Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Pattern Recognition Letters 135 (2020) 346–353

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

Trends in IoT based solutions for health care: Moving AI to the edge

Luca Greco*, Gennaro Percannella, Pierluigi Ritrovato, Francesco Tortorella, Mario Vento

Dept. of Computer and Electrical Engineering and Applied Mathematics (DIEM), University of Salerno, Italy

ARTICLE INFO

Article history:
Received 31 January 2020
Revised 28 April 2020
Accepted 11 May 2020
Available online 13 May 2020

MSC:
41A05
41A10
65D05
65D17

Keywords:
Atrificial intelligence
Internet of things
Healthcare
Edge computing
Fog computing

ABSTRACT

In recent times, we assist to an ever growing diffusion of smart medical sensors and Internet of things devices that are heavily changing the way healthcare is approached worldwide. In this context, a combination of Cloud and IoT architectures is often exploited to make smart healthcare systems capable of supporting near realtime applications when processing and performing Artificial Intelligence on the huge amount of data produced by wearable sensor networks. Anyway, the response time and the availability of cloud based systems, together with security and privacy, still represent critical issues that prevents *Internet of Medical Things* (IoMT) devices and architectures from being a reliable and effective solution to the aim. Lately, there is a growing interest towards architectures and approaches that exploit Edge and Fog computing as an answer to compensate the weaknesses of the cloud. In this paper, we propose a short review about the general use of IoT solutions in health care, starting from early health monitoring solutions from wearable sensors up to a discussion about the latest trends in fog/edge computing for smart health.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

In recent times, healthcare is increasingly leveraging information technologies for delivering smart systems aimed at speeding up health diagnostics and treatment. Such systems provide intelligent services for health monitoring and medical automation in different contexts and environments (hospitals, offices, home, on-the-go...), thus allowing a substantial reduction of physician visit costs and a general enhancement of patient care quality [3].

In this context, the wide diffusion of powerful embedded hardware together with the development of smart medical sensors and devices for ubiquitous healthcare has made the *Internet of Medical Things* (IoMT) drastically change the way healthcare is approached worldwide, so that the number of healthcare devices using IoT and wearable technologies is expected to reach 162 million by the end of 2020 [3].

Data captured by wearable, ingestible and embedded sensors, mobility patterns, device usage patterns allow to track user habits and can be effectively collected and processed to reveal critical conditions by using state of the art Artificial Intelligence (AI) and Machine/Deep Learning (ML/DL) based approaches. Traditional cloud based architectures for Big Data analysis are able to provide good performance and reliability when supporting non-safety and latency critical IoT applications [20,23,45]. But when the end user is a patient with critical and time-sensitive needs, a higher degree of robustness and accessibility is required, since the occurrence of disconnection from the core network or bandwidth/latency variations could have a dramatically negative impact and also lead to fatal consequences in emergency situations [4].

A growing interest towards architectures that realize the cooperation of Cloud, Fog and Edge computing is lately emerging. The main goal is to exploit the full potential of edge nodes and low level fog nodes to handle not only functional tasks but also data processing, analysis, correlation and inference [7].

Such approaches represent a promising answer towards the implementation of reliable distributed healthcare applications and services, since an intelligent mapping of computational and resource management tasks across the nodes proves to meet the stringent requirements of IoMT systems [4].

In this context, we assist to the spread of “Edge/Fog Health” solutions that make use of suitable computing paradigms to distribute health sensors data processing and storage among multiple nodes, located at different levels of proximity to users, as follows:

- *Edge computing* occurs directly on the devices to which the sensors are attached or a gateway device that is physically close to the sensors: examples of edge nodes are wearable devices such

* Corresponding author.
E-mail address: lgreco@unisa.it (L. Greco).

https://doi.org/10.1016/j.patrec.2020.05.016
0167-8655/© 2020 Elsevier B.V. All rights reserved.

as smartphones, smartwatches or portable "ad hoc" embedded systems like single-board computers, microcontrollers;
- *Fog computing* nodes act at a local area network level, so they can include bigger and more powerful devices such as PCs, local servers and gateways that may be physically more distant from the sensors and actuators.

Both paradigms (increasingly often implemented together) leverage the proximity to the user in order to provide location aware health services with reduced latency and high availability [12]. Some methods relying on hierarchical computing strategies have been proposed to allocate and distribute the inference tasks of AI and ML methods between the cloud, the fog and the edge levels (or fog/edge peers), trying to push the (limited) computational capacities of edge devices to their top [1,5,26,50].

A transition from the mobile cloud computing model (MCC), characterized by high data transmission costs and limited coverage towards a mobile edge computing model (MEC) [22] with low-latency and reliable edge ML models is then progressively happening in the smart healthcare domain.

In this paper, we propose a review about IoMT solutions, focusing on health monitoring. We aim to illustrate the evolution of IoT based health care systems, starting from early monitoring solutions, that involve edge nodes only for functional tasks, up to recent proposals that leverage cooperative edge/fog computing for smart health. We therefore discuss trends in Edge ML techniques and models.

2. A general architecture for IoMT systems

The use of Fog/Edge computing architectures in the health care domain typically deals with the design of remote monitoring solutions that leverage wearable and field sensor networks for implementing protective, preventive and responsive systems [3]. In this context, most contributions involve fog nodes acting as local servers that collect and process health data to quickly respond to the service requirements [49]. The scientific community has been investigating for more than ten years solutions for monitoring patients' health status with the aim of remotely providing reports to clinicians.

