technologies make it challenging to monitor off-road vehicles digitally. Strong dependence on cloud/fog computing in off-grid distant places makes network accessibility and expertise problematic. Using edge devices, such as cell phones with computing capabilities, is a solution that has yet to be marketed. High-end technology is needed since DL is an effective tool for analyzing large amounts of data from the IoT.

Designing a DL model for an embedded device with limited resources is still challenging. In gathering and transferring data to servers and performing analysis, we may experience network difficulties and data leaks. There is a growing trend to create a framework for cloud-based learning that uses both the cloud and popular devices. A cloud-based device exploits the edge to minimize latency, increase safety and security, and employ sophisticated methods for data storage. Moreover, it may train top-notch computational models and share data around cutting-edge technologies via the cloud.

In Table 11, we presented a comprehensive overview of open issues and challenges, along with corresponding future directions, within the domain of DL-based IoT. Table 11 addresses critical aspects of DL-based IoT applications, ranging from privacy protection and dataset constraints to security, resource constraints, and the integration of edge-to-cloud computing.

## 7 Future agenda

DL has significantly impacted the IoT by enabling intelligent data analysis and decision-making in various IoT applications. In this section, we discuss research areas, trends, emerging applications, advancements, and limitations in DL-based IoT areas [79, 81, 82].

### 7.1 Research areas, trends, and emerging applications

Table 12 illuminates the broad spectrum of applications and challenges in harnessing DL to IoT capabilities across diverse domains. Ongoing research and innovative endeavors in these spheres are pivotal for unleashing the complete potential of DL-infused IoT solutions to tackle real-world issues and enhance quality of life. Additionally, trends delineate the evolving panorama of DL-driven IoT research and development, underscoring nascent opportunities and imperative areas for advancement and innovation in the field. The ensuing applications unveil the dynamic landscape of DL-fueled IoT research and development, elucidating burgeoning prospects and pivotal focal points for innovation and advancement within the domain.

### 7.2 Advancements and limitations

In the realm of emerging technologies, significant advancements are continuously reshaping the landscape. In the context of DL-based IoT systems. However, alongside these advancements, it is essential to acknowledge the existing

**Table 11** Open issue/challenge and future directions in the DL-based IoT domain

| Open issue/challenge | Future direction |
|---|---|
| Privacy protection in IoT networks [31, 38, 69] | • Improving and developing different privacy protection solutions using DL<br>• Research and development in DL models for privacy protection with higher accuracy<br>• Integrating state-of-the-art encryption techniques into DL models to protect data<br>• Establishing evaluation criteria to relate DL models to IoT data better<br>• Research to improve the performance of DL models in privacy protection<br>• Developing solutions to deal with challenges and skills in data security in the IoT<br>• Developing new methods to evaluate and measure the effectiveness of privacy protection solutions in real environments |
| Constraints on datasets for DL in IoT [36, 37, 79] | • Research to increase the size of DL datasets in the field of IoT<br>• Development of up-to-date techniques for data preprocessing to improve the accuracy of DL models<br>• Research in the field of solutions that improve the quality of data in the IoT<br>• Developing solutions for optimizing DL models in resource-constrained environments |
| Security in DL for diagnosis [19, 24, 34] | • Providing solutions to reduce time and computational consumption in DL to detect attacks in IoT<br>• Integration of different techniques to increase accuracy in detecting attacks in DL models<br>• Development of new solutions to evaluate the effectiveness of DL models in detecting attacks<br>• Researching integrating DL with other security methods, such as blockchain and ensemble learning |
| Designing DL in resource-constrained networks [35, 52, 80] | • Development of low-volume and high-performance DL models in resource-constrained networks<br>• Integrating strategies to transfer learning in resource-constrained environments to improve DL<br>• Research combining DL with technologies such as blockchain to enhance security and efficiency<br>• Development of criteria and methods to evaluate the efficiency of DL models in real conditions of IoT |
| Energy efficiency [51, 75, 76] | • Research on hardware architectures for energy-efficient DL computing in IoT devices<br>• Development of adaptive power management strategies to dynamically regulate the energy consumption of IoT devices based on the processing needs of DL tasks<br>• Research on compression algorithms for model parameters and data to minimize energy consumption during communication and storage<br>• Designing DL training algorithms that consider energy efficiency as a critical optimization criterion<br>• Research on integrating renewable energy sources, such as solar or kinetic energy, to power IoT devices to host DL models |
| Real-time processing [37, 74] | • Explore predictive analytics to predict future events in IoT data streams for proactive and timely decision-making<br>• Research into forecasting models that use historical data to predict trends and patterns in real-time<br>• Investigate low-latency communication protocols and technologies to minimize latency in data transmission between IoT devices and processing units<br>• Research on edge computing architectures for ultra-low latency processing<br>• Explore load-balancing strategies to ensure optimal distribution of processing tasks across edge and cloud resources |

limitations that pose challenges to further innovation and deployment. Addressing these limitations will be crucial for unlocking the full potential of DL-based IoT technologies and ensuring their sustainable growth and adoption in diverse domains. In Table 13 we list the advancements and limitations of DL-based IoT.

## 8 Conclusion

Considering the increasing importance of AI, especially DL, and people's involvement with IoT applications, challenges, and research gaps, we systematically examined, analyzed, and compared 56 research articles published between 2019

**Table 12** Research areas, trends, and emerging applications in DL-based IoT

Research Areas

- **Edge Computing and DL:** DL models are being optimized and deployed on edge devices to process data locally, reducing latency and bandwidth usage. Research focuses on developing lightweight DL models tailored for resource-constrained edge devices [42, 44, 80]
- **Anomaly Detection:** DL techniques such as autoencoders and RNNs are used for anomaly detection in IoT data streams. This includes identifying unusual patterns in sensor data indicating potential faults or security breaches [29, 33, 34]
- **Predictive Maintenance:** DL models analyze sensor data to predict equipment failures before they occur, enabling proactive maintenance. Techniques like LSTM networks and CNNs are commonly used for time-series analysis [73]
- **Energy Efficiency:** DL is employed to optimize energy consumption in IoT systems. This includes developing energy-efficient algorithms for data processing, communication, and resource allocation in IoT networks [17, 50, 52]
- **Security and Privacy:** DL-based approaches are utilized for enhancing IoT security by detecting intrusions, identifying malicious activities, and protecting sensitive data. Federated learning and homomorphic encryption are explored to preserve privacy in IoT data sharing scenarios [38, 69]
- **Resource Management:** DL models optimize resource utilization in IoT networks by dynamically allocating resources based on demand and workload. Reinforcement learning (RL) is applied to adaptively manage resources in real-time [17, 35, 37]
- **Dynamic Adaptation and Self-Organization:** DL models capable of dynamically adapting to changing environmental conditions and self-organizing IoT networks for improved efficiency and resilience [35]
- **Time Series Forecasting:** DL techniques for accurate forecasting of time series data in IoT applications, including demand forecasting, weather prediction, and financial market analysis [18, 37, 74]
- **Distributed Computing and Communication:** DL-driven optimization of communication protocols, network routing, and data aggregation strategies in distributed IoT environments for enhanced scalability and reliability [79, 83]
- **Autonomous Systems and Robotics:** Integration of DL techniques into autonomous systems and robotic platforms for perception, navigation, and decision-making in dynamic IoT environments [8, 37]
- **Health Monitoring:** DL-powered wearable devices and IoT sensors for continuous health monitoring, disease prevention, and personalized wellness recommendations [6, 73]
- **Urban Mobility and Transportation:** DL-based solutions for traffic prediction, route optimization, and public transportation management in urban mobility systems, enhancing efficiency and reducing congestion [8, 66, 77]

Trends

- **Federated Learning:** With the proliferation of distributed IoT devices, federated learning is gaining traction. It enables model training across multiple edge devices without centrally aggregating sensitive data, addressing privacy concerns [10, 34]
- **Explainable AI:** As DL models are increasingly deployed in critical IoT applications, there's a growing need for explainable AI techniques to interpret model decisions and ensure transparency and trustworthiness [22, 36]
- **Multi-Modal Data Fusion:** Integrating data from diverse IoT sensors (e.g., audio, video, and environmental sensors) presents opportunities for DL models to extract richer insights and improve decision-making accuracy [50, 66]
- **Blockchain Integration:** Incorporating blockchain technology with DL-based IoT systems to enhance data integrity, traceability, and security, particularly in applications requiring transparent and tamper-proof transaction records [31, 69, 80]
- **Differential Privacy:** Integrating differential privacy mechanisms with DL models to protect sensitive information and preserve individual privacy in IoT data sharing and collaborative learning scenarios [31, 69]
- **Explainable Reinforcement Learning:** Advancing explainable reinforcement learning techniques to interpret and visualize decision-making processes of RL agents in IoT control and automation tasks, enhancing transparency and trustworthiness [22, 36]
- **Health and Environmental Monitoring:** Increasing focus on DL-powered IoT solutions for real-time monitoring of air quality, water quality, and environmental pollutants, supporting public health initiatives and sustainability efforts [68, 73]

Emerging Applications

- **Smart Cities:** DL-based IoT solutions are deployed for traffic management, waste management, energy optimization, and public safety in smart city initiatives [48, 55, 69]
- **Healthcare:** IoT devices integrated with DL models enable remote patient monitoring, personalized treatment recommendations, and early disease detection [2, 57, 79]
- **Precision Agriculture:** DL-powered IoT systems monitor soil conditions, crop health, and weather patterns to optimize agricultural practices and maximize yield [68]
- **Industrial IoT (IIoT):** DL techniques are applied in IIoT for predictive maintenance, quality control, supply chain optimization, and process automation [31, 46]
- **Remote Sensing and Environmental Monitoring:** Deploying DL-based IoT platforms for remote sensing applications, including forest fire detection, wildlife tracking, and habitat monitoring, leveraging satellite imagery and UAVs for environmental conservation efforts [17, 18, 73]
- **Smart Water Management:** Implementing DL-powered IoT solutions for monitoring and managing water resources, including water quality monitoring, leak detection, and irrigation optimization, to enhance water conservation efforts and ensure sustainable water management practices [74]

and April 2024 in DL-based IoT. In this SLR article, we answered six research questions, including the applications of DL in IoT, the DL-based IoT research fields, the tools and simulators used in the field of DL-based IoT, the datasets used in the field of DL-based IoT, and the future agendas and open issues in the domain of DL-based IoT. According to the previous sections, in the field of DL-based IoT, the articles are categorized into five categories: security, edge

**Table 13** Advancements and limitations of DL-based IoT

| Advancements | Limitations |
|---|---|
| • Transfer learning and meta-learning approaches facilitate knowledge transfer across different IoT domains, reducing the need for large labeled dataset<br>• Advances in hardware accelerators (e.g., GPUs, TPUs) and model compression techniques improve the efficiency of DL inference on edge devices<br>• Automated Machine Learning (AutoML): AutoML frameworks streamline the process of model selection, hyperparameter tuning, and architecture optimization, making DL more accessible to non-experts and reducing the barrier to entry for developing IoT applications<br>• Hybrid Models and Ensemble Learning: Combining DL with traditional machine learning algorithms and rule-based systems through hybrid models and ensemble learning approaches improves model robustness, interpretability, and generalization in IoT applications with heterogeneous data sources and modalities<br>• Explainable AI (XAI) Tools: Advances in XAI techniques and interpretability frameworks enable DL models to provide transparent explanations of their decisions and predictions, fostering trust, accountability, and regulatory compliance in safety–critical IoT deployments<br>• Edge Intelligence Platforms: Development of edge intelligence platforms and frameworks facilitates seamless integration and deployment of DL models on edge devices, enabling real-time inference and decision-making at the network edge while minimizing latency and bandwidth requirements<br>• Privacy-Preserving Techniques: Advancements in privacy-preserving DL techniques, such as federated learning, differential privacy, and secure multiparty computation, enable collaborative model training and inference across distributed IoT devices while preserving data privacy and confidentiality<br>• Quantum Computing Integration: Exploration of quantum computing techniques for accelerating DL computations and solving optimization problems in IoT applications with exponential computational complexity, unlocking new possibilities for quantum-enhanced DL in resource-constrained environments | • DL models require large amounts of labeled data for training, which can be challenging to obtain in IoT domains with limited data availability<br>• Data Quality and Bias: Challenges related to data quality, imbalance, and bias in IoT datasets can lead to suboptimal model performance, generalization errors, and unintended consequences, requiring careful data preprocessing and bias mitigation strategies<br>• Scalability and Resource Constraints: Scaling DL models to large-scale IoT deployments with thousands or millions of devices poses scalability challenges in terms of computational resources, memory requirements, and communication overhead, necessitating efficient distributed training and inference techniques<br>• Energy Consumption and Efficiency: DL inference on resource-constrained edge devices may incur high energy consumption and latency, impacting battery life and operational efficiency, prompting the need for energy-efficient model architectures and optimization strategies for IoT deployments<br>• Adversarial Attacks and Security: DL models deployed in IoT systems are vulnerable to adversarial attacks and security breaches, including data poisoning, model inversion, and evasion attacks, necessitating robust defense mechanisms and adversarial training techniques to enhance resilience and security<br>• Data Privacy and Sovereignty: Concerns about data privacy, sovereignty, and ownership in IoT ecosystems raise legal, ethical, and regulatory challenges regarding data collection, storage, sharing, and access control, necessitating privacy-preserving technologies and regulatory frameworks to safeguard user rights and interests<br>• Cost and Resource Allocation: The high cost of DL model development, training, and deployment, coupled with resource allocation constraints in IoT environments with limited computational resources and budgetary constraints, may impede widespread adoption and scalability, requiring cost-effective solutions and resource management strategies to maximize ROI and sustainability |

