Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 232 (2024) 397–405

Procedia Computer Science

www.elsevier.com/locate/procedia

5th International Conference on Industry 4.0 and Smart Manufacturing

Edge Intelligence for Industrial IoT: Opportunities and Limitations

Claudio Savaglio[a,*], Pasquale Mazzei[b], Giancarlo Fortino[a]

[a]*University of Calabria, DIMES Dept., Rende (CS), 87036 Italy*
[b]*Neosperience S.P.A., Rende (CS), 87036 Italy*

Abstract

Industrial Internet of Things (IIoT) nowadays represents a fundamental technology for the Industry 4.0 initiatives for connecting devices, operators and processes so as to eventually fuel digital transformation. However, if based exclusively on Cloud Computing, the IIoT faces inevitable limitations and barriers. Therefore, in this paper we deal with the Edge Intelligence (EI), a novel paradigm at the confluence of Edge Computing and Artificial Intelligence, and, particularly, we analyze its impact on the processing of the IIoT data in the proximity of the events of interest rather than on Cloud servers. Therefore, a general discussion on the most relevant benefits, limitations and open challenges of the IIoT-EI duo is proposed hereinafter and exemplified through an emblematic IIoT use case: this is related to safety-critical tasks based on video analytics and provides a full-fledged comparison between Cloud- and EI-based IIoT deployments (in terms of reliability, responsiveness, bandwidth usage, energy footprint and usability), finally outlining a general strategy for large-scale EI-aided IIoT systems engineering.

Peer-review under responsibility of the scientific committee of the 5th International Conference on Industry 4.0 and Smart Manufacturing

Keywords: Industrial Internet of Things; Edge Intelligence; Cloud Computing; Amazon AWS

1. Introduction

Internet of Things (IoT) empowers today factories to become more efficient, agile, and adaptive by enabling data-driven decision-making, automation, and optimization across the entire manufacturing ecosystem [1]. In particular, IoT technology enables the connection and communication between various sensors, actuators and processing units within a factory environment thus enabling, just to name a few, precise monitoring and predictive maintenance, efficient inventory management and asset tracking, integrated quality control and trust evaluation [2, 5]. Beside improving productivity and operational efficiency by streamlining the whole supply chain, wearable IoT devices (whose adoption is continuously increasing in work places and, in some cases, mandatory) can ensure worker safety by promptly analyzing movements, monitoring vital signs, and detecting potential safety hazards [3].

* Corresponding author.
E-mail address: csavaglio@dimes.unical.it

1877-0509
Peer-review under responsibility of the scientific committee of the 5th International Conference on Industry 4.0 and Smart Manufacturing
10.1016/j.procs.2024.01.039

The success gained by the IoT in making the factories "smart" gave origin to the Industrial IoT (IIoT) [4] whose impact spans across various industries, including manufacturing, energy, healthcare, transportation, and agriculture. Therefore, there emerged several IoT platforms that cater specifically to industrial applications and provide the necessary tools and infrastructure to connect, manage, and analyze IoT devices and data in an industrial environment [6] (e.g., IBM Watson IoT Platform, Siemens MindSphere, PTC ThingWorx,Cisco IoT Cloud Connect). Whereas featured by different levels of scalability, elasticity, integration with external systems as well as different pricing options (which all represent the drivers for their selection), the majority of these IoT Platforms leverage Cloud Computing infrastructure to store and process large volumes of data, being hence identified as Cloud-based IoT Platforms. If, on the one hand, the abundance of resources available on the Cloud allows potentially unlimited scalability and flexibility without investment on physical asset purchase and management, on the other hand the Cloud Computing exposes well-known intrinsic limitations (e.g., dependence on Internet connectivity, bandwidth consumption, data sovereignty and compliance, privacy concerns) which are even more relevant in the industrial domain, especially in safety-critical cases [7]. For example, IoT devices can collect data from machinery's operation, forward them to the Cloud for the execution of anomaly detection, and receive actuation commands in the case of dangerous events: such a kind of workflow may involve sensitive industrial data and it is safety-critical, but it is also exposed to cybersecurity risks and prone to fail if unpredictable events (e.g., a fire or an earthquake) affect the reachability of remote server.

From here it emerges the need of a paradigm shift (pushed also by other application domains, for example intelligent transport system, e-Health, entertainment) for moving computation from remote sites close to the data source rather than forwarding local data towards Internet. The Edge Intelligence (EI) [7] [14], also referred as Edge AI, embodies this principle and brings together Edge Computing infrastructures with approaches, models and techniques of Artificial Intelligence, Data Science, Distributed and Pervasive Computing, in order to enable intelligence at the edge of the network. Rapidly gaining traction as also demonstrated by the increasing number of both research papers and EI-enabled IoT Platforms [10] (e.g., Amazon AWS, Crosser, Azure, Bosch), the EI is a very promising and already mainstream, despite its young age, approach but its adoption is a strategical choice which needs to be carefully and comprehensively assessed.