Early contributions, as in Chen et al. [10], propose mainly PC based systems to remotely monitor patients, for example, using data from ECG and accelerometers in order to advise clinicians about periods of elevated heart rate and filter expected critical situations. Subsequent implementations use micro-controllers and single board computers to gather physiological data from specialized sensors (ECG, E-health sensor platform) to be further processed and analyzed in a PC based environment [39,53].

In recent days, the maturity of IoT technologies has paved the way to the development of several smarter solutions exploiting both network architectures and software platforms domains. Such solutions aim to address healthcare at different levels: pediatric and elderly care, chronic disease supervision, epidemic disease monitoring, medical cyber-physical systems, private health and fitness management [27].

In this work we mainly focus on systems addressing health monitoring problems, that can be also split in two categories: static remote monitoring, where the patient is supposed to be in a building (home, hospital...), and dynamic monitoring, that assumes the patient to be monitored on-the-go.

Actually, a general solution employs a multilevel architecture [19], shown in a very simplified way in Fig. 1 that could involve:

- *an Edge level*, where portable devices (smartphones, smartwatches, compact embedded systems and compact gateways...) perform pre-processing and some low level elaborations on data collected from Wireless Body Sensor Networks (WBSN).
- *a Fog level*, where PCs, servers/gateways gather data from field sensor networks (and edge devices) to perform local processing and/or storage.
- *a Cloud level*, where cloud services are called for high performance computing tasks and remote data storage.

The three levels do not have to be all implemented. For example, in static monitoring problems, data from sensors could be directly collected by fog devices and be elaborated with optional support from cloud services. In the same way, in some dynamic monitoring scenarios where a fog level could not be implemented, edge devices would directly interact with cloud services.

3. Review of IoMT monitoring solutions

In less then ten years, several contributions presenting IoT architectures for smart health monitoring have been proposed. For our review, we selected a number of studies (reported in Table 1) that we found relevant to illustrate the evolution of IoT based health care systems, starting from early monitoring solutions, that involve edge nodes only for functional tasks, up to recent proposals that leverage cooperative edge/fog computing and allow for Edge ML.

The studies cover different application fields of smart health systems (e.g. monitoring of physiological parameters, speech, motion, posture, skin, environmental conditions and so on), and employ different sets of devices and sensors. We grouped the contributions according to the problems they aim to address.

3.1. Analysis of the physiological parameters

Several proposals focus on the development of systems aimed at sensing physiological parameters as health indicators to assess critical circumstances leading to accidents (protective systems).

A first interesting contribution is shown in Magaña Espinoza et al. [30] where a WBSN is used to monitor heart rate and motion rate of people within their homes. The edge node is connected to the internet and allows to send an alert (on smartphone) to family members or specialists whenever rapid changes in measured values occur (early detection of falls, tachycardia or bradycardia). In the same way, Villarrubia et al. [51] propose a system for tracking patients at home and monitoring their cardiac function by performing a basic analysis of ECG data. Patients can interact with the system by means of a common TV interface.

The work of [28] investigates the use of Bluemix cloud technology to store physiological data, allowing remote access by clinicians that can visualize results of their analysis by means of IBM Watson IoT platform, while [2] propose a case study of fever diagnosis by using an embedded system which continuously monitors the temperature of the patient.

In [46] a real time IoT based ECG telemetry is proposed. This is one of the first studies where quality assessment algorithms are implemented directly on an android smartphone allowing a real-time evaluation. The authors demonstrate the effectiveness of the method under different physical activities.

Static monitoring can leverage the use of field sensors that allow to gather contextual data for multimodal activity recognition tasks. An interesting study is proposed in Pham et al. [41] where environmental sensors, an optitrack camera, and smartwatch embedded sensors are used to collect video, motion and audio signals jointly with specialized wearables for physiological parameters gathering. This is actually a cooperative fog to cloud architecture, where data preprocessing, indoor localization and activity recognition algorithms are performed by means of a home gateway while a private cloud is used for storing data to be accessed remotely.

Table 1
Overview of the surveyed contributions.