and fog computing, energy management, smart cities, and healthcare. We found that most of the articles used implementation to evaluate their innovation; on the other hand, others used the simulation method to evaluate their idea. Also, some articles used both methods, i.e., simulation and implementation. Most reviewed articles used Python, especially the Tensor Flow framework, for implementation or simulation. Also, studies used datasets to evaluate their innovation, and some did not use any datasets. We can mention limited data sources, data preprocessing, data quality, data security and confidentiality, complexity and costs, limited resources, and latency minimization as open issues in this field.

**Author contributions** P. Raoufi: Writing original draft, methodology, software, A. Hemmati: Preparation, Visualization, and Investigation A. M. Rahmani: conceptualization, writing, reviewing and editing, validation, and supervision. All authors reviewed the manuscript.

**Funding** No funding was received.

**Data availability** No datasets were generated or analysed during the current study.

## Declarations

**Competing interests** The authors declare no competing interests.

## References

1. Zobaed SM, Hassan M, Islam MU and Haque ME (2021) Deep learning in IOT-based healthcare applications. In: Deep learning for internet of things infrastructure. CRC Press, pp 183–200
2. Bolhasani H, Mohseni M, Rahmani AM (2021) Deep learning applications for IoT in health care: a systematic review. Inform Med Unlocked 23:100550
3. Bhattacharya S, Somayaji SRK, Gadekallu TR, Alazab M, Maddikunta PKR (2022) A review on deep learning for future smart cities. Internet Technol Lett 5(1):e187

4. Lakshman K, Kaluri R, Gundluru N, Alzamil Z, Rajput D, Khan A, Haq MA, Alhussen A (2022) A review on deep learning techniques for IoT data. Electronics 11:1604
5. Zikria Y, Afzal M, Kim S, Marin A, Guizani M (2020) Deep learning for intelligent IoT: opportunities, challenges and solutions. Comput Commun 164:50
6. Kant Singh K, Singh A, Lin J-W, Elngar AA (eds) (2021) Deep learning and IoT in healthcare systems: paradigms and applications, 1st edn. Apple Academic Press. https://doi.org/10.1201/9781003055082
7. Saleem TJ, Chishti MA (2021) Deep learning for the internet of things: potential benefits and use-cases. Digit Commun Netw 7(4):526–542
8. Hemmati A, Rahmani AM (2022) The internet of autonomous things applications: a taxonomy, technologies, and future directions. Internet Things 20:100635
9. Chen W, Qiu X, Cai T, Dai HN, Zheng Z, Zhang Y (2021) Deep reinforcement learning for internet of things: a comprehensive survey. IEEE Commun Surv Tutorials 23(3):1659–1692
10. Hosseinzadeh M, Hemmati A, Rahmani AM (2022) Federated learning-based IoT: a systematic literature review. Int J Commun Syst 35(11):e5185. https://doi.org/10.1002/dac.5185
11. Hosseinzadeh M, Hemmati A, Rahmani A (2022) 6G-enabled internet of things: vision, techniques, and open issues. Comput Model Eng Sci 133:509–556
12. Hosseinzadeh M, Hemmati A, Rahmani A (2022) Clustering for smart cities in the internet of things: a review. Clust Comput 25:1–31
13. Wang L, Ma C, Feng X et al (2024) A survey on large language model based autonomous agents. Front Comput Sci 18:186345. https://doi.org/10.1007/s11704-024-40231-1
14. Chang Y, Wang X, Wang J, Wu Y, Yang L, Zhu K, Chen H, Yi X, Wang C, Wang Y, Ye W, Zhang Y, Chang Y, Yu P, Yang Q, Xie X (2014) A survey on evaluation of large language models. ACM Trans Intell Syst Technol 15:1
15. Guo T, Chen X, Wang Y, Chang R, Pei S, Chawla NV, Wiest O, Zhang X (2024) Large language model based multi-agents: a survey of progress and challenges. arXiv preprint. arXiv:2402.01680
16 Thakur D, Saini JK, Srinivasan S (2023) DeepThink IoT: the strength of deep learning in internet of things. Artif Intell Rev 56(12):14663–14730
17. Heidari A, Navimipour NJ, Unal M (2022) Applications of ML/DL in the management of smart cities and societies based on new trends in information technologies: a systematic literature review. Sustain Cities Soc 85:104089
18 Amiri Z, Heidari A, Navimipour NJ, Esmaeilpour M, Yazdani Y (2024) The deep learning applications in IoT-based bio- and medical informatics: a systematic literature review. Neural Comput Appl 36(11):5757–5797
19 Thamilarasu G, Chawla S (2019) Towards deep-learning-driven intrusion detection for the internet of things. Sensors 19:1977. https://doi.org/10.3390/s19091977
20. Otoum Y, Liu D, Nayak A (2022) DL-IDS: a deep learning–based intrusion detection framework for securing IoT. Trans Emerg Telecommun Technol 33(3):e3803
21. Albishari M, Li M, Zhang R, Almosharea E (2023) Deep learning-based early stage detection (DL-ESD) for routing attacks in Internet of Things networks. J Supercomput 79(3):2626–2653
22. El Houda ZA, Brik B, Senouci SM (2022) A novel IoT-based explainable deep learning framework for intrusion detection systems. IEEE Internet Things M 5(2):20–23
23. Ishaque NBM, Florence SM (2022) Internet of things enabled waste detection and classification using optimal deep learning model. In: Sharma R, Sharma D (eds) New trends and applications in internet of things (IoT) and big data analytics. Springer International Publishing, Cham, pp 15–28
24. Qazi E-U-H, Imran M, Haider N, Shoaib M, Razzak I (2022) An intelligent and efficient network intrusion detection system using deep learning. Comput Electr Eng 99:107764
25. Nasir M, Javed AR, Tariq MA, Asim M, Baker T (2022) Feature engineering and deep learning-based intrusion detection framework for securing edge IoT. J Supercomput 78(6):8852–8866
26. Sriram S, Vinayakumar R, Alazab M, Soman KP (2020) Network flow based IoT botnet attack detection using deep learning. In: IEEE INFOCOM 2020 - IEEE conference on computer communications workshops (INFOCOM WKSHPS), Toronto, ON, Canada, pp 189–194. https://doi.org/10.1109/INFOCOMWKSHPS50562.2020.9162668
27. Saharkhizan M, Azmoodeh A, Dehghantanha A, Choo KKR, Parizi RM (2020) An ensemble of deep recurrent neural networks for detecting IoT cyber attacks using network traffic. IEEE Internet Things J 7(9):8852–8859
28 Ullah S, Khan MA, Ahmad J, Jamal SS, Huma ZE, Hassan MT, Pitropakis N, Arshad, Buchanan WJ (2022) HDL-IDS: a hybrid deep learning architecture for intrusion detection in the internet of vehicles. Sensors 22:1340. https://doi.org/10.3390/s22041340
29. Abusitta A, De Carvalho GHS, Wahab OA, Halabi T, Fung BCM, Mamoori SA (2023) Deep learning-enabled anomaly detection for IoT systems. Internet Things 21:100656
30. Dina AS, Siddique AB, Manivannan D (2023) A deep learning approach for intrusion detection in Internet of things using focal loss function. Internet Things 22:100699
31. Kumar P, Kumar R, Gupta GP, Tripathi R, Srivastava G (2022) P2TIF: a blockchain and deep learning framework for privacy-preserved threat intelligence in industrial IoT. IEEE Trans Industr Inf 18(9):6358–6367
32. Naeem H, Ullah F, Naeem MR, Khalid S, Vasan D, Jabbar S, Saeed S (2020) Malware detection in industrial internet of things based on hybrid image visualization and deep learning model. Ad Hoc Netw 105:102154
33. Yazdinejad A, Kazemi M, Parizi RM, Dehghantanha A, Karimipour H (2023) An ensemble deep learning model for cyber threat hunting in industrial internet of things. Digit Commun Netw 9(1):101–110
34. Wang X, Wang Y, Javaheri Z, Almutairi L, Moghadamnejad N, Younes OS (2023) Federated deep learning for anomaly detection in the internet of things. Comput Electr Eng 108:108651
35 Nazir A, He J, Zhu N, Qureshi SS, Qureshi SU, Ullah F, Wajahat A, Pathan MS (2024) A deep learning-based novel hybrid CNN-LSTM architecture for efficient detection of threats in the IoT ecosystem. Ain Shams Eng J 15:102777
36. Sharma B, Sharma L, Lal C, Roy S (2024) Explainable artificial intelligence for intrusion detection in IoT networks: a deep learning based approach. Expert Syst Appl 238:121751
37. Lilhore UK, Dalal S, Simaiya S (2024) A cognitive security framework for detecting intrusions in IoT and 5G utilizing deep learning. Comput Secur 136:103560
38. Joychandra Singh N, NazrulHoque K, Singh R, Bhattacharyya DK (2024) Botnet-based IoT network traffic analysis using deep learning. Secur Privacy 7(2):e355
39. Zhao R, Wang X, Xia J, Fan L (2020) Deep reinforcement learning based mobile edge computing for intelligent Internet of Things. Phys Commun 43:101184
40. Wang S, Tong S (2022) Analysis of high-level dance movements under deep learning and internet of things. J Supercomput 78(12):14294–14316
41. Chen M, Wenhu Du (2023) The predicting public sentiment evolution on public emergencies under deep learning and internet of things. J Supercomput 79(6):6452–6470
42. Belmonte-Fernández Ó, Sansano-Sansano E, Trilles S, Caballer-Miedes A (2022) A reactive architectural proposal for fog/edge