Along such a line, in this paper we discuss opportunities, strengths and weaknesses about the EI adoption in the IIoT domain and we present a use case in which the same scenario is developed according to both Cloud- and EI-based deployment, being both quantitatively as well as qualitatively assessed. Such a comprehensive comparative analysis is typically missing in the state-of-the-art where the focus is typically either on theoretical discussion or on the implementation of "stand-alone" EI solutions (algorithms, models, techniques), which are detached from any real deployment and aimed only to shown their feasibility over resource constrained IIoT devices [11], [12]. Instead, we consider a full-fledged scenario and we finally outline a strategy that, starting from the preliminary yet interesting results obtained, aims to outline the optimal configuration for the IoT system. The rest of the paper is organized as follows: in Section 2.1 the foundations, target domains and research directions of EI are introduced, while a focus on its application in the IIoT arena is provided in Section 2.2 and some related works reported in Section 2.3. Then, the developed use case pivoted around video analytics tasks is shown and discussed in both its deployment settings in Section 3 and a general strategy for large-scale EI-aided IIoT systems engineering is outlined in Section 4. Final remarks and considerations conclude the paper in Section 5.

2. Edge Intelligence and IIoT

2.1. EI foundations, application domains and research directions

According to recent estimates, the sum of present IoT devices generate more than 90 Zettabytes of data at the network edge [8]: to keep up with both the increasing amount and speed of data as well a to support the new ambitious services and scenarios of the IoT, the EI paradigm emerged at the confluence of Edge Computing and AI [14]. A formal definition of EI is still missing [7] but the researchers agree with its goal, namely to prioritize local operations on edge devices over centralized and remote data centers. Such a truly distributed and pervasive approach allows for the local processing of data by re-engineering, combining techniques from Cloud/Edge/Distributed/Pervasive Computing AI, data science and customizing them according to the actual available resources of edge devices.

Regardless the proliferation of EI chips and processors, ML libraries, and frameworks, a reference architecture specifically designed for EI is still missing. The Multi-access edge system promoted by European Telecommunications Standards Institute (ETSI) is a valid starting point [9] but it does not actually provide built-in building blocks and components to handle most common EI task such as Edge Training (for which Federated learning and Knowledge Transfer Learning are two commonly cited techniques) and Edge Inference (common techniques in this field include Model Compression, Model Partitioning, and Early-Exit of Inference). The most common architectural setting for EI, instead, foresees the Cloud level for heavyweight and time-insensitive tasks and the Edge level, populated by increasingly powerful, smaller, and cheap IoT boards and microcomputers, deputed to sensing, actuation and lightweight computing tasks such as caching and pre-processing. Indeed, it is important to underline that the EI it is not opposed to Cloud Computing but pushes for the creation of a *continuum* in which exclusively as a matter of opportunity.

The discussed features have made the EI very appealing for different application domains: a more comprehensive list can be found in [14] but a selection is reported hereinafter. In healthcare, for example, EI can be applied to wearable devices or medical sensors to collect patient data (such as heart rate, blood pressure, or glucose levels), analyze it locally (aiming to detect critical health events, trigger alerts, or to provide immediate feedback to patients) and free of privacy concerns, finally sending to the cloud aggregated or anonymized reports. EI plays a crucial role to enhances safety and responsiveness in autonomous vehicles, where real-time decision-making is critical: by deploying EI capabilities within vehicles, data from various sensors can be processed locally to detect objects, recognize traffic signs, and make immediate navigation decisions, while on the Cloud offline analysis of the driving style can be performed [15]. Another suitable application domain are smart cities, aiming to contribute to sustainable urban planning and improved quality of life: indeed EI can greatly support environmental monitoring by collecting and immediately processing data on air quality, noise levels, or weather conditions, thus enabling real-time detection of pollution, identification of patterns, and immediate response to environmental events. Likewise, also the IIoT is a key sector for EI but this will be discussed in the following subsection. Said that, although EI represents a highly promising field with broad appeal and usefulness, the path towards its full realization is still long: on-going and future research directions aim to standard APIs, data models, programming paradigms and supporting infrastructures (mandatory including 5G and 6G technologies) as well as to software platforms for EI, lightweight operating systems for edge devices and full-fledged development methodology. In the absence of such a sound and well-grounded specific enablers, these open issues represent nowadays a not negligible entry barrier for the adoption of EI solution, reason why simulation approaches as the one reported in [17] are being recently developed to preliminary assess its pros and cons, especially for large-scale deployments.