Study	Objectives	Sensors	Fog/edge computing devices	Notes on methodology
Orha and Oniga [39]	Automatic recording of the main physiological parameters of the human body by the use of an Arduino microcontroller.	Specialized sensors	Arduino microcontroller, PC	Data transferred to a PC for processing.
Yakut et al. [53]	Measuring ECG signal by using E-Health Sensor Platform connected to a Raspberry Pi.	E-health sensor platform	Raspberry Pi	Raspberry Pi saves data to a text file to be further processed in a Matlab computer environment.
Magaña Espinoza et al. [30]	Detecting and alerting professionals about persons falling to the ground; event-based monitoring to report tachycardia and bradycardia	Heart rate sensor and triple axis accelerometer	LCP2148 ARM7 microcontroller	Encryption scheme for wireless sensors communication. Mobile application and web page for easy access and push notifications of abnormal events.
Villarrubia et al. [51]	Tracking and monitoring patients carrying a Holter within their homes	Triple axis accelerometer, ecg sensor	Raspberry Pi, Arduino	Multi-agent system based on the PANGEA platform. Indoor tracking by using accelerometers and wifi networks.
Kaur and Jasuja [28]	Monitoring pulse rate, body temperature	Temperature sensors and heart rate sensor	Raspberry Pi	Remote health monitoring by using Bluemix cloud.
Azimi et al. [4,5]	Classifying abnormalities in ECG signals	ECG sensors	Linux based PCs or GPU dedicated hardware (NVIDIA Jetson TK1)	Partitioning of a linear machine learning method (linear SVM) and distributed deploying of a deep learning algorithm
Alwan and Rao [2]	Measuring body temperature	Temperature sensors	Raspberry Pi, Arduino	Data exchanged between Arduino and Raspberry Pi units by means of ZigBee
Satija et al. [46]	ECG monitoring with signal quality assessment	ECG sensors	Arduino, Android phone	ECG signal quality assessment based on Discrete Fourier Transform based filtering. ECG signals collected during various physical activities.
Mathur et al. [32]	Monitoring temperature and gait to predict the health of the residual limb in lower limb amputee	Movement sensors, two thermistors	Raspberry Pi, Arduino uno, Android mobile device	Gaussian processes for machine learning used to predict the residual limb skin temperature of the amputee (MATLAB offline)
Muhammad et al. [36], [37]]	Voice pathology detection	Microphone, body temperature, electrocardiogram, ambient humidity	Smartphone	Data captured by IoT devices sent by bluetooth to the phone app. Feature extraction and classification (ELM) performed in cloud. Local binary pattern on a Mel-spectrum representation of the voice signal.
Dubey et al. [18]	Speech monitoring of patient with Parkinson's disease (PD) and ECG monitoring	Smartwatch microphone	Intel Edison	Dynamic Time Warping (DTW) is adopted for mining patterns in ECG time series. Clinical speech processing by means of average magnitude function for estimating pitch
Monteiro et al. [35]	Teletreatment of patients with Parkinson's disease through speech analysis	Smartwatch microphone	Intel Edison	Acoustic features extracted by Intel Edison are sent to the cloud for classification.
Pham et al. [41]	Collecting physiological, motion and audio signals for daily health monitoring at home	Environmental sensors (Passive infrared, grid-eye thermopile array), optitrack camera, wearable (ECG and breath - smart shirt), smartwatch sensors	Arduino Mega	Data from the environmental sensors and wearable sensors processed by a home gateway (pre-processing, indoor localization and activity recognition algorithms) Physiological data with contextual information sent to the private cloud for storing and local/remote access.

(*continued on next page*)

Table 1 (*continued*)

Study	Objectives	Sensors	Fog/edge computing devices	Notes on methodology
Sood and Mahajan [48]	Detecting and controlling the spread of Chikungunya virus	Wearable and environmental sensors	Not provided	Fog layer for realtime processing and analysis of data collected from sensors. Cloud services used for storing and deeper analysis.
Sareen et al. [47]	Preventing Zika virus outbreak	Mosquito sensors	Mobile phone, fog servers	Data collected and processed by Fog servers. Depth analysis performed in cloud.
Hegde et al. [24]	COVID-19 pre-screening, fever and cyanosis non-contact detection	Raspberry Pi Camera v2, FLIR Lepton 3.5 Radiometry Long-Wave Infrared Camera	Raspberry Pi 4, Google Coral USB accelerator	Real-time detection and segmentation of forehead and lip regions using PoseNet. Temperature estimation from infrared camera image and cyanosis assessed from lips' image in the visible spectrum.
Greco et al. [21]	Detecting anomalies in physiological parameters in real time	Accelerometers, gyroscopes and magnetometers	Raspberry Pi 3	Edge stream computing architecture with distributed implementation of HTM algorithm for anomaly detection
Abdellatif et al. [1]	Analysis of electroencephalography (EEG data)	EEG sensors	Not provided	Data compression achieved with stacked autoencoders; edge-based feature extraction (five discriminative frequency features are manually selected); event detection at the edge (simple classification rule compared with different ML techniques).
Yeh [56]	Securing IoT based communication through BSN architecture	Generic wearable sensor	Local processing unit (handheld device) and fog server	Robust crypto primitives to ensure transmission confidentiality and entity authentication among smart objects.
Uddin [50]	Wearable sensor based activity prediction	ECG sensor, magnetometer, accelerometer, gyroscope	Local fog server with GPU	Recurrent neural network LSTM used for activity recognition
Liu et al. [9]	Food recognition for dietary assessment	Smartphone camera	Smartphone	Image pre-processing and segmentation on mobile device. CNN classification on cloud.
Dai et al. [15]	On device inference app for skin cancer detection	Smartphone camera	Apple Iphone	CNN pre-trained model running on iOS device (Core ML framework) allowing to classify skin lesions without cloud support.
Queralta et al. [43]	Fall detection system with cardiovascular or diabetes monitoring	ECG, EEG, EMG, blood pressure	Edge gateway Raspberri Pi 3	LSTM RNN for fall detection implemented on Edge Node
Ram et al. [44]	Multimodal activity recognition	ECG sensors, accelerometers, gyroscopes magnetometers on chest, ankles and arms (movement tracking)	PC server	Random Forest and Support Vector Machines for activity prediction
Abdel-Basset et al. [6]	Decision making model for detecting and monitoring patients with type-2 diabetes	Blood pressure, heart rate, respiratory rate, motion activity, glucose recognition	not provided	Hybrid technique based on type-2 neutrosophic with VIKOR method for predicting type-2 diabetes risks
Devarajan et al. [17]	Monitoring, predicting and controlling the risk of remote diabetic patients in real-time based on physiological conditions	Glucose level, ECG, smartphone embedded for physical activities	Smartphone	J48Graft decision tree classifier to discover the risk level of a diabetic patient (hypoglycaemia, normal, pre-diabetes and hyperglycaemia) from blood glucose level, body weight, physical activities and diet.
Priyadarshini et al. [42]	Prediction of Stress Types, Diabetes and Hypertension Attacks	(simulated) EDA, HR, SpO2,temperature, 3-dimensional accelerometer data, PGC, DBP, 2-h serum insulin, body mass index (BMI), diabetes, pedigree function, SBP, DBP, total cholesterol (TC), HDL, LDL, PGC and HR	PC server	Deep learning model to detect a person's mental state with an early detection for type-2 diabetes from sensors' captured data