computing in the internet of things paradigm with application in deep learning. In: Pardalos PM, Rassia ST, Tsokas A (eds) Artificial intelligence, machine learning, and optimization tools for smart cities: designing for sustainability. Springer International Publishing, Cham, pp 155–175
43. Lv Z, Lou R (2022) Edge-fog-cloud secure storage with deep-learning-assisted digital twins. IEEE Internet Things M 5(2):36–40
44. Chen Y, Lin Q, Wei W, Ji J, Wong K-C, Coello CAC (2022) Intrusion detection using multi-objective evolutionary convolutional neural network for Internet of things in fog computing. Knowl-Based Syst 244:108505
45. Peláez-Rodríguez C, Pérez-Aracil J, De Lopez-Diz A, Casanova-Mateo C, Fister D, Jiménez-Fernández S, Salcedo-Sanz S (2023) Deep learning ensembles for accurate fog-related low-visibility events forecasting. Neurocomputing 549:126435
46. Xu Q, You Q, Gong Y, Yang X, Wang L (2024) RIS-assisted UAV-enabled green communications for industrial IoT exploiting deep learning. IEEE Internet Things J. https://doi.org/10.1109/JIOT.2024.3369687
47. Wang J, Dai B, Li Y, He Y, Sun Y, Shen W (2024) An intelligent Edge-IoT platform with deep learning for body condition scoring of dairy cow. IEEE Internet Things J 11(10):17453–17467. https://doi.org/10.1109/JIOT.2024.3357862
48 Xin Q, Alazab M, Díaz VG, Montenegro-Marin CE, Crespo RG (2022) A deep learning architecture for power management in smart cities. Energy Rep 8:1568–1577
49. Teng T, Ma Li (2022) Deep learning-based risk management of financial market in smart grid. Comput Electr Eng 99:107844
50. Li J, Zhou T (2023) Evolutionary multi-agent deep meta reinforcement learning method for swarm intelligence energy management of isolated multi-area microgrid with internet of things. IEEE Internet Things J 10(14):12923–12937
51. Tomazzoli C, Scannapieco S, Cristani M (2023) Internet of things and artificial intelligence enable energy efficiency. J Ambient Intell Humaniz Comput 14(5):4933–4954
52. Puri V, Jha S, Kumar R, Priyadarshini I, Hoang Son L, Abdel-Basset M, Elhoseny M, Viet Long H (2019) A hybrid artificial intelligence and internet of things model for generation of renewable resource of energy. IEEE Access 7:111181–111191
53. Raj RJS, Shobana SJ, Pustokhina IV, Pustokhin DA, Gupta D, Shankar K (2020) Optimal feature selection-based medical image classification using deep learning model in internet of medical things. IEEE Access 8:58006–58017
54. Zhou Z, Yu H, Shi H (2020) Human activity recognition based on improved bayesian convolution network to analyze health care data using wearable IoT device. IEEE Access 8:86411–86418
55. Vaiyapuri T, Lydia EL, Sikkandar MY, Díaz VG, Pustokhina IV, Pustokhin DA (2021) Internet of things and deep learning enabled elderly fall detection model for smart homecare. IEEE Access 9:113879–113888
56. Guan Y, Qiu Y, Tian C (2022) Trajectory planning in college football training using deep learning and the internet of things. J Supercomput 78(17):18616–18635
57. Sahu AK, Sharma S, Raja R (2022) Deep learning-based continuous authentication for an IoT-enabled healthcare service. Comput Electr Eng 99:107817
58. Jeba Sheela A, Gowthami M, Raj Kumar VS, Charles Prabu V, Queen Mary Vidya M (2022) A hybrid DL with the Internet of Things to monitor human activities using wearable sensors. Meas: Sensors 24:100496
59. Qiu H, Zheng Q, Memmi G, Lu J, Qiu M, Thuraisingham B (2021) Deep residual learning-based enhanced JPEG compression in the internet of things. IEEE Trans Industr Inf 17(3):2124–2133
60. Khanna A, Selvaraj P, Gupta D, Sheikh TH, Pareek PK, Shankar V (2023) Internet of things and deep learning enabled healthcare disease diagnosis using biomedical electrocardiogram signals. Exp Syst 40(4):e12864
61 Obayya M, Arasi MA, Almalki NS, Alotaibi SS, Sadig MA, Sayed A (2023) Internet of things-assisted smart skin cancer detection using metaheuristics with deep learning model. Cancers 15:5016. https://doi.org/10.3390/cancers15205016
62. Wang C, Ko YC (2023) Emotional representation of music in multi-source data by the Internet of Things and deep learning. J Supercomput 79(1):349–366
63 Kumar P, Kumar R, Gupta GP, Tripathi R, Jolfaei A, Najmul Islam AKM (2023) A blockchain-orchestrated deep learning approach for secure data transmission in IoT-enabled healthcare system. J Parallel Distrib Comput 172:69–83
64. Srivastava A, Neog S, Medhi K (2023) An efficient deep learning architecture for internet of medical things. In: 2023 4th international conference on computing and communication systems (I3CS), Shillong, India, pp 1–6. https://doi.org/10.1109/I3CS58314.2023.10127240
65. Mohamed Shakeel P, Aboobaider BBM, Salahuddin LB (2022) A deep learning-based cow behavior recognition scheme for improving cattle behavior modeling in smart farming. Internet Things 19:100539
66. Yang J, Zhang J, Wang H (2021) Urban traffic control in software defined internet of things via a multi-agent deep reinforcement learning approach. IEEE Trans Intell Transp Syst 22(6):3742–3754
67. Liu P, Wang J, Sangaiah AK, Xie Y, Yin X (2019) Analysis and prediction of water quality using LSTM deep neural networks in IoT environment. Sustainability 11:2058. https://doi.org/10.3390/su11072058
68 Contreras-Castillo J, Guerrero-Ibañez JA, Santana-Mancilla PC, Anido-Rifón L (2023) SAgric-IoT: an IoT-based platform and deep learning for greenhouse monitoring. Appl Sci 13:1961. https://doi.org/10.3390/app13031961
69. Singh SK, Jeong Y-S, Park JH (2020) A deep learning-based IoT-oriented infrastructure for secure smart City. Sustain Cities Soc 60:102252
70. Vinayakumar R, Alazab M, Srinivasan S, Pham QV, Padannayil SK, Simran K (2020) A visualized botnet detection system based deep learning for the internet of things networks of smart cities. IEEE Trans Ind Appl 56(4):4436–4456
71. Simla A J, Rekha C, Leo LM (2023) Agricultural intrusion detection (AID) based on the internet of things and deep learning with the enhanced lightweight M2M protocol. Soft Comput. https://doi.org/10.1007/s00500-023-07935-1
72. Rezaee K, Khosravi MR, Attar H, Menon VG, Khan MA, Issa H, Qi L (2023) IoMT-assisted medical vehicle routing based on uav-borne human crowd sensing and deep learning in smart cities. IEEE Internet Things J 10(21):18529–18536
73. Arepalli PG, JairamNaik K (2024) A deep learning-enabled IoT framework for early hypoxia detection in aqua water using light weight spatially shared attention-LSTM network. J Supercomput 80(2):2718–2747
74. Arepalli PG, JairamNaik K (2024) Water contamination analysis in IoT enabled aquaculture using deep learning based AODEGRU. Ecol Inform 79:102405
75. Xu Y, He H, Liu J, Shen Y, Taleb T, Shiratori N (2023) IDADET: iterative double-sided auction-based data-energy transaction ecosystem in internet of vehicles. IEEE Internet Things J 10(11):10113–10130
76. Ding X, Gan Q, Shaker MP (2023) Optimal management of parking lots as a big data for electric vehicles using internet of things and long-short term memory. Energy 268:126613
77. Rani P, Sharma R (2023) Intelligent transportation system for internet of vehicles based vehicular networks for smart cities. Comput Electr Eng 105:108543