2.2. EI and IIoT.

As anticipated, the localized processing power offered by the EI generically brings intelligence close to the data sources and, when possible, directly on-device. Such a paradigm shift with respect to the conventional Cloud-based scenario has relevant impact especially in the IIoT domain [47]. First, by locally performing data-driven tasks such as predictive maintenance, anomaly detection and automated quality control, the time required for data transmission to the cloud and back is avoided, facilitating faster decision-making and, consequently, improving operational efficiency and reducing downtime. Provided with local intelligence, the IIoT edge devices can adapt, optimize, and self-regulate their operations and energy consumption based on real-time data: in such a way, they are able to make on-the-spot decisions without relying on constant commands from a central server, improving overall system performance with a smaller environmental footprint. Moreover, they can continue to operate even if the cloud connection is lost (certainly not a remote occurrence in industrial environments, often characterized by conditions adverse to connectivity, e.g., elevate humidity or temperature values), thus enhancing system reliability and reducing dependence on the network [26]. As a further consequence, the adoption of EI reduces the amount of the IIoT data that needs to be transmitted to the remote servers for processing (only relevant and actionable insights are sent, conserving bandwidth and minimizing network congestion and operational cost) and addresses data privacy concerns: indeed, to keep sensitive information local and to transmit only aggregated/anonymized information mitigate the risk of unauthorized access, data breaches and, in general, cybersecurity risks, which are extremely relevant for industries where large volumes of valuable data are generated and have to travel over the network [46]. By relegating the Cloud to time-insensitive tasks and offline procedures (reporting, data exploration, etc.) the productivity, efficiency and stability of industrial operations can be greatly improved, with extremely relevant benefits also for the workers. Moreover, by providing them with wearable IoT devices fully integrated within the Smart Factory ecosystem, indeed, accidents can be prevented as well as workers safety and well-being monitored and improved [24]. These wearables, adapted or purposely designed for industrial use, span from a range of different categories (e.g., Smart Suite, Health and Biometric trackers, Smart Helmets) and allow real-time monitoring workers' vital signs and environmental conditions as well as detecting/notifying fatigue, abnormal heart rate, or dangerous oxygenation/high stress levels [25]. If implemented according to EI principles and supported by low-power technologies, their effectiveness and efficacy is notable enhanced and privacy concerns mitigated.

On the basis of these considerations about synergies and opportunities, in the last five years the duo EI-IIoT has gained traction both academia and industry. Beside a good number of secondary work focused on theoretical aspects and analysis with macroscopic lens [27], [30], [35], [28] a few research papers concretely explored EI solution in comparison with the Cloud-based implementation. In this respect, particular emphasis can be given to [29], where an on-demand cooperative inference framework exploiting deep neural networks (DNN) model for inference with high accuracy and responsiveness is presented, and to [26], which conversely focused on the model training and proposed a federated active transfer learning model leverages on personalization and privacy preservation. Another line of research, instead, dealt with service placements and resource allocations [41], [42] taking into account devices' features like mobility and energy, Finally, an increasing number of works have been presented integrating EI and IIoT with Blockchain [27], [37] and/or Digital Twins, [38], [44] while the concept of IIoT continuum is mostly unexplored [13].

2.3. Related work

Video analytics enables very important activities in the Factory 4.0 but it is also a heavyweight (in term of both computation and data size) task which represents a perfect benchmark for the joint exploitation of Cloud and Edge Computing. In particular, the limited literature on EI-enabled video analytics for the IIoT scenario exposes three main reasons [19]: *(i)* the uplink network to the cloud is unsuitable for shipping continuous video streams; *(ii)* video should be mandatory processed on-premise because of private or sensitive data; and *(iii)* frequent outages or fails have place on (the path towards) the remote server. In particular, CEVAS [21] is a video analytics system spanning the Cloud and Edge that can automatically and dynamically identify the optimal Cloud-Edge partition points for concurrent video analytics pipelines and orchestrates their executions between the Cloud and Edge. Similarly, [20] focuses on hybrid Cloud-Edge deployments to enable video analytics closer to the edge, while delivering the Cloud Computing experience of scalability and manageability. In [43], instead, an edge-centric video analytics for IIoT systems is introduced with a particular emphasis on a real testbed and its architecture. Only [23] introduces an accurate comparative approach between Edge- and Cloud-based deployments but it exclusively focuses on the latency analysis.

3. Use case

In this Section we present a typical IIoT scenario, comprising heterogeneous devices, human and tasks with different requirements, which can greatly benefit from EI. Let imagine a workshop into a mechanical factory where workers interact with robots and machinery in presence of flammable material containers. In this scenario, an increase in temperature dictated by any event could trigger a fire and endanger the lives of the workers first of all, and finally compromise the assets and the workshop's operations. Therefore, the whole area is assumed to be monitored by temperature sensors, a videocamera and, for security reasons, workers carry wearable IoT devices to monitor their vital signs and to provide tactile feedback to notify the reception of alerts or notifications. This is necessary since fumes and noises produced by machinery could make other means of signaling (sirens, flashing lights, etc.) less visible/audible and reduce their effectiveness, with clear impact on the workers' safety. Given the outlined scenario, when the temperature exceeds a certain threshold, the videocamera monitoring the area might start processing the video, recognize the source of the fire as well as the identity of the workers in its proximity with the final goal to alert them through their wearable IoT devices. A Cloud- and an EI-based deployment are presented hereinafter and compared later with two possible architectures both exploiting a RaspberryPi4 as IoT device and Amazon Web Services (AWS), a comprehensive computing platform whose several services (ranging from computing power,

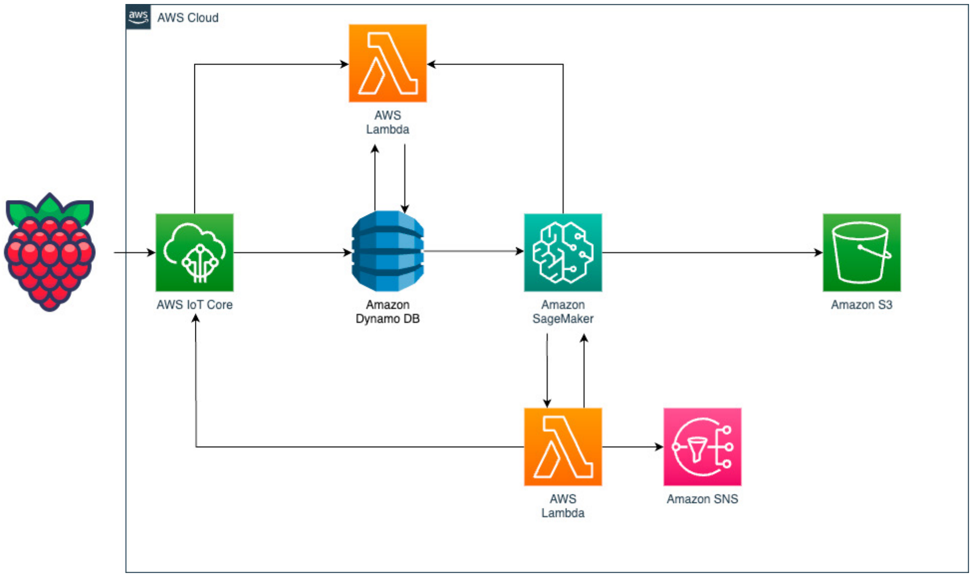

Fig. 1. Proposed Cloud-based deployment

storage options, networking, security etc.) constitute a wide and continuously expanding ecosystem [16].

Cloud-based deployment. Figure 1 shows a cloud-based deployment for the outlined scenario. The *AWS IoT Core* is the component deputed to the management and to the secure connectivity of the Raspberry Pi4; this microcomputer, through its temperature sensors, gathers raw data and exploits the *AWS IoT Core* to transmit them to the Cloud via the HTTP protocol. Here, an *AWS Lambda* function compares the current temperatures with the previous ones and decides whether to trigger the RaspberryPi4 for broadcasting the real time video through its videocamera, always by means of the aforementioned *AWS IoT Core*. The system hence process the incoming video streams in real-time through *Amazon SageMaker*, a key component for performing custom machine learning models. Once the source of fire is detected, it starts the worker's face recognition task and through the *Amazon Simple Notification Service (SNS)* alerts are set-up and sent to the identified workers. Finally, temperatures or video metadata, such as detected workers or GPS fire locations, are stored in an *Amazon DynamoDB* database for further analysis and retrieval while the raw video is stored in *Amazon S3* (a Simple Storage Service providing high durability, scalability, and security).

Edge-based deployment. Figure 2, instead, shows an EI-based deployment for the same scenario. By installing *AWS IoT Greengrass* on the IoT device it is possible to extend AWS managed services to the edge device (hence, remotely to manage large-scale deployments) as well as to run local processing, messaging, and data caching by means of ad-hoc components. In particular, a Python function can monitor the temperature data directly acquired by the RaspberryPi4's sensor, save it in on an *InfluxDB* (a time-series database designed specifically for handling timestamped data) instance and decides if trigger the video analysis: in the EI deployment, *TensorFlow* allow the offline training of the video analytics model while *TensorFlowLite* optimizes it for deployment on the edge device through. Then, another *Python function* is in charge of performing the inference and, based on the outcome of the worker recognition, triggering the notification through a local communicating protocol. As for the previous cloud-based deployment, also in this scenario it is possible to store small chunks of (meta)data like the outcome of the video analysis.

3.1. Comparative results analysis

Each of the two outlined deployments has benefits and drawbacks. Therefore, we have compared them (Table 1 provides readers with a bird's eye view of both quantitative and qualitative aspects), by considering the same data

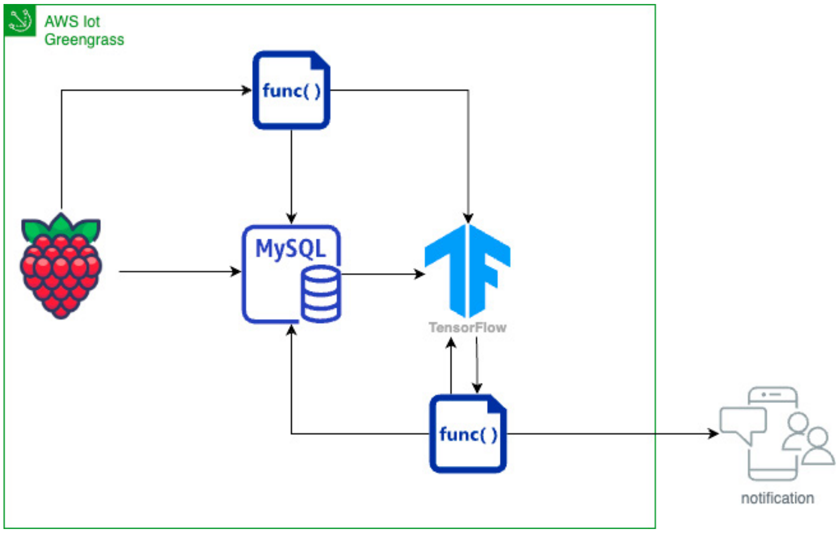

Fig. 2. Proposed EI-based deployment

stream of temperature to trigger the face recognition task over the same example video and according to the following metrics:

- Responsiveness: the inference time per frame varies from 0.01 second per frame in the Cloud-based deployment (Amazon EC2 P3.2xlarge instance) to the 0.28 second per frame of the RaspberryPi4 (ARM Cortex-A72 quad-core up to 1.5GHz). However, if using a Google Coral Board as Edge device, the inference time drops below 0.07 second per frame thanks to its Edge TPU, a co-processor which is a small ASIC that accelerates TensorFlow Lite models efficiently and that results faster than an desktop Intel i7 CPU (0.1 frame per second). Beside the processing time, the other relevant contribution is the communication time: 4 seconds are needed to upload the video in the Cloud-based deployment while the notifications on the wearable are delivered into an equivalent time, in both cases below 1 second.
- Bandwidth usage: in the Cloud-based deployment there is the need to upload the video on the remote server. The more the video is in high resolution, the more bandwidth will be used (with relative time). Such an aspect is particularly relevant if we consider that many videocamera might be located in the industrial area and proportionally consume bandwidth, with potential bottleneck effect. The video compression can be an alternative to alleviate such data traffic but it introduces an overhead and extra-time, and can also reduce the accuracy of the inference up to 35%. Conversely, in the EI-based deployment the bandwidth usage is negligible (few KB).
- Energy footprint: in the EI-based deployment, the RaspberryPi4 consumes approximately 3.1 watts during the temperature reading and 5.5 watts at its maximum load, i.e., inference operation. In the same conditions, the power consumption of the Cloud server is an order of magnitude higher even with power management techniques, but additional contributions due to its cooling, networking and auxiliary systems are not negligible.

3.2. *Discussion: opportunities and limitations*

As exemplified by the use case, the EI-based deployment guarantees undeniable opportunities to the IIoT. From a quantitative viewpoint, the recent advancements on tiny and low-cost hardware have significantly reduced the processing latency of edge devices; since the latter do not have to transfer raw data, there is no network latency and therefore the *responsiveness* of EI-based systems is easily comparable to that of traditional Cloud systems. Moreover, the reduced networking activity also impacts on the lower power consumption of edge devices (which are designed to be mobile and energy-saving by default) thus reducing the overall *energy footprint* and the *bandwidth usage*. Also from a qualitative viewpoint the the “sense-and-forward” approach featuring the Cloud-based deployment makes it strongly dependant by the stability of the internet connection and, therefore, the communication with the workers’ wearable

Table 1. Qualitative and Quantitative comparison of Cloud- and EI-based deployment

	Cloud-based Deployment	**EI-based Deployment**
Inference time	0.01 second per frame	0.28 second per frame (RaspberryPi4) 0.07 second per frame (Google Coral Board)
Upload time	4 seconds	<1 second
Bandwidth usage	Order of GB	Order of KB
Energy footprint	3.1 watts (temperature reading) 5.5 watts (inference)	Order of kWatts
Reliability	●○○	●●●
Usability	●●○	●○○

might not be ensured. Conversely, the EI-based guarantees the operational continuity of the machinery and, foremost, the stability of the alert service which is based on local inference and local (SIM-based, in our case, or ZigBee- or LoraWAN-based as alternative) communication technology. If *reliability* is definitively a selling point of EI, the discussion is more complex as regards *usability*. From-scratch EI implementations typically need relevant time and expertise to be built-up, configured and maintained: for example, in our case, we used AWS IoT Greengrass to ease offline and managed services, but we had to develop ad-hoc functions for the business logic. Cloud-based IoT Platforms like AWS or Microsoft Azure, instead, provide at different prices and on-demand a wide set of solid services, already implemented and ready-to-use over secure internet connection. Also the migration from an existing Cloud-based to an EI-based deployment demands time and resources since it is typically not transparent. Fortunately, platforms like AWS IoT SiteWise Edge [45] allows to mirror the same functional architecture in the cloud as at the edge and collect, organize, process, and monitor data on-premises, thus streamlining the whole development process. However, generally speaking, the entry barrier to fully master such a complex IoT Platforms (or better, IoT ecosystems) should not be overlooked, especially for the EI-based deployments.

Wrapping-up, albeit the EI still faces some challenges and limitations in the engineering of IIoT systems, it promises relevant advancements to the state-of-the-art of Cloud-based deployment. Moreover, the increasing interest on this paradigm and on the Edge-Cloud continuum will surely pave the way towards high-performance edge devices and new development approaches, as the one reported in the following Section.

4. A general strategy towards large-scale EI-aided IIoT deployments

Moving from small-scale use case to large-scale EI-aided IIoT deployments, the need of general and more comprehensive strategies goes naturally. Indeed, as theoretically discussed first and later exemplified by the presented use case, EI tasks have technical goals and requirements (e.g., accuracy, responsiveness, privacy) which are often orthogonal and interdependent with the inherent features of Cloud- and EI-based deployments (reliability, usability, cost, etc.). Moreover, also the choice of a specific hardware (Raspberry and Google Coral board), ML model (pre-trained or customized with transfer-learning) or communication protocol (ZigBee, LoraWAN, Bluetooth Low-Energy, MQTT have their own datarate, bandwidth, etc.) greatly impact on the overall performance of the system, especially when it has to scale-up. Therefore, in order to properly engineer complex systems as the IIoT ones, systematic and comprehensive approach are necessary, even better if simulation-based (so to preliminary assess the most suitable deployment by avoiding unnecessary waste of time, resources and money). Along this direction, the EdgeMiningSim methodology presented in [18] can be successfully applied to breakdown the development process in well-defined phases (i.e., domain analysis, data analysis, task analysis, deployment modelling, simulation, results evaluation and validation) and conduct rigorous, long-lasting and scale-variable performance test. In such a way, it is possible to explore the wide space of possible configuration and settings (how many and what kind of devices to deploy, which task offloading policies, data rate and bandwidth, etc.) and finally outline the most suitable deployment case-by-case [22]. In particular, in that article a Smart Monitoring scenario were considered, studying how relaxing the constraint of absolute accuracy or completely determinism without affecting the overall performance but rather improving the efficiency of the considered clustering task. With respect to our use case, the application of the EdgeMiningSim would allow performing interesting scalability tests to verify, for example, the resource consumption of different ML models on the IIoT edge devices, the maximum number of wearables simultaneously reachable within a certain time by the notification system, their energy consumption, etc. Following such a kind of analysis is strategic because it allows for more informed decision making, shortening the time-to-market and developing more performing system, aligned with the initial intents.

5. Conclusion

As the technology continues to advance, EI is certainly to become an even more integral part of the industrial ecosystem. In this paper we have surveyed opportunities and limitations of the EI-IIoT duo with a practical approach, firstly through a general discussion and secondly through an emblematic use case, aiming to compare both qualitatively and quantitatively compare Cloud- and EI-based deployments. Given the complexity of this challenge task, we have finally proposed a general strategy to systematically drive the engineering of large-scale IIoT scenarios according to given requirements and intents. As future work, we intend to further improve and extend the current use case and provide an integrated testbed for safety operations in the IIoT scenario, developed according to the aforementioned EdgeMiningSim methodology.

Acknowledgements

The research leading to this work was carried out under the Italian MIUR, PRIN 2017 Project "Fluidware" (CUP H24I17000070001), and under the "MLSysOps Project" (Grant Agreement 101092912) funded by the European Community's Horizon Europe Programme.

References

[1] Fortino, G., Russo, W., Savaglio, C., Shen, W. & Zhou, M. Agent-oriented cooperative smart objects: From IoT system design to implementation. *IEEE Transactions On Systems, Man, And Cybernetics: Systems*. **48**, 1939-1956 (2017)
[2] Li, S., Xu, L. & Zhao, S. The internet of things: a survey. *Information Systems Frontiers*. **17** pp. 243-259 (2015)
[3] Svertoka, E., Saafi, S., Rusu-Casandra, A., Burget, R., Marghescu, I., Hosek, J. & Ometov, A. Wearables for industrial work safety: A survey. *Sensors*. **21**, 3844 (2021)
[4] Sisinni, E., Saifullah, A., Han, S., Jennehag, U. & Gidlund, M. Industrial internet of things: Challenges, opportunities, and directions. *IEEE Transactions On Industrial Informatics*. **14**, 4724-4734 (2018)
[5] Fortino, G., Messina, F., Rosaci, D., Sarné, G. & Savaglio, C. A trust-based team formation framework for mobile intelligence in smart factories. *IEEE Transactions On Industrial Informatics*. **16**, 6133-6142 (2020)
[6] Fortino, G., Guerrieri, A., Savaglio, C. & Spezzano, G. A Review of internet of things Platforms through the IoT-A Reference Architecture. *International Symposium On Intelligent And Distributed Computing*. pp. 25-34 (2021)
[7] Barbuto, V., Savaglio, C., Chen, M. & Fortino, G. Disclosing edge intelligence: A systematic meta-survey. *Big Data And Cognitive Computing*. **7**, 44 (2023)
[8] Rydning, D., Reinsel, J. & Gantz, J. The digitization of the world from edge to core. *Framingham: International Data Corporation*. **16** pp. 1-28 (2018)
[9] Porambage, P., Okwuibe, J., Liyanage, M., Ylianttila, M. & Taleb, T. Survey on multi-access edge computing for internet of things realization. *IEEE Communications Surveys & Tutorials*. **20**, 2961-2991 (2018)
[10] Yu, W., Liang, F., He, X., Hatcher, W., Lu, C., Lin, J. & Yang, X. A survey on the edge computing for the Internet of Things. *IEEE Access*. **6** pp. 6900-6919 (2017)
[11] Sun, W., Liu, J. & Yue, Y. AI-enhanced offloading in edge computing: When machine learning meets industrial IoT. *IEEE Network*. **33**, 68-74 (2019)
[12] Liang, F., Yu, W., Liu, X., Griffith, D. & Golmie, N. Toward edge-based deep learning in industrial Internet of Things. *IEEE Internet Of Things Journal*. **7**, 4329-4341 (2020)
[13] Cabrini, F., Valiante Filho, F., Rito, P., Barros Filho, A., Sargento, S., Venâncio Neto, A. & Kofuji, S. Enabling the industrial Internet of Things to cloud continuum in a real city environment. *Sensors*. **21**, 7707 (2021)
[14] Zhou, Z., Chen, X., Li, E., Zeng, L., Luo, K. & Zhang, J. Edge intelligence: Paving the last mile of artificial intelligence with edge computing. *Proceedings Of The IEEE*. **107**, 1738-1762 (2019)
[15] Lin, K., Li, C., Li, Y., Savaglio, C. & Fortino, G. Distributed learning for vehicle routing decision in software defined Internet of vehicles. *IEEE Transactions On Intelligent Transportation Systems*. **22**, 3730-3741 (2020)
[16] Wittig, M. & Wittig, A. Amazon web services in action. (Simon,2018)

[17] Casadei, R., Fortino, G., Pianini, D., Placuzzi, A., Savaglio, C. & Viroli, M. A methodology and simulation-based toolchain for estimating deployment performance of smart collective services at the edge. *IEEE Internet Of Things Journal*. **9**, 20136-20148 (2022)

[18] Savaglio, C. & Fortino, G. A simulation-driven methodology for IoT data mining based on edge computing. *ACM Transactions On Internet Technology (TOIT)*. **21**, 1-22 (2021)

[19] Bhardwaj, R., Xia, Z., Ananthanarayanan, G., Jiang, J., Shu, Y., Karianakis, N., Hsieh, K., Bahl, P. & Stoica, I. Ekya: Continuous learning of video analytics models on edge compute servers. *19th USENIX Symposium On Networked Systems Design And Implementation (NSDI 22)*. pp. 119-135 (2022)

[20] McCann, J., Quinn, L., McGrath, S. & Flanagan, C. Video Surveillance Architecture from the Cloud to the Edge.. *International Journal For Computers & Their Applications*. **29** (2022)

[21] Zhang, M., Wang, F., Zhu, Y., Liu, J. & Wang, Z. Towards cloud-edge collaborative online video analytics with fine-grained serverless pipelines. *Proceedings Of The 12th ACM Multimedia Systems Conference*. pp. 80-93 (2021)

[22] Alam, M., Hassan, M., Uddin, M., Almogren, A. & Fortino, G. Autonomic computation offloading in mobile edge for IoT applications. *Future Generation Computer Systems*. **90** pp. 149-157 (2019)

[23] Zen, K., Mohanan, S., Tarmizi, S., Annuar, N. & Sama, N. Latency Analysis of Cloud Infrastructure for Time-Critical IoT Use Cases. *2022 Applied Informatics International Conference (AiIC)*. pp. 111-116 (2022)

[24] Svertoka, E., Saafi, S., Rusu-Casandra, A., Burget, R., Marghescu, I., Hosek, J. & Ometov, A. Wearables for industrial work safety: A survey. *Sensors*. **21**, 3844 (2021)

[25] Roda-Sanchez, L., Garrido-Hidalgo, C., Hortelano, D., Olivares, T. & Ruiz, M. OperaBLE: an IoT-based wearable to improve efficiency and smart worker care services in Industry 4.0. *Journal Of Sensors*. **2018** (2018)

[26] Foukalas, F. & Tziouvaras, A. Edge Artificial Intelligence for Industrial Internet of Things Applications: An Industrial Edge Intelligence Solution. *IEEE Industrial Electronics Magazine*. **15**, 28-36 (2021)

[27] Zhang, Y., Huang, H., Yang, L., Xiang, Y. & Li, M. Serious challenges and potential solutions for the industrial Internet of Things with edge intelligence. *IEEE Network*. **33**, 41-45 (2019)

[28] Qiu, T., Chi, J., Zhou, X., Ning, Z., Atiquzzaman, M. & Wu, D. Edge computing in industrial internet of things: Architecture, advances and challenges. *IEEE Communications Surveys & Tutorials*. **22**, 2462-2488 (2020)

[29] Zeng, L., Li, E., Zhou, Z. & Chen, X. Boomerang: On-demand cooperative deep neural network inference for edge intelligence on the industrial Internet of Things. *IEEE Network*. **33**, 96-103 (2019)

[30] Chen, B., Wan, J., Lan, Y., Imran, M., Li, D. & Guizani, N. Improving Cognitive Ability of Edge Intelligent IIoT through Machine Learning. *IEEE Network*. **33**, 61-67 (2019)

[31] Zhang, K., Zhu, Y., Maharjan, S. & Zhang, Y. Edge intelligence and blockchain empowered 5G beyond for the industrial Internet of Things. *IEEE Network*. **33**, 12-19 (2019)

[32] Tang, S., Chen, L., He, K., Xia, J., Fan, L. & Nallanathan, A. Computational intelligence and deep learning for next-generation edge-enabled industrial IoT. *IEEE Transactions On Network Science And Engineering*. (2022)

[33] Le Minh, K. & Le, K. Odlie: On-demand deep learning framework for edge intelligence in industrial internet of things. *2021 8th NAFOSTED Conference On Information And Computer Science (NICS)*. pp. 458-463 (2021)

[34] Yu, Y., Chen, R., Li, H., Li, Y. & Tian, A. Toward data security in edge intelligent IIoT. *IEEE Network*. **33**, 20-26 (2019)

[35] Hafeez, T., Xu, L. & Mcardle, G. Edge intelligence for data handling and predictive maintenance in IIOT. *IEEE Access*. **9** pp. 49355-49371 (2021)

[36] Ren, L., Liu, Y., Wang, X., Lü, J. & Deen, M. Cloud–edge-based lightweight temporal convolutional networks for remaining useful life prediction in IIoT. *IEEE Internet Of Things Journal*. **8**, 12578-12587 (2020)

[37] Khezr, S., Yassine, A., Benlamri, R. & Hossain, M. An edge intelligent blockchain-based reputation system for IIoT data ecosystem. *IEEE Transactions On Industrial Informatics*. **18**, 8346-8355 (2022)

[38] Lu, Y., Huang, X., Zhang, K., Maharjan, S. & Zhang, Y. Low-latency federated learning and blockchain for edge association in digital twin empowered 6G networks. *IEEE Transactions On Industrial Informatics*. **17**, 5098-5107 (2020)

[39] Lu, Y., Huang, X., Zhang, K., Maharjan, S. & Zhang, Y. Communication-efficient federated learning for digital twin edge networks in industrial IoT. *IEEE Transactions On Industrial Informatics*. **17**, 5709-5718 (2020)

[40] Bellavista, P., Della Penna, R., Foschini, L. & Scotece, D. Machine learning for predictive diagnostics at the edge: An IIoT practical example. *ICC 2020-2020 IEEE International Conference On Communications (ICC)*. pp. 1-7 (2020)

[41] Gong, Y., Yao, H., Wang, J., Li, M. & Guo, S. Edge intelligence-driven joint offloading and resource allocation for future 6G industrial internet of things. *IEEE Transactions On Network Science And Engineering*. (2022)

[42] Wang, T., Zhang, Y., Xiong, N., Wan, S., Shen, S. & Huang, S. An effective edge-intelligent service placement technology for 5G-and-beyond industrial IoT. *IEEE Transactions On Industrial Informatics*. **18**, 4148-4157 (2021)

[43] Bazhenov, N., Harkovchuk, A. & Korzun, D. Edge-centric video data analytics for smart assistance services in industrial systems. *Proc. 14th Int'l Conf. On Mobile Ubiquitous Computing, Systems, Services And Technologies (UBICOMM)*. (2020)

[44] Fortino, G. & Savaglio, C. Integration of Digital Twins & Internet of Things. *The Digital Twin*. pp. 205-225 (2023)

[45] McCarthy, D. AWS at the Edge: A Cloud Without Boundaries. *International Data Corporation Accessed Via Https://d1. Awsstatic. Com/IoT/IDC-AWS-at-the-Edge-White-Paper. Pdf*. (2020)

[46] Fotia, L., Delicato, F. & Fortino, G. Trust in edge-based internet of things architectures: state of the art and research challenges. *ACM Computing Surveys*. **55**, 1-34 (2023)

[47] Rajawat, A., Goyal, S., Chauhan, C., Bedi, P., Prasad, M. & Jan, T. Cognitive Adaptive Systems for Industrial Internet of Things Using Reinforcement Algorithm. *Electronics*. **12**, 217 (2023)