

Sensor data fusion for the industrial artificial intelligence of things

Youcef Djenouri¹ | Asma Belhadi² | Gautam Srivastava^{3,4}  |
Essam H. Houssein⁵  | Jerry Chun-Wei Lin⁶ 

¹SINTEF Digital, Oslo, Norway

²Kristiania University College, Oslo, Norway

³Brandon University, Brandon, Canada

⁴Research Centre for Interneural Computing, China Medical University, Taichung, Taiwan

⁵Minia University, Minya, Egypt

⁶Western Norway University of Applied Sciences, Bergen, Norway

Correspondence

Jerry Chun-Wei Lin, Western Norway University of Applied Sciences, Bergen, Norway.

Email: jerrylin@ieee.org

Abstract

The emergence of smart sensors, artificial intelligence, and deep learning technologies yield artificial intelligence of things, also known as the AIoT. Sophisticated cooperation of these technologies is vital for the effective processing of industrial sensor data. This paper introduces a new framework for addressing the different challenges of the AIoT applications. The proposed framework is an intelligent combination of multi-agent systems, knowledge graphs and deep learning. Deep learning architectures are used to create models from different sensor-based data. Multi-agent systems can be used for simulating the collective behaviours of the smart sensors using IoT settings. The communication among different agents is realized by integrating knowledge graphs. Different optimizations based on constraint satisfaction as well as evolutionary computation are also investigated. Experimental analysis is undertaken to compare the methodology presented to state-of-the-art AIoT technologies. We show through experimentation that our designed framework achieves good performance compared to baseline solutions.

KEYWORDS

AIoT, cognitive smart cities, deep learning, multi-agent system

1 | INTRODUCTION

2.5 Billion citizens are expected to occupy smart cities by 2050 and will need to be accommodated while using their smart devices (Rose et al., 2021). The understanding of how cities will function is a subtle balance of business, retail, leisure, energy, housing, ecology and mobility as well as being able to manage any external systems (Rose et al., 2021). The use of heterogeneous data in these smart cities is growing at a tremendous pace in both volume and type (Lom & Pribyl, 2021). This leads to the importance of intelligent systems to improve the role of city planning through smart city modelling (Arafah et al., 2021; Djenouri & Comuzzi, 2017; Goodwin & Russomanno, 2009). When one looks at the current focus within smart industry, it can be seen that smart devices, as well as sensors and smart devices, are everywhere in smart environments used to learn while collecting data from events that occur in the artificial intelligence of things (AIoT) (Tange et al., 2020), which is a sub-area of internet of things (IoT) that uses AI in many different applications (Lin Jerry, