Edge Intelligence for Empowering IoT-Based Healthcare Systems

Vahideh Hayyolalam, Moayad Aloqaily, Öznur Özkasap, and Mohsen Guizani

Abstract

The demand for real-time, affordable, and efficient smart healthcare services is increasing exponentially due to the technological revolution and burst of population. To meet the increasing demands on this critical infrastructure, there is a need for intelligent methods to cope with the existing related challenges. In this regard, edge computing technology can reduce latency and energy consumption by moving processes closer to the data sources in comparison to the traditional centralized cloud and IoT-based healthcare systems. In addition, by bringing automated insights into the smart healthcare systems, artificial intelligence (AI) provides the possibility of detecting and predicting high-risk diseases in advance, decreasing medical costs for patients, and offering efficient treatments. The objective of this article is to highlight the benefits of the adoption of the edge intelligent technology, along with the use of AI in smart healthcare systems. Moreover, a novel smart healthcare model is proposed to boost the utilization of AI and edge technology in smart healthcare systems. Additionally, we discuss potential challenges and future research directions arising when integrating these different technologies.

Introduction

Improved healthcare systems represent the cornerstone of a healthy society. Due to the rapid population growth and the rise of various diseases, there is a drastic surge in requests for robust and efficient healthcare systems. Also, with the technological evolution and the emergence of the Internet of Things (IoT) coupled with the improvement of next-generation wireless communications, the concept of smart healthcare or connected healthcare systems has been introduced as a developed version of conventional healthcare systems. In fact, smart healthcare refers to a health service that exploits technologies such as IoT, wearable devices, and advanced communication protocols to dynamically connect patients, care givers, and health institutions and transfer information among them [1]. Smart healthcare services provide the possibility of handling and responding to medical requests remotely in an intelligent way, which diminishes hospitalization remarkably and helps people and doctors to predict, detect, diagnose, and treat diseases intelligently. Also, smart healthcare is used to prevent and control the outbreak of contagious and infectious diseases such as the Ebola virus, Avian influenza, Chickungunya virus [2], and recently, the COVID-19 pandemic. Taking all the above mentioned issues into account, adoption of efficient smart healthcare services would certainly improve the health level of society.

In order to expand the healthcare facilities and satisfy the vast amount of users' requests, smart healthcare systems employ numerous smart devices and IoT technologies. Thus, a massive amount of heterogeneous medical records generated by smart devices need to be processed and evaluated accurately based on the target of the system [3]. Therefore, there is a need for effective communications among various entities of smart healthcare community and data centers to ensure low response time for emergency cases in real-time health needs, since the slow response time and high latency at data centers are critical risks in smart healthcare systems and can lead to the irreparable catastrophes [1]. Keeping these facts in mind, edge computing technology and artificial intelligence (AI) are the best choices for these issues. Specifically, adopting the combination of these two technologies can pave the way for solving various challenges in the scope of smart healthcare.

Edge computing technology moves some processes close to the data sources, which reduces the load of transmitted data conspicuously [4]. In fact, edge technology enables smart healthcare systems to conduct some processes and store some data near the end users rather than transferring all the records to the cloud data centers [5]. In this way, smart healthcare systems only need to send the results of processes and some raw data to the remote cloud data centers. Various research works, by utilizing edge technology in the field of smart healthcare, can conduct big data analysis, store and process sensitive medical data, diminish latency, and reduce energy consumption, cost, network congestion, and response time.

Moreover, by emulating human cognition in analyzing data through complex algorithms, AI can smooth the path of estimating the conclusions without the direct involvement of humans. In fact, AI is responsible for investigating relations between the treatment, prevention, and detection techniques of diseases based on the data collected from patients. In the edge context, the resource constraints necessitate innovative and lightweight methods to be able to execute them at the edge environment. Therefore, machine learning (ML) and deep learn-

Vahideh Hayyolalam and Öznur Özkasap are with Koç University; Moayad Aloqaily and Mohsen Guizani are with Qatar University.

ing (DL) as subsets of AI are mostly preferred in edge technology to train the system for learning, making decisions, and actuating. The integration of edge computing and AI introduces an edge intelligence concept, which leads to an emerging revolution in smart healthcare systems [6]. Researchers have proposed several smart healthcare systems adopting AI methods such as DL, deep neural network (DNN), reinforcement learning (RL), *K*-means, and other ML-based algorithms.

The contributions of this article can be summarized as follows:
- It classifies the current AI-assisted edge-based smart healthcare solutions via three viewpoints, namely, the application cases in IoT smart healthcare, the proper AI deployment location at the edge, and the most suitable type of adopted AI technique.
- It proposes a novel smart healthcare model leveraging AI and edge technology together to reduce the limitations of the existing models.
- It adopts the use of deep reinforcement learning (DRL). According to the number of vital sign types that need to be monitored and the amount of collected medical data, the proposed system is designed to distribute the DRL layers and processes optimally.
- It generalizes the design of the proposed DRL model in such a way that it can be used in parallel mode at the edge layer. This means that the processes with the same functionality are distributed among different devices or DRL layers to reduce response time, latency, and congestion.

The article is organized as follows. The next section discusses how edge computing affects smart healthcare systems. Then we state the AI implication in edge-based smart healthcare systems as well as the pros and challenges of integrating AI and edge technology. Following that, we introduce a new framework for smart healthcare exploiting AI and edge technology. Open issues and challenges are then discussed. The final section provides the conclusion of this research.

EMERGING EDGE TECHNOLOGY IN SMART HEALTHCARE

Many healthcare systems have adopted cloud computing to deliver affordable and scalable solutions in order to process and store massive amounts of data recorded via numerous biosensors. However, since cloud-oriented healthcare systems mostly include mobile devices, networks, and cloud servers, there are often long distances between the systems' units, which intensifies a high latency issue in these systems [5]. Thus, emergency and real-time medical demands cannot rely on cloud-based healthcare systems. Furthermore, the vast amounts of records produced by sensors need to regularly be transferred to the cloud for being processed and stored, which leads to high energy consumption and high cost. Also, most chronic patients require low-cost mobile environments, which is not supported by cloud-oriented solutions [3].

Edge computing as an additional layer between end users' devices and remote cloud data centers is an emerging solution to overcome cloud-based healthcare limitations. In fact, edge computing is
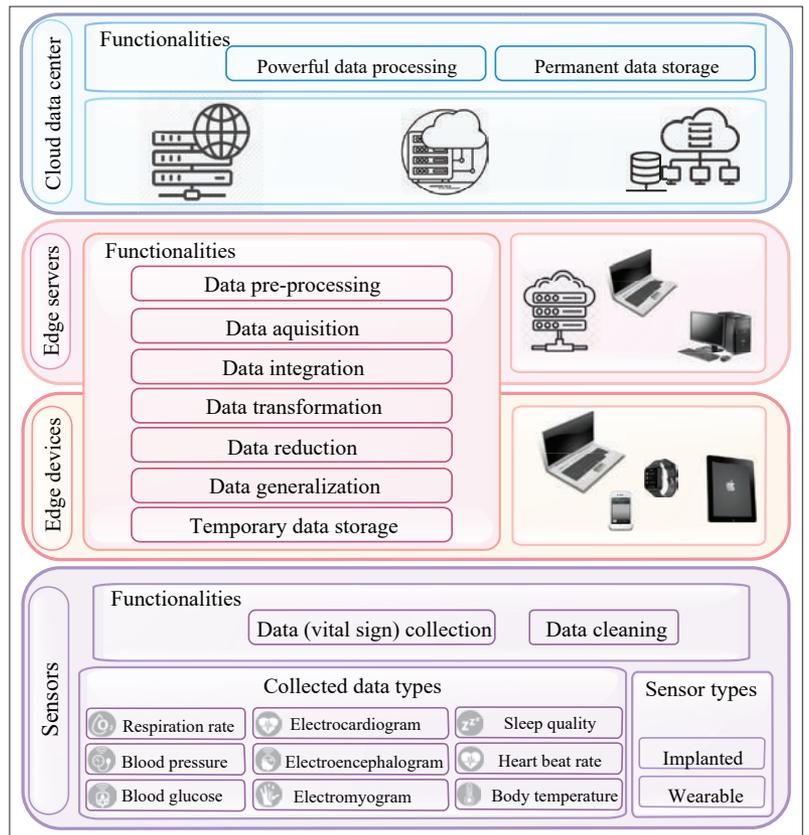

FIGURE 1. Edge-assisted smart healthcare functionality flow.

a smart gateway that plays the role of a bridge between devices and cloud data servers [7]. Edge-assisted solutions transfer processes close to the end users, which provides short response time and reduces energy usage. Indeed, edge technology enables smart healthcare systems to mine and process the collected data on the edge servers and devices near the user rather than transferring it to the cloud servers. Processing at the edge also assists and boosts security, privacy, mobility, low network bandwidth usage, geographical distribution, and location awareness, and facilitates online diagnosis and analytics, which reduces clinic and hospital visits [3]. Hence, edge technology contributes to the development of smart healthcare services by providing swift, more comprehensive, and ubiquitous treatments.

By and large, edge computing architecture in smart healthcare applications consists of four layers, including the sensors, edge devices, edge servers, and cloud data center layers. Figure 1 illustrates this architecture as a functionality flow, which indicates the functionality and the common devices used in each layer. The sensor layer contains different sensors, including implanted and wearable sensors that are responsible for collecting vital signs and data from patients. We indicate some data types that can be collected via sensors in a smart healthcare system. The sensor layer communicates with the edge devices layer via short-range and low-power wireless communication protocols such as ZigBee, Bluetooth, and WiFi [5]. The edge devices layer includes users' devices such as smartphones, Raspberry Pi, tablets, and smart watches, which are capable of conducting some processes and short-term data storage. The information
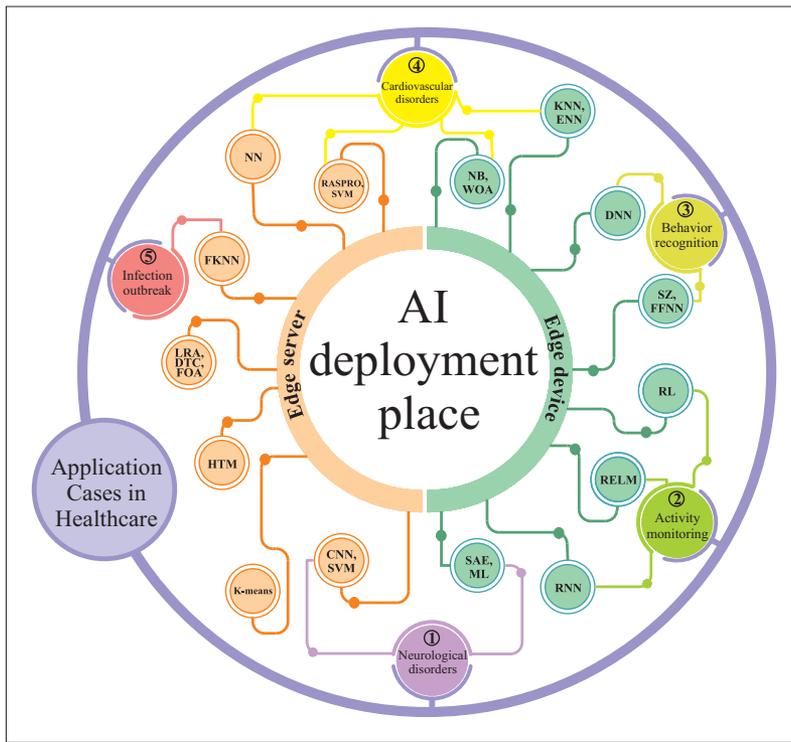
FIGURE 2. The classification of state-of-the-art solutions based on the AI deployment place and their application case.

is sent via Bluetooth/WiFi/Internet from the edge devices layer to the edge servers layer. The edge servers layer comprises micro data centers, and is able to conduct more processes and store more data in comparison to the edge devices layer [1]. As shown in Fig. 1, the functionality of the edge devices and edge servers layers are similar. The only difference is that the storage capacity and the computational power in the edge server layer is more powerful than in the edge device layer. The cloud data center layer encompasses more powerful processors and has the ability to store the vast amount of data. This layer communicates with the edge server layer via the Internet.

AI in Edge-Assisted Smart Healthcare Systems

This section first provides an overview and discusses the benefits of applying AI in edge-assisted smart healthcare systems. Then some related research works are examined in detail. Three different classifications of the investigated works are provided and illustrated in Figs. 2 and 3. Finally, a brief discussion along with Table 1, including the summary of inspected papers, are provided in this section.

The first classification considers the AI deployment place in edge computing to classify the papers. Regarding this, papers are classified into two groups: *edge device deployment* and *edge server deployment*. The second classification categorizes the research works based on their applied cases in smart healthcare into five groups (see the application case column in Table 1). The five mentioned groups include:

- *Neurological disorders* (e.g., epileptic seizure, chronic brain disorders)
- *Activity monitoring* (e.g., elderly monitoring, fitness tracking, emergency care, and tele-health advice)

- *Behavior recognition* (e.g., stress detection, feeling recognition)
- *Cardiovascular disorders* (e.g., heart problems, blood pressure)
- *Infection outbreak* (e.g., various viruses such as Chickungunya, COVID 19)

These two classifications are outlined in Fig. 2, which illustrates the adopted algorithms, their associated healthcare cases, and their deployment places at the edge. For instance, the DNN algorithm is employed for behavior recognition, which is deployed at an edge device. However, it can be seen that three papers (CNN, SVM [7], K-means [8], and HTM [9]) are not assigned to any of the application case groups, since their application healthcare cases have not been clarified by the authors. Moreover, Fig. 3 depicts the third categorization, which is based on the adopted AI technique, including three subcategories: *ML*, *DL*, and *hybrid* methods. The ML-based techniques include *supervised learning*, *unsupervised learning*, and *RL*; hybrid technologies encompass the hybrid of ML and DL, and the hybrid of ML and heuristic methods.

Overview and Benefits

The integration of AI and edge technology enables smart healthcare systems to run AI techniques locally on end-user devices where data are produced, while in traditional systems data processing procedures have been performed by AI techniques placed in the cloud far from end users. This integration provides the possibility of processing data on edge devices in a few milliseconds, which delivers real-time results. Since healthcare services are directly related to human health, supplying real-time services and results is crucial in critical conditions [10].

Researchers have recently focused on moving the inference portion of the AI workflow to the edge devices. The smart healthcare system can conduct the training portion in one environment and run the other parts in another place. In order to train ML/DL, the vast amount of data and powerful computing items are needed; thus, the cloud is more suitable for these AI techniques, whereas the inference part or running the trained models in novel records can be executed at edge devices. Moreover, some techniques, such as model compression methods by reducing large AI models into slight ones, can move some training to the edge devices gradually [5].

All in all, the adoption of AI techniques for processing medical data on the edge of the network can contribute to providing smarter healthcare services with prompt response, high privacy, high accuracy, improved reliability, enhanced bandwidth usage, low power consumption, and low latency. Also, it can reduce the burden on the network by diminishing the amount of data transmitted to the cloud servers [5].

Review and Classification

In this subsection, AI-based edge-assisted smart healthcare solutions are discussed in two groups according to the first classification mentioned earlier (Fig. 2).

AI in Edge Devices: Researchers in [6] have outlined a novel architecture for human activity recognition that utilizes Dockers technology on

edge devices accompanied by AI techniques. The proposed architecture consists of a sensor network, a central cloud server controlling the sensors, and edge computing as a bridge between sensors and cloud. The regularized extreme learning machine (RELM) is used in edge computing to classify the data collected from the sensors.

Also, in [1], the authors have proposed a context-aware edge-assisted smart healthcare framework by exploiting AI techniques. They have adopted DL for data compression in mobile edge devices to diminish bandwidth and energy usage. They have employed stacked auto-encoders (SAEs), which are a type of neural network and hierarchically extract data. Also, in order to guarantee high reliability and swift response in detecting or predicting disorders in epileptic seizure detection, the authors have adopted a data classification method leveraging ML algorithms.

The researchers in [11] proposed a human activity prediction system adopting edge technology and AI techniques. In the proposed system, data is collected by sensors and transferred to the edge device (laptop), where features are extracted from the collected data. Then, with regard to the features, a recurrent neural network (RNN) is trained on the edge device. The trained RNN is utilized for predicting human activities.

The authors of [12] proposed a smart healthcare system using edge technology and DL for real-time stress level detection by taking physical activity into account. The proposed system adopts wristband sensors to collect the required data. Also, DNN is utilized on edge devices. In the employed DNN, the input layer receives the collected data from sensors. Some hidden layers are responsible for data and stress level analysis, the output layer reports the detected stress level, and eventually, the results are stored in the cloud. In the proposed network, a supervised learning method is adopted for data training, the gradient descent algorithm is employed for optimizing the training algorithm, and logistic regression is used for classification of stress levels.

Also, the researchers in [13] designed an offloading model for edge-aided smart healthcare systems adopting AI methods to preserve user privacy in emergency care and telehealth advice application. The proposed model uses the RL-based offloading algorithm to protect the privacy of healthcare data during the processing procedure. The researchers in [14] designed a real-time elderly monitoring system on a heterogeneous multi-core edge device using AI techniques. The proposed system utilizes the shimmer device for data acquisition; then the collected data is compressed by a compressive sensing (CS) method. Next, the compressed data is transferred to the edge gateway through Bluetooth. The ODROID XU4 is used as an edge gateway where the compressed data is reconstructed and classified to identify various abnormal events.

An energy-aware data reduction method for edge-assisted smart healthcare systems adopting AI methods was proposed in [5]. This system uses a body area network (BAN), a wireless sensor network (WSN), and wearable devices to collect vital signs. The fast error bounded lossy compression pattern (SZ) is utilized for feature extraction in wearable devices. Afterward, the compressed data
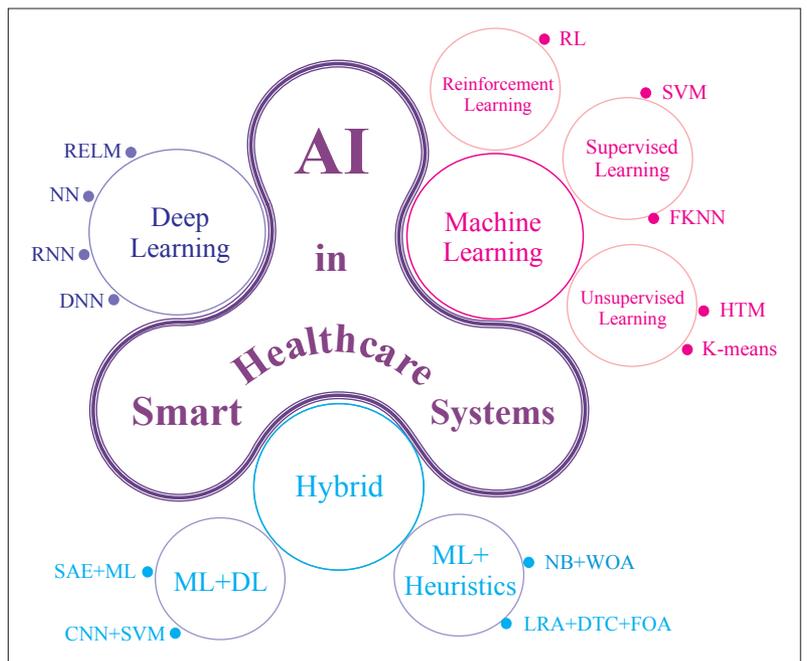

FIGURE 3. The classification of state-of-the-art solutions based on the adopted AI techniques.

are transferred to a local PC for further processing and analysis. The researchers deployed part of the DL on edge devices and the training network at the cloud. Moreover, a feedforward neural network (FFNN) is deployed in an edge device (PC) to classify the compressed data, which helps with detecting the stress level for the case study. The researchers in [10] proposed a real-time smart remote monitoring framework using edge technology and AI methods. The proposed framework consists of four layers, including the sensing layer for data collection, the edge device layer for offline pre-processing, the edge server layer, and the cloud layer for further online processes. The offline classification methods are trained in the cloud for predicting the health state of the patient in disconnected emergency events. The proposed method adopts the whale optimization algorithm (WOA) and naïve Bayes classifier (NB) for choosing the slight set of features with high accuracy level.

AI in Edge Servers: In [4], the authors proposed an edge-based smart healthcare system to detect severity and transmit alerts for cardiovascular diseases. Leveraging ML-based algorithms like rapid active summarization for effective prognosis (RASPRO) and support vector machine (SVM), the proposed system has conducted data compression, data fusion, and data classification in an Android smartphone as an edge device. Moreover, the authors in [2] have introduced a new model for controlling the outbreak of the Chickungunya virus. The proposed smart healthcare system leverages edge technology as a bridge between cloud storage and end users. BANs and smartphones are adopted to collect data and transfer them to the edge servers, where an ML-based fuzzy k-nearest neighbor (FKNN) algorithm classifies the data coming from edge devices. The results are sent to cloud data centers. Also, in [3], the authors proposed a new smart healthcare framework that utilizes edge technology and DL technique for automat-

Reference	Algorithm	AI category	AI contribution	AI deployment place	Healthcare application case	Main objective of the system	AI benefits
[6]	RELM	DL	Data classification	Edge device (Raspberry Pi)	Activity monitoring	Human activity detection	Low communication latency, swift decision making, and low cost
[1]	DL(SAE) +ML	Hybrid	Data compression (feature extraction), data classification	Edge device (smartphone)	Neurological disorders	Epileptic seizure detection	Low bandwidth usage, low energy usage, low response time, and high reliability
[11]	DL (RNN)	DL	Predicting the activities	Edge device (laptop)	Activity monitoring	Human activity prediction	Real-time support and high prediction rate (99.69 percent)
[12]	DL (DNN)	DL	Stress analysis	Edge device (smartphone)	Behavior recognition	Stress level detection	High detection accuracy (99.7 percent), rapid detection, and real-time supporting
[13]	RL	RL	Data privacy	Edge device	Activity monitoring	Data privacy	High privacy, low energy consumption, and low computational latency
[14]	CS, ENN/ KNN	Hybrid	Data compression, data classification	Sensors, edge gateway (ODROID XU4)	Cardiovascular disorders	Resource allocation	Low execution time, high accuracy, low latency, and real-time support
[5]	SZ algorithm, FFNN	Hybrid	Data compression, data classification	Edge devices (PC, wearable device)	Behavior recognition	Data reduction	High lifetime, low energy consumption, and high accuracy
[10]	NB, WOA	Hybrid	Feature selection and data classification	Edge devices	Cardiovascular disorders	Remote monitoring	Real-time support, big data analysis, and high accuracy
[4]	RASPRO, SVM	Supervised	Data compression, data fusion, and data classification	Edge gateway	Cardiovascular disorders	Severity detection and alerts transmission for cardiovascular diseases	Low bandwidth usage, high availability, and low cost
[2]	FKNN	Supervised	Data classification	Edge servers	Infection outbreak	Control the Chickungunya virus infection	Real-time support, low execution time, and low latency
[3]	NN	DL	Data classification	Worker node in Fog (Raspberry Pi)	Cardiovascular disorders	Automatic heart disease analysis	High prediction accuracy and low execution time
[7]	CNN, SVM	Hybrid	Feature extraction, data classification	Edge server	Neurological disorders	Data analysis	High accuracy, high precision, high sensitivity, and reduced network congestion
[9]	HTM	Unsupervised	Data anomaly detection	Edge server	Not mentioned	Data analysis	High accuracy
[8]	K-means	Unsupervised	Data dissemination	Edge server	Not mentioned	Data dissemination	High reliability, high accuracy
[15]	LRA, DTC, FOA	Hybrid	Resource allocation, classification of requests	Edge server (cloudlet)	Not mentioned	Resource allocation	Low energy consumption, low execution time, and low cost

TABLE 1. Summary of AI-based edge-assisted smart healthcare solutions.

ically analyzing and predicting cardiac diseases. This system delivers a smart healthcare service as a fog service. The proposed model includes BANs (sensors), edge gateways (smartphone), and a Fog-Bus module that consists of a broker node (Laptop), a worker node (Raspberry Pi), and a cloud data center (Microsoft Azure). The set partitioning in hierarchical trees (SPIHT) and singular value decomposition (SVD) techniques are adopted for data filtering in the broker node, and an ensemble DL technique is used for data classification in the worker node.

The researchers in [7] designed an automatic epileptic seizure prediction and localization system exploiting edge technology and AI techniques. They have proposed a novel multimodal data processing method on the mobile edge computing environment using convolutional deep learning (in a CNN) and wavelet transform for extracting high order features and a nonlinear SVM classifier along with generalized radial basis function (GRBF) kernel and long short-term memory (LSTM) DL for classifying the extracted electroencephalogram (EEG) and DL features.
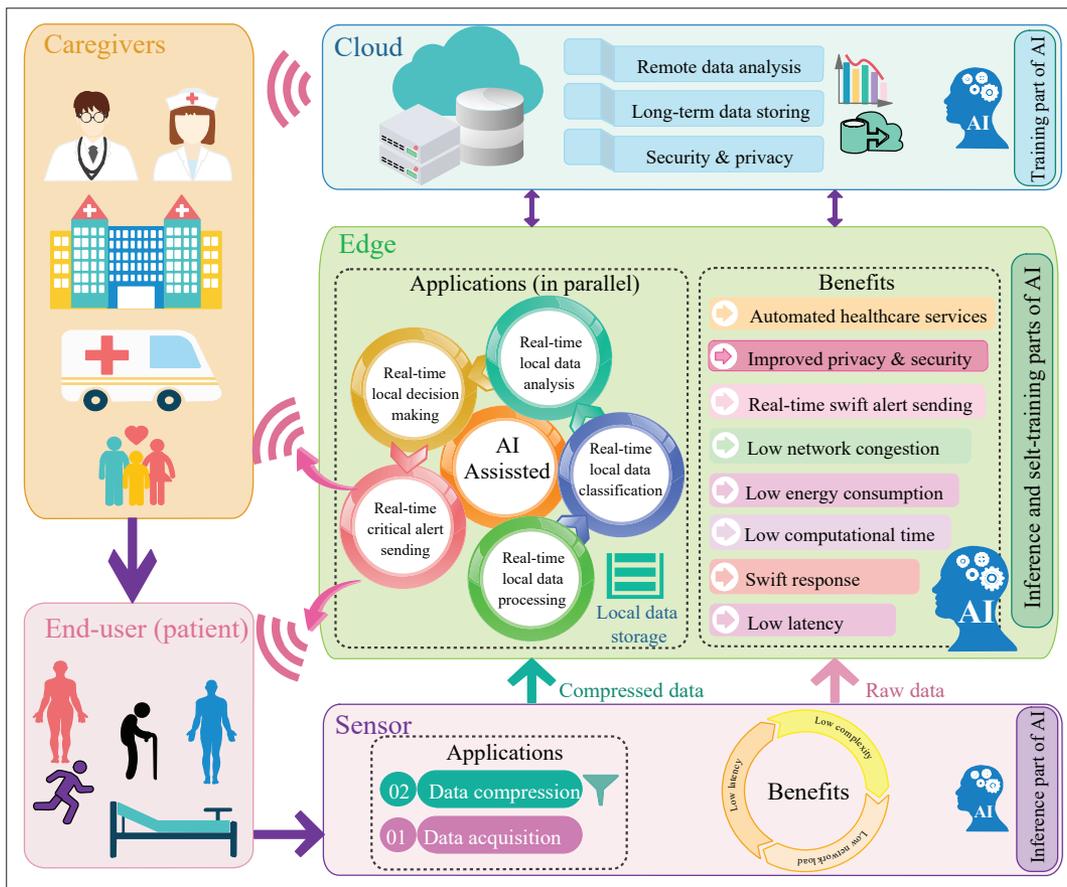
FIGURE 4. An overview of the system model.

All in all, the adoption of AI techniques for processing medical data on the edge of the network can contribute to providing smarter healthcare services with prompt response, high privacy, high accuracy, improved reliability, enhanced bandwidth usage, low power consumption and low latency. Also, it can reduce the burden on the network by diminishing the amount of transmitted data to the cloud servers.

Also, in [9], a real-time data analysis solution for medical data stream analysis was performed, which leveraged edge technology and AI methods for detecting data anomaly. The proposed architecture comprises four layers. The sensor layer fetches and dispatches medical data using wearable devices. The pre-process layer adopts Raspberry Pi to convert the collected raw medical data streams into resource description framework (RDF) streams. In the cluster processing layer, a hierarchical temporal memory (HTM) algorithm is employed to cope with the anomaly detection issue. The last layer, the persistent layer, stores the analyzed data to provide the possibility of further evaluation.

In [8], the authors designed an edge-based data dissemination model in which the mobile edge servers are utilized for disseminating various data segments. The proposed model includes the installation method and data propagation technique, which are responsible for ensuring the integrity of data objects and preventing the transmission collision, respectively. Moreover, in order to provide efficient data dissemination in heterogeneous edge networks, researchers used adaptive protocol and the *k*-means clustering algorithm in this work.

The researchers in [15] suggested a task offloading and resource allocation model for edge-assisted smart healthcare systems using AI methods. The authors considered edge computing with three layers: devices, edges, and clouds layers. Cloudlets with micro data centers are deployed at the edge of the network as edge servers. The proposed model uses a linear regression algorithm (LRA) at the edge servers to analyze the user requests and

a pipeline tree classifier (PTC) for clustering the jobs. Moreover, a fruitfly optimization algorithm (FOA) is used to optimize the resource allocation procedure.

DISCUSSION AND SUMMARY

With regard to the investigated papers, most of the researchers adopted AI on edge technology to manage the collected medical data. In most cases, data management contributes to reducing the amount of transmitted data. Since a lower amount of data needs to be conveyed to the cloud, cost, response time, latency, and network congestion will be decreased, bandwidth and energy usage will be diminished, and accuracy and reliability will be enhanced. Also, the smart healthcare system will be able to make decisions swiftly. The last column of Table 1 provides more information about AI benefits in edge-assisted smart healthcare solutions. Th ensuing section discusses a new system model that tries to provide a single system covering most of the mentioned AI contributions and benefits.

SYSTEM MODEL

This section describes a novel smart healthcare framework leveraging edge technology and employing different AI techniques like DRL. Figure 4 demonstrates the proposed system model with three main layers, including *sensor layer*, *edge layer*, and *cloud layer*. In this model, the end user can be a patient at a smart home being monitored remotely, a patient in a health institute, elderly people, or even a healthy person who cares for

The proposed system adopts different AI techniques, including DRL, distributed from sensors to the cloud, to facilitate smart healthcare services. Being flexible, the proposed framework can automatically define a proper solution map for the encountered issue. Based on the size of the faced problem and the volume of possible collected data, it can distribute the processes from sensors to cloud or from edge to cloud.

AI technique	Benefits	Drawbacks
ML	Improves over time, adapts without human intervention	Needs high amounts of time and resources for learning, high level of error susceptibility
RL	High learning speed, real-time response, dynamic, capable of learning from small data	Lack of clear interpretability, lack of integration of prior knowledge
DL	High-dimensional (with many layers), handles large amounts of data	Just provides approximate statistics, not accurate results, needs tons of data for training, low-speed evolution and data quality (medical data are highly ambiguous and heterogeneous; thus, the training phase may face some problems like data sparsity, redundancy, and missing value), transferring it to other problems is not easy
DRL	Attains complex smart healthcare states, best accuracy, powerful model, scalability, complex decision making,	Potential dead zones, may incur complexity states, off-policy learning, heavy computations, continuous data adjustment

TABLE 2. Merits and demerits of common AI techniques in the smart healthcare domain.

his/her fitness. Care givers can be a health professional, a health institute, a mobile ambulance, family members, or a drugstore.

The proposed system adopts different AI techniques, including DRL, distributed from sensors to the cloud, to facilitate smart healthcare services. Being flexible, the proposed framework can automatically define a proper solution map for the encountered issue. Based on the size of the problem faced and the volume of possible collected data, it can distribute the processes from sensors to cloud or from edge to cloud. Regarding this, various layers of DRL can be deployed on different layers of the proposed framework. As a case in point, if the purpose of the system is fitness tracking for a healthy person, there is no need to perform some of the processes on the cloud. Therefore, the system can accomplish processes on sensors and edge layers. The results can be stored either temporarily on the edge or permanently on the cloud according to the end user's request. The system is flexible enough to choose the appropriate place for hosting the input and output layers of DRL, depending on the current healthcare issue.

The sensor layer is typically responsible for collecting data from the end user and transferring the raw records to the edge layer, either to the edge devices or directly to the edge servers. We believe that compressing data in the sensor layer helps the system to reduce the overall latency, network congestion, and computation complexity on the edge layer. The data compression can be handled by adopting lightweight AI techniques such as CS on the sensors.

The edge layer consists of edge devices and edge servers hosting AI techniques. Many research attempts have been made in the field of smart healthcare leveraging AI and edge technology. However, there are still some unaddressed challenges. For example, DL needs a vast volume of data to train the algorithms. But, in some cases such as tele-surgery, there is not enough data available to train. Also, the trained data for a specific disease using DL is not adoptable for another similar illness; thus, the data should be re-trained. Moreover, RL, as an autonomous self-training system, should make several small decisions to achieve a sophisticated objective. In the case of complex decisions with many states, like monitoring a mobile patient with a chronic disease, which may have countless states due to the unstable context information of a moveable patient, the RL algorithm may not be able to handle the problem since learning from all states and regulating the reward path can be uncontrollable for the algorithm. A potential solution for these limitations can be an AI technique that uses both estimation and mapping the states together. In fact, the proposed framework can use a real-time self-training model or training from datasets depending on the encountered circumstances. This characteristic makes the system flexible for use in various issues based on the demands of end users, like remote monitoring, self-monitoring, and clinical monitoring. Table 2 summarizes the pros and cons of the common AI techniques utilized in the smart healthcare domain. Hence, considering the results of Table 2 and the discussion, DRL is seen as the best choice since it can cover the cases mentioned above. According to the issue size, the system can automatically decide to map all the states, use estimation, or both.

In the case of critical health issues, real-time alert sending and delivering health services in time is so crucial to save a person's life. Hence, reducing latency and time of decision making are vital in smart healthcare systems. DRL can be adopted in parallel on the edge layer by distributing data and processes on various devices or different hidden layers of DRL. It is crystal clear that the distributed parallel model contributes to more reduction of latency and swifter decision making in comparison to an unparalleled model. Since the processes are performed faster than the prior methods, the power consumption can also be diminished noticeably. Therefore, the lifetime of battery-powered devices will be increased.

The medical data, either compressed or raw, are transferred from sensors to the edge layer, where they are locally stored, classified, and analyzed adopting an AI technique in parallel mode to deliver real-time decisions for critical healthcare requests. Soon after making a decision, the corresponding alert would instantly be sent to the care givers, end user, or both. Thus, they can take action in time and prevent a possible catastrophe. Exploiting autonomous and smart self-training AI methods lying on edge technology, the proposed model can offer an automated smart healthcare system that can properly work even without human intervention.

Furthermore, considering the critical nature of IoT devices and medical data, which are vulnerable to attacks by hackers, security and privacy management are on the list of vital concerns in smart healthcare systems and should be taken into account. In our vision, security and privacy management are considered at different levels distributed from sensor to cloud layer adopting automated AI techniques like DRL. The proposed system model can cover crucial issues like network security, data security, data privacy, user privacy, communication privacy, encryption, and authentication of users to prevent illegal access to the system.

Eventually, the cloud layer performs remote and further data analysis and long-term data storage. With regard to the constrained resources of IoT and edge devices, in the case of too complex processes, the cloud can be the best place for conducting further analysis that is not possible to do on the edge. For instance, should the smart healthcare system need to train a vast amount of data, the cloud can be the best choice for this purpose.

In summary, the smart components of the proposed system model adopt sensors to collect data related to real-time patients' status while completely integrated with physical items. All the components are connected to a cloud-oriented system that collects and processes all the data produced by sensors. The proposed system model can be considered as a digital twin system that adopts real-world and real-time data to perform simulation leveraging edge intelligence and wireless communication, which can estimate how a process can be established to get better results in terms of predicting a disease in advance or saving lives in critical circumstances. Moreover, the proposed system model has reasonable complexity since there are no complex algorithms used in the model. The overall complexity of the proposed system model depends on the health issue and the amount of data. However, it cannot be greater than the complexity of DRL itself.

CHALLENGES AND FUTURE RESEARCH DIRECTIONS

According to our investigation, the ensuing open issues and challenges in the adoption of AI in edge-assisted smart healthcare solutions are detected. Future research directions in this area include both technical and non-technical challenges, summarized below. The first three challenges are considered as non-technical, and the rest are known to be technical challenges.

Sensor Usage: The most important issue in terms of sensor adoption is patients' comfort in using sensors. Sensors can be either implanted or wearable. Elderly people may have difficulties in using wearable ones. Some others may refuse to adopt the implanted sensors. This challenge also implies a personal comfort zone; in other words, what seems comfortable to one patient cannot be assumed to be the same for another.

Alert Receiving: The most salient concern in sending alerts is how reliable and responsive family members/trusted relatives/care givers are in reacting once they get an alert. For instance, they may not get the alert notification in time due to different reasons (e.g., a dead battery of their device, being busy with other tasks, being far from the patient). These issues may result in irreparable losses. Thus, fault tolerance measures need to be considered.

Patients' Concerns: Some patients are concerned about the treatment procedure and how they can trust a digital system for distinguishing or predicting their diseases. Thus, they usually resist using smart technologies, since they do not have enough information about how the system works, and how doctors interact with the system and indirectly with the patients.

Data Loss: Although efficient data compression contributes to reducing the amount of transmitted data, which may diminish the latency, power usage, and bandwidth usage, data integrity and information loss as a result of data compression is still a crucial open issue for smart healthcare solutions.

Lightweight AI Techniques: Since edge computing apparatuses, which are mostly used in smart healthcare applications, consist of constrained resources and limited power supply, there is a need for lightweight AI techniques that can be executed on them. Although AI techniques usually contribute to obtaining high accuracy, in some cases, it is achieved at the expense of augmented computational cost. Effective resource utilization may lead to the enhancement of lightweight AI techniques, which can help the smart healthcare system by analyzing enormous volumes of medical records produced continuously. RL techniques, in comparison to other AI methods, need lower memory resources and computational complexity. The RL techniques usually learn from their prior experiences; then they do not need training with datasets. Thus, using RL can lead to preserving computational resources and time as well, which are considered as the critical criteria in the healthcare systems.

Redundant Training Data: The DL techniques adopt a huge sample space of datasets for training. Each sample consists of various features that are adopted in the detection/prediction of diseases. Nevertheless, there is substantial redundancy in samples transferring almost identical information. Hence, the computational resources must work on processing similar information many times, which increases the time and computational complexity. These two factors are essential in the case of real time critical medical situations to save the patient's life. Therefore, proposing a strong and efficient algorithm for diminishing redundant training and providing effective resource utilization is still a challenging issue.

Privacy and Security: Since data is stored close to the end users thanks to the edge technology, individual users own the data entirely and can protect them. However, due to the constrained resources of users' devices, they cannot benefit from strong privacy and security protection algorithms. Thus, there is still considerable room for enhancing the privacy and security of smart healthcare systems leveraging edge technology and AI. A lightweight AI technique for coping with this issue can be a promising research area.

Automatic Network Management: With regard to the critical nature of smart healthcare systems, it is vital for these systems to have the capability of real-time and automatic management of network issues. The adoption of lightweight AI techniques such as lightweight RL can lead to smart and robust network traffic management in smart healthcare systems. Also, in

The proposed system model can be considered as a digital twin system that adopts real-world and real-time data to perform simulation leveraging edge intelligence and wireless communication, which can estimate how a process can be established to get better results in terms of predicting a disease in advance or saving lives in critical circumstances.

The proposed model considers security and privacy requirements for the entire smart healthcare system, including the sensitive medical data, the associated network connections, and users' access to the smart healthcare system. Although the proposed system model covers some limitations of prior smart healthcare frameworks, there are still other challenges that need to be addressed, comprising data loss and autonomous network management.

order to prevent security gaps and learn the network for autonomously diminishing any possible attacks, the smart healthcare system can utilize an ML-based intrusion detection system. Moreover, software-defined networking plays a crucial role in the implementation of autonomous network management. The use of learning algorithms can help to control the virtualization of network components and functions to deliver a self-regulated smart healthcare system.

CONCLUSION

This article illuminates how edge technology, along with AI techniques, can enhance smart healthcare systems. Processing locally, edge technology improves smart healthcare systems by reducing latency, network burden, and power consumption. Also, AI injects smartness into the system. The integration of AI and edge technology makes the components of smart healthcare systems smarter and brings many advantages to them. We introduce a novel smart healthcare framework that leverages different AI techniques in a parallel mode on edge technology. The proposed model distributes the processes from sensors to cloud servers. Conducting some processes using a lightweight AI technique on sensors results in diminishing the latency, complexity, and network load. Furthermore, adopting parallel processing on the edge layer employing a flexible AI method contributes to more reduction of latency, swift response and decision making, and sending real-time alerts to care givers, which are vital in smart healthcare services. The proposed model considers security and privacy requirements for the entire smart healthcare system, including the sensitive medical data, the associated network connections, and users' access to the smart healthcare system. Although the proposed system model covers some limitations of prior smart healthcare frameworks, there are still other challenges that need to be addressed, mainly comprising data loss and autonomous network management.

REFERENCES

[1] A. A. Abdellatif et al., "Edge Computing for Smart Health: Context-Aware Approaches, Opportunities, and Challenges," IEEE Network, vol. 33, no. 3, May/June 2019, pp. 196–203.
[2] S. Rani, S. H. Ahmed, and S. C. Shah, "Smart Health: A Novel Paradigm to Control the Chickungunya Virus," IEEE Internet of Things J., vol. 6, no. 2, 2018, pp. 1306–11.
[3] S. Tuli et al., "Healthfog: An Ensemble Deep Learning Based Smart Healthcare System for Automatic Diagnosis of Heart Diseases in Integrated IoT and Fog Computing Environments," Future Generation Computer Systems, vol. 104, 2020, pp. 187–200.
[4] R. K. Pathinarupothi, P. Durga, and E. S. Rangan, "IoT-Based Smart Edge for Global Health: Remote Monitoring with Severity Detection and Alerts Transmission," IEEE Internet of Things J., vol. 6, no. 2, 2018, pp. 2449–62.
[5] J. Azar et al., "An Energy Efficient Iot Data Compression Approach for Edge Machine Learning," Future Generation Computer Systems, vol. 96, 2019.
[6] M. Al-Rakhami et al., "A Lightweight and Cost Effective Edge Intelligence Architecture Based on Containerization Technology," World Wide Web, 2019, pp. 1–20.
[7] M.-P. Hosseini et al., "Multimodal Data Analysis of Epileptic EEG and Rs-Fmri via Deep Learning and Edge Computing," Artificial Intelligence in Medicine, vol. 104, 2020, p. 101,813.
[8] C. Shu et al., "Mobile Edge Aided Data Dissemination for Wireless Healthcare Systems," IEEE Trans. Computational Social Systems, vol. 6, no. 5, 2019, pp. 898–906.
[9] L. Greco, P. Ritrovato, and F. Xhafa, "An Edge-Stream Computing Infrastructure for Real-Time Analysis of Wearable Sensors Data," Future Generation Computer Systems, vol. 93, 2019, pp. 515–28.
[10] M. K. Hassan et al., "A Hybrid Real-Time Remote Monitoring Framework with Nb-Woa Algorithm for Patients with Chronic Diseases," Future Generation Computer Systems, vol. 93, 2019, pp. 77–95.
[11] M. Z. Uddin, "A Wearable Sensor-Based Activity Prediction System to Facilitate Edge Computing in Smart Healthcare System," J. Parallel and Distributed Computing, vol. 123, 2019, pp. 46–53.
[12] L. Rachakonda et al., "Stress-Lysis: A DNN-Integrated Edge Device for Stress Level Detection in the IOMT," IEEE Trans. Consumer Electronics, vol. 65, no. 4, 2019, pp. 474–83.
[13] M. Min et al., "Learning-Based Privacy-Aware Offloading for Healthcare Iot With Energy Harvesting," IEEE Internet of Things J., vol. 6, no. 3, 2018, pp. 4307–16.
[14] H. Djelouat et al., "Real-Time ECG Monitoring Using Compressive Sensing on a Heterogeneous Multicore Edge-Device," Microprocessors and Microsystems, vol. 72, 2020, p. 102,839.
[15] K. Lin et al., "Task Offloading and Resource Allocation for Edge-of-Things Computing on Smart Healthcare Systems," Computers & Electrical Engineering, vol. 72, 2018, pp. 348–60.

BIOGRAPHIES

VAHIDEH HAYYOLALAM received her B.S. degree in applied mathematics from Payam-Noor University of Tabriz, Iran, and her M.Sc. degree in computer engineering from the Science and Research Branch, Islamic Azad University, Iran, in 2016. Currently, she is a Ph.D. student at Koç University, Istanbul, Turkey. Her research interests include IoT, edge intelligence, and machine learning applications.

MOAYAD ALOQAILY [S'12, M'17] received his Ph.D. degree in electrical and computer engineering from the University of Ottawa, Ontario, Canada, in 2016. He was an instructor in the SYSC Department at Carleton University, Ottawa, Canada, in 2017. He was the managing director of xAnalytics Inc., Canada, in 2019. Currently, he is with the Faculty of Engineering, Al Ain University, United Arab Emirates. His current research interests include the applications of AI and ML, connected and autonomous vehicles, blockchain solutions, and sustainable energy and data management. He has chaired and co-chaired many IEEE conferences and workshops. He has served as a Guest Editor in many publications, including IEEE Wireless Communications, IEEE Network, and Computer Networks. He is an Associate Editor for Cluster Computing, Security and Privacy, IEEE Access, and IET Quantum Computing. He was appointed the Co-Editor-in-Chief of IEEE CommSoft TC eLetter in 2020. He started his own Special Interest Group (SIG) on Blockchain and Application as well as Internet of Unmanned Aerial Networks. He is an ACM member, and a Professional Engineer Ontario (P.Eng.).

ÖZNUR ÖZKASAP [SM'19] is a professor with the Department of Computer Engineering, Koç University, where she leads the Distributed Systems and Reliable Networks (DISNET) Research Laboratory. She received her Ph.D. degree in computer engineering from Ege University in 2000. She was a graduate research assistant with the Department of Computer Science, Cornell University, where she completed her Ph.D. dissertation. Her research interests include distributed systems, peer-to-peer systems, energy efficiency, mobile and vehicular ad hoc networks, and security in distributed systems. She serves as an Area Editor for Future Generation Computer Systems, Elsevier Science, and Cluster Computing (Springer). She is a recipient of the Career Award of TUBITAK (the Scientific and Technological Research Council of Turkey) and the Informatics Association of Turkey 2019 Prof. Aydın Köksal Computer Engineering Science Award.

MOHSEN GUIZANI [F] received his B.S. (Hons.) and M.S. degrees in electrical engineering, and M.S. and Ph.D. degrees in computer engineering from Syracuse University, New York, in 1984, 1986, 1987, and 1990, respectively. He served in different academic and administrative positions at the University of Idaho, Western Michigan University, the University of West Florida, the University of Missouri–Kansas City, the University of Colorado–Boulder, and Syracuse University. He is currently a professor with the Computer Science and Engineering Department, Qatar University. He is the author of nine books and more than 500 publications in refereed journals and conferences. His research interests include wireless communications and mobile computing, computer networks, mobile cloud computing, security, and smart grid. He is also a Senior Member of ACM. He also served as a member, the Chair, and the General Chair for a number of international conferences. Throughout his career, he received three teaching awards and four research awards.