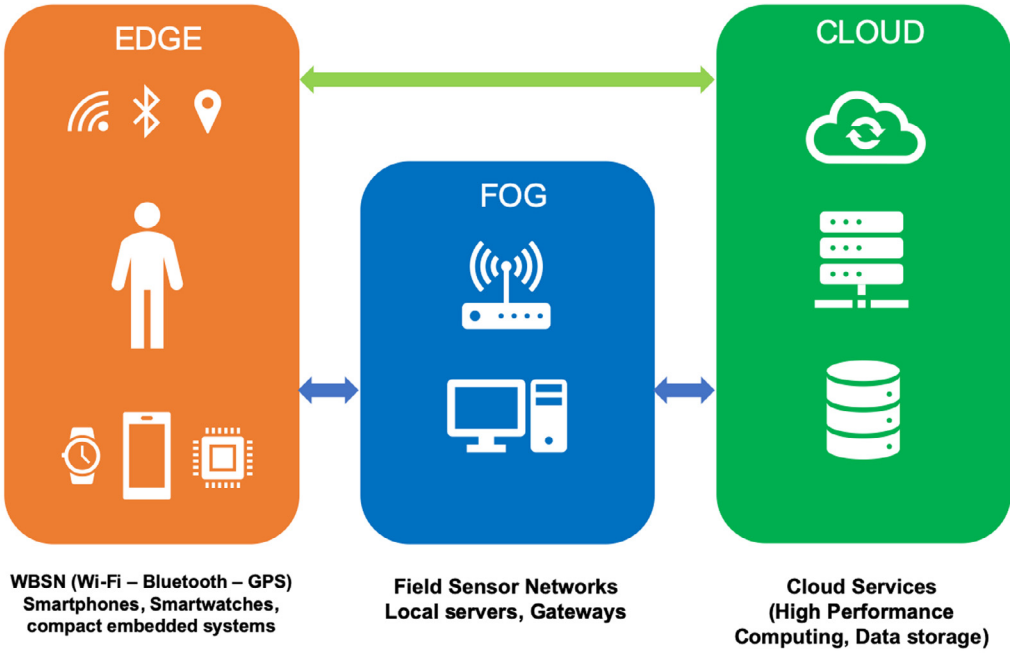

Fig. 1. A three tier based architecture for IoMT systems.

On the same line, fog architectures for activity recognition are presented in Uddin [50] and Ram et al. [44]. In particular, Uddin [50] present a solution for recognizing twelve different human activities using wearable sensors and a LSTM recurrent neural network running on local fog server with GPU acceleration, while [44] employ additional sensors for movement tracking and investigate the use of support vector machines (SVM) and random forest (RF) classifiers for activity prediction.

Recent proposals in physiological data analysis from wearable sensors start investigating the application of edge ML techniques. The work of [21] addresses the problem of detecting anomalies in physiological parameters by means of an edge stream computing architecture. Here, a distributed implementation of the HTM algorithm [38], running on edge nodes only, is adopted for the inference. Queralta et al. [43] also propose a fall detection solution relying on a LSTM recurrent neural network implemented at the edge level.

The work of [1] discusses the Multi Access Edge Computing approach with a case study on electroencephalography (EEG) data. The authors individuate the main functions to be implemented at the edge side to meet the application requirements (data compression, feature extraction and classification) and compare the accuracy of some traditional classifiers (Random Forests, Naive Bayes, k-Nearest Neigh-bors, and classification/regression trees).

Another interesting approach for classifying abnormalities in ECG signals has been proposed by Azimi et al. [4,5]. The solution, named Hierarchical Computing Architecture for Healthcare (HiCH), implements a variant of the MAPE-K model introduced by IBM [25] to distribute the computations through the three layers (edge, fog and cloud). The model is composed of four main computing elements: the *monitor* element that implements a bridge between the sensors and the edge, where pre-processing, aggregation and data storage take place; the *analyze* element that includes the cloud machines responsible for heavy computation tasks (model training); the *plan* element, located at the edge and periodically updated by the analyze component, is responsible for running the trained model and provide a local decision, thus conferring a high availability to the whole system. This approach has been tested in two cases: the partitioning of a linear machine learning method (linear SVM) [4] and the deploying of a deep learning (nonlinear) algorithm [5].

3.2. Rehabilitation systems

The use of protective IoMT solutions is effective also in post-operative situations where health tracking is aimed at detecting infections or complications in rehabilitation systems [3]. Mathur et al. [32] propose a solution to predict the health status of the residual limb in lower limb amputee by monitoring temperature and gait; here, a combination of edge devices (including an Android mobile) is used to capture and send data to a fog station that performs ML based prediction. In the same direction, Villeneuve et al. [52] contribute with a study on the estimation of simplified human limb kinematics based on measurements from two low-power accelerometers placed on the forearm.

Some proposals following the MCC paradigm have addressed voice pathology and speech monitoring. Muhammad et al. [36], [37]] discuss a voice pathology detection system that analyzes data gathered from smartphone microphone and wearable sensors and exploits an extreme learning based machine classification in the cloud. The works in Dubey et al. [18] and Monteiro et al. [35] use a fog architecture to implement a speech monitoring system for the tele-treatment of patients with Parkinson's disease. Audio signals are captured from a smartwatch microphone and sent to a fog node for acoustic feature extraction, while classification is performed in the cloud.

3.3. Skin pathologies and dietary assessment

With the diffusion of mobile deep learning frameworks, whose goal is to execute DL inference directly on smartphones [55], some interesting solutions are beginning to appear in the smart health care field. For example, Dai et al. [15] propose a solution for skin cancer detection that uses a pretrained CNN model running on a mobile device and performs a classification of skin lesions without cloud support. Liu et al. [9] discuss a solution where visual food recognition is performed for dietary assessment. In this work, im-

age pre-processing and segmentation is executed on device while CNN classification is done in the cloud.

3.4. Epidemic diseases treatment and location aware solutions

IoMT systems are becoming a viable solution also in diagnosis and treatment of epidemic diseases. In these contexts, real time operation, mobility, location awareness and smart data fusion from different sensors (environment, location, bio sensors...) are actually a matter. Such systems prove to be effective when diagnosing virus infections in an initial state so that proper treatment can be given on time to enable fast recovery. An interesting case is presented in Sood and Mahajan [48] where a solution to diagnose and prevent the outbreak of Chikungunya is discussed. A fog-based system analyzes user's health symptoms and surrounding environmental conditions for diagnosing the presence of the virus; it's also able to send alert messages to make users aware of infected and risk-prone areas. Sareen et al. [47] also propose a system to prevent and control the spread of Zika virus disease by implementing a MCC paradigm; the goal is to diagnose the possibly infected users and represent them together with mosquito-dense sites and breeding sites on the Google map to help the government healthcare authorities to control risk-prone areas. In both cases, a fog layer is used for processing and a preliminary analysis while cloud services are used for storage, deep analysis and remote access. The recent COVID-19 pandemic has raised a dramatic need for rapid screening tools to evaluate, in contactless conditions (at least 1 m away), the presence of the main symptoms of infection (fever, cyanosis, tiredness...), so that some IoT prototypes, based on low cost hardware, are beginning to appear. For instance, the AutoTriage system, presented in Hegde et al. [24], is able to run in real-time DL algorithms at the edge level to detect forehead and lip regions, allowing to estimate the temperature of the forehead-eye area by an infrared camera, while cyanosis is assessed from lips region in the visible spectrum. In a similar way, Maghdid et al. [31] propose a design study for predicting COVID-19 disease by applying multimodal AI approaches on smartphone embedded sensors readings. The possibility to share information about the location allows smart IoT devices to maximize the safety of their human owners. Al-Hamadi and Chen [11] discuss a trust-based approach for information sharing in a IoMT framework where data from sensors are used to derive a health loss risk for a user entering a given location at a given time, according to her vulnerabilities. The authors focus on an architecture where there is a centralized cloud collecting and analyzing sensing reports submitted by individual IoT devices.

IoT based health tracking is also increasingly adopted to support medical treatment and provide on-time medication or rehabilitation. In this field, the effects of edge to cloud computing have been evaluated by Masip-Bruin et al. [8] that discuss a breath assistance system for Chronic Obstructive Pulmonary Disease (COPD). In the proposed scenario, a COPD patient carries a portable oxygen concentrator (POC) with a smart mechanism aimed at adjusting and tailoring the oxygen doses to the patients' real-time context and needs. This would ensure a therapy tuned to patients' activity by continuously collecting and processing contextual information.

3.5. Diabetes treatment

An increasing number of IoMT solutions are also appearing in diabetes treatment where many commercial wearable and portable monitoring devices such as blood glucose monitors, insulin pens, continuous glucose monitors, insulin pumps, and closed-loop artificial pancreas systems are capable of wireless communication with smartphones or tablets, that act as edge nodes and provide basic analytic services without cloud support [29]. Several recent contributions in this field are focused on the definition of models to predict the onset of the disease. For example, Abdel-Basset et al. [6] discuss a novel technique to predict type-2 diabetes risks, while [17] use a J48Graft decision tree classifier, implemented on a smartphone, to discover the risk level of diabetic patients. Finally, a deep learning model to predict diabetes, stress types and hypertension attacks from wearable sensor data is discussed in Priyadarshini et al. [42], where a fog architecture is also proposed.

4. Discussion and trends

In the previous section, we have provided an overview of the most relevant IoT applications and solutions in literature that use a combination of edge, fog and cloud computing for addressing health care problems. The aforementioned studies have confirmed that a wide range of AI and ML based approaches can be effectively adopted for decision making, anomaly detection, predictive risk monitoring, treatment support and so on. Most cases still offload deep analysis tasks to the cloud, due to limited resources on edge low power devices, but an increasing trend towards solutions implementing a functional cooperation between edge, fog and cloud services can be actually observed. Many applications using ML/DL techniques such as multimodal activity prediction or recognition, pre-incident monitoring (fall detection and prevention), disease treatment (medical automation) and emergency care require the inference to be performed as close as possible to the sensor nodes to reduce latency, meet the requirement of realtime operation and minimize energy consumption. In static monitoring, fog architectures with GPU enhanced local servers seem to be the most adopted solution for addressing high demanding ML inference tasks. In dynamic monitoring problems, where a small sized low power computing node is employed, other optimization strategies come into play. On one hand, edge stream-computing techniques that aim at parallelizing the inference process between different edge peers can be a concrete answer; on the other, the adoption of DL lightweight models, optimized for running inference on embedded devices, allows to achieve very interesting results [14]. In this context, the advent of deep learning mobile frameworks has been accompanied by the proliferation of studies about strategies and techniques to transform and adapt DL models for running in constrained resource conditions. Such transformations range from model compression and quantization up to approximation and pruning aimed at obtaining networks with a very small footprint without penalizing performance [34]. A different matter concerns on-line or active learning systems, where training performed at the edge level could suffer from:

- *biased training samples* - performing at a local level, the edge device is able to capture polarized samples (about the same user or case) that can lead to overfitting when updating/re-training the model;
- *unsufficient computational power* - re-training can be computationally very expensive also if compressed/lightweight models are employed. This often leads to offloading the retraining stage to the cloud.

To address such issues, a number of interesting works discussing the suitability of AI methods for the IoT concept are being proposed [33]. Leveraging early studies about distributed deep networks training [16], some proposals investigate P2P approaches as in Chung and Yoo [13], where the authors discuss a distributed learning system with each edge node implementing a common neural network and exchanging weights with other peer nodes. On the same line, distributed approaches to deep networks training and collaborative machine learning at the edge level are also widely discussed by Park et al. [40], where both theoretical and technical enablers for decentralized edge training are investigated.

In addition to other features, these approaches can be beneficial for facing the problems deriving from changes in data distribution over time that, particularly in health applications, could negatively affect the performance of the whole system. Covariate shift, for example, can be efficiently faced when segmentation of deep networks on the edge [54] is introduced. In this case, in fact, only the device containing the first layers of the network has to be modified in order to correct the shift. We strongly expect these findings about *edge ML* to flourish in health care applications within the next few years, triggering further investigations in different settings.

5. Concluding remarks

IoT solutions for healthcare are evolving from simple architectures to collect, transmit and visualize data acquired by field and wearable sensor networks towards complex smart systems able to provide analytics, recognize activities and making decisions. In this transition, AI and machine learning techniques play an important role but their implementation require computational capacity that is often available only by means of cloud services. Actually, with the exponential growth of the data produced by the sensors, a cloud centric vision of the ML processing presents several weaknesses for different reasons. The quality of the service is strongly influenced by the quality of the internet connection, which also poses an availability problem. In particular, where a real time performance must be ensured, as in early-detection, risk prevention or activity recognition, the response time must be low to allow healthcare providers to proactively react to possible degradation of health conditions. Data storage and security are also critical issues when dealing with health related services, due to the large amount of personal data to be managed. After all, the choice of a fully local management is still impracticable due to limitations in processing power and storage, especially in the case of dynamic monitoring. Some proposals that try to map a proper computing model to the IoT tiers are being introduced to meet the requirements of IoMT systems. They aim at distributing the DL workload between the cloud and fog nodes, often relegating the model training to the cloud and moving the decision making to the edge. The interest towards these approaches is well motivated by the diffusion of compact GPU embedded hardware which allows the deployment of powerful and effective fog nodes. In cases where the DL training phase is still computationally unmanageable by the fog/edge node, DNN partitioning and distributed training represents a promising trend at present.

Declaration of Competing Interest

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

References

[1] A.A. Abdellatif, A. Mohamed, C.F. Chiasserini, M. Tlili, A. Erbad, Edge computing for smart health: context-aware approaches, opportunities, and challenges, IEEE Netw. 33 (3) (2019) 196–203.

[2] O.S. Alwan, K.P. Rao, Dedicated real-time monitoring system for health care using ZigBee, Healthc. Technol. Lett 4 (4) (2017) 142–144.

[3] A.O. Akmandor, N.K. Jha, Smart health care: an edge-side computing perspective, IEEE Consum. Electron. Mag. 7.1 (2017) 29–37.

[4] I. Azimi, A. Anzanpour, A.M. Rahmani, T. Pahikkala, M. Levorato, P. Liljeberg, N. Dutt, HiCH: hierarchical fog-assisted computing architecture for healthcare IoT, ACM Trans. Embed. Comput. Syst. (TECS) 16 (5s) (2017) 174.

[5] I. Azimi, J. Takalo-Mattila, A. Anzanpour, A.M. Rahmani, J.P. Soininen, P. Liljeberg, Empowering healthcare IoT systems with hierarchical edge-based deep learning, in: 2018 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE), IEEE, 2018, pp. 63–68.

[6] M. Abdel-Basset, G. Manogaran, A. Gamal, V. Chang, A novel intelligent medical decision support model based on soft computing and IoT, IEEE Internet Things J. (2019) 1–11.

[7] K. Bierzynski, A. Escobar, M. Eberl, Cloud, fog and edge: cooperation for the future? in: 2017 Second International Conference on Fog and Mobile Edge Computing (FMEC), IEEE, 2017.

[8] X. Masip-Bruin, E. Marin-Tordera, A. Alonso, J. Garcia, Fog-to-cloud computing (F2C): the key technology enabler for dependable e-health services deployment, in: 2016 Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net), 2016, pp. 1–5.

[9] C. Liu, Y. Cao, Y. Luo, G. Chen, V. Vokkarane, M. Yunsheng, P. Hou, A new deep learning-based food recognition system for dietary assessment on an edge computing service infrastructure, IEEE Trans. Serv. Comput. 11 (2) (2017) 249–261.

[10] C.-M. Chen, H. Agrawal, M. Cochinwala, D. Rosenblut, Stream query processing for healthcare bio-sensor applications, in: 20th International Conference on Data Engineering, IEEE, 2004.

[11] H. Al-Hamadi, I.R. Chen, Trust-based decision making for health IoT systems, IEEE Internet Things J. 4 (5) (Oct. 2017) 1408–1419.

[12] M. Chiang, T. Zhang, Fog and IoT: an overview of research opportunities, IEEE Internet Things J. 3 (6) (Dec. 2016) 854–864.

[13] K. Chung, H. Yoo, Edge computing health model using P2P-based deep neural networks, Peer-to-Peer Netw. Appl. 13 (2019) 694–703.

[14] S. Dey, J. Mondal, A. Mukherjee, Offloaded execution of deep learning inference at edge: challenges and insights, in: 2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), Kyoto, Japan, 2019.

[15] X. Dai, I. Spasi, B. Meyer, S. Chapman, F. Andres, Machine learning on mobile: an on-device inference app for skin cancer detection, in: 2019 Fourth International Conference on Fog and Mobile Edge Computing (FMEC), Rome, Italy, 2019, pp. 301–305.

[16] J. Dean, G. Corrado, R. Monga, K. Chen, M. Devin, M. Mao, Q.V. Le, Large scale distributed deep networks, in: Advances in Neural Information Processing Systems, 2012, pp. 1223–1231.

[17] M. Devarajan, V. Subramaniyaswamy, V. Vijayakumar, L. Ravi, Fog-assisted personalized healthcare-support system for remote patients with diabetes, J. Ambient Intell. Humaniz. Comput. 10 (2019) 3747–3760.

[18] H. Dubey, J. Yang, N. Constant, A.M. Amiri, Q. Yang, K. Makodiya, Fog data: fnhancing telehealth big data through fog computing, in: In Proceedings of the ASE Bigdata & Socialinformatics 2015, ACM, October 2015, p. 14.

[19] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, M. Ayyash, Internet of things: A survey on enabling technologies, protocols, and applications, IEEE Commun. Surv. Tutor. 17 (4) (2015) 2347–2376. Fourthquarter

[20] P. Gaba, R.S. Raw, Vehicular cloud and fog computing architecture, applications, services, and challenges, in: IoT and Cloud Computing Advancements in Vehicular Ad-Hoc Networks, IGI Global, 2020, pp. 268–296.

[21] L. Greco, P. Ritrovato, F. Xhafa, An edge-stream computing infrastructure for real-time analysis of wearable sensors data, Future Gener. Comput. Syst. 93 (2019) 515–528.

[22] M. Hartmann, U.S. Hashmi, A. Imran, Edge computing in smart health care systems: review, challenges, and research directions, Trans. Emerg. Telecommun. Technol. e3710 (2019).

[23] W. He, G. Yan, L.D. Xu, Developing vehicular data cloud services in the IoTenvironment, IEEE Trans. Ind. Inf. 10 (2) (May 2014) 1587–1595.

[24] C. Hegde, P.B. Suresha, J. Zelko, Z. Jiang, R. Kamaleswaran, M.A. Reyna, G.D. Clifford, Autotriage-an open source edge computing raspberry pi-based clinical screening system, 2020, In medRxiv. 10.1101/2020.04.09.20059840

[25] An architectural blueprint for autonomic computing, IBM White Paper 31.2006(2006) 1–6.

[26] W.N.H. Ibrahim, A. Selamat, O. Krejcar, J. CHAUDHRY, Recent advances on fog health - a systematic literature review, New Trends in Intelligent Software Methodologies, Tools and Techniques, IOS Press, 2018.

[27] S.M.R. Islam, D. Kwak, M.H. Kabir, M. Hossain, K.S. Kwak, The internet of things for health care: acomprehensive survey, IEEE Access 3 (2015) 678–708.

[28] A. Kaur, A. Jasuja, Health monitoring based on IoT using raspberry PI, in: 2017 International Conference on Computing, Communication and Automation (ICCCA), Greater Noida, 2017, pp. 1335–1340.

[29] D.C. Klonoff, Fog computing and edge computing architectures for processing data from diabetes devices connected to the medical internet of things, J. Diabetes Sci. Technol. 11 (4) (2017) 647–652.

[30] P. Magaña Espinoza, R. Aquino-Santos, N. Cárdenas-Benitez, J. Aguilar-Velasco, C. Buenrostro-Segura, A. Edwards-Block, et al., WiSPH: a wireless sensor network-based home care monitoring system, Sensors 14 (4) (2014) 7096–7119.

[31] H.S. Maghdid, K.Z. Ghafoor, A.S. Sadiq, K. Curran, K. Rabie, A novel ai-enabled framework to diagnose coronavirus covid-19 using smartphone embedded sensors: Design study, 2020, ArXiv preprint arXiv:2003.07434.

[32] N. Mathur, G. Paul, J. Irvine, M. Abuhelala, A. Buis, I. Glesk, A practical design and implementation of a low cost platform for remote monitoring of lower limb health of amputees in the developing world, IEEE Access 4 (2016) 7440–7451.

[33] A. Poniszewska-Maranda, D. Kaczmarek, N. Kryvinska, F. Xhafa, Studying usability of AI in the IoTsystems/paradigm through embedding NN techniques into mobile smart service system, Computing 101 (11) (2019) 1661–1685.

[34] P. Molchanov, S. Tyree, T. Karras, T. Aila, J. Kautz, Pruning convolutional neural networks for resource efficient inference, 2016, ArXiv preprint arXiv:1611.06440.

[35] A. Monteiro, H. Dubey, L. Mahler, Q. Yang, K. Mankodiya, Fit: a fog computing device for speech tele-treatments, in: 2016 IEEE International Conference on Smart Computing (SMARTCOMP), IEEE, May 2016, pp. 1–3.
[36] G. Muhammad, S.M.M. Rahman, A. Alelaiwi, A. Alamri, Smart health solution integrating IoT and cloud: acase study of voice pathology monitoring, IEEE Commun. Mag. 55 (1) (January 2017) 69–73.
[37] G. Muhammad, M.F. Alhamid, M. Alsulaiman, B. Gupta, Edge computing with cloud for voice disorder assessment and treatment, IEEE Commun. Mag. 56 (4) (2018) 60–65.
[38] N. Community, Introduction to HTM, 2018. https://numenta.org.
[39] I. Orha, S. Oniga, Automated system for evaluating health status, design and technology in electronic packaging (SIITME), in: IEEE 19th International Symposium for, 2013, pp. 219–222.
[40] J. Park, S. Samarakoon, M. Bennis, M. Debbah, Wireless network intelligence at the edge, Proc. IEEE 107 (11) (2019) 2204–2239.
[41] M. Pham, Y. Mengistu, H. Do, W. Sheng, Delivering home healthcare through a cloud-based smart home environment (coSHE), Future Gener. Comput. Syst. 81 (2018) 129–140.
[42] R. Priyadarshini, R. Barik, H. Dubey, Deepfog: fog computing-based deep neural architecture for prediction of stress types, diabetes and hypertension attacks, Computation 6 (4) (2018) 62.
[43] J.P. Queralta, T.N. Gia, H. Tenhunen, T. Westerlund, Edge-AI in lora-based health monitoring: Fall detection system with fog computing and LSTM recurrent neural networks, in: In 2019 42nd International Conference on Telecommunications and Signal Processing (TSP), IEEE, July 2019, pp. 601–604.
[44] S.S. Ram, B. Apduhan, N. Shiratori, A machine learning framework for edge computing to improve prediction accuracy in mobile health monitoring, in: In International Conference on Computational Science and Its Applications, Springer, Cham, July 2019, pp. 417–431.
[45] B.B.P. Rao, P. Saluia, N. Sharma, A. Mittal, S.V. Sharma, Cloud computing for internet of things & sensing based applications, in: In Sixth International Conference on Sensing Technology (ICST), Kolkata, 2012, pp. 374–380.
[46] U. Satija, B. Ramkumar, M.S. Manikandan, Real-time signal quality-aware ECG telemetry system for IoT-based health care monitoring, IEEE Internet Things J. 4 (3) (June 2017) 815–823.
[47] S. Sareen, S.K. Gupta, S.K. Sood, An intelligent and secure system for predicting and preventing Zika virus outbreak using fog computing, Enterp. Inf. Syst. 11 (9) (2017) 1436–1456.
[48] S.K. Sood, I. Mahajan, A fog-based healthcare framework for chikungunya, IEEE Internet Things J. 5 (2) (2017) 794–801.
[49] W. Tang, K. Zhang, D. Zhang, J. Ren, Y. Zhang, X.S. Shen, Fog-enabled smart health: toward cooperative and secure healthcare service provision, IEEE Commun. Mag. 57 (5) (May 2019) 42–48.
[50] M.Z. Uddin, A wearable sensor-based activity prediction system to facilitate edge computing in smart healthcare system, J. Parallel Distrib. Comput. 123 (2019) 46–53.
[51] G. Villarrubia, J. Bajo, D. Paz, F. Juan, J.M. Corchado, Monitoring and detection platform to prevent anomalous situations in home care, Sensors 14 (6) (2014) 9900–9921.
[52] E. Villeneuve, W. Harwin, W. Holderbaum, B. Janko, R.S. Sherratt, Reconstruction of angular kinematics from wrist-worn inertial sensor data for smart home healthcare, IEEE Access (2016) 2351–2363.
[53] O. Yakut, S. Solak, E.D. Bolat, Measuring ECG signal using e-health sensor platform, in: International Conference on Chemistry, Biomedical and Environment Engineering (ICCBEE'14), 2014, pp. 65–69.
[54] X. Wang, Y. Han, V.C.M. Leung, D. Niyato, X. Yan, X. Chen, Convergence of edge computing and deep learning: acomprehensive survey, IEEE Commun. Surv. Tutor. (Early Access) (2020) 1–36.
[55] M. Xu, J. Liu, Y. Liu, F.X. Lin, Y. Liu, X. Liu, A first look at deep learning apps on smartphones, in: In The World Wide Web Conference, ACM, May 2019, pp. 2125–2136.
[56] K.H. Yeh, A secure IoT-based healthcare system with body sensor networks, IEEE Access 4 (2016) 10288–10299.