78. Ijemaru GK, Ang L-M, Seng KP (2023) Swarm intelligence internet of vehicles approaches for opportunistic data collection and traffic engineering in smart city waste management. Sensors 23:2860. https://doi.org/10.3390/s23052860
79. Aminizadeh S, Heidari A, Dehghan M, Toumaj S, Rezaei M, Navimipour NJ, Stroppa F, Unal M (2024) Opportunities and challenges of artificial intelligence and distributed systems to improve the quality of healthcare service. Artif Intell Med 149:102779
80. Luong NC, Xiong Z, Wang P, Niyato D (2018) Optimal auction for edge computing resource management in mobile blockchain networks: a deep learning approach. In: 2018 IEEE international conference on communications (ICC), Kansas City, MO, USA, 2018, pp 1–6. https://doi.org/10.1109/ICC.2018.8422743
81 Heidari A, Navimipour NJ, Unal M, Zhang G (2023) Machine learning applications in internet-of-drones: systematic review, recent deployments, and open issues. ACM Comput Surv 55(12):247
82 Amiri Z, Heidari A, Darbandi M, Yazdani Y, Navimipour NJ, Esmaeilpour M, Sheykhi F, Unal M (2023) The personal health applications of machine learning techniques in the internet of behaviors. Sustainability 15:12406. https://doi.org/10.3390/su151612406
83. Diro AA, Chilamkurti N (2018) Distributed attack detection scheme using deep learning approach for Internet of Things. Fut Gener Comput Syst 82:761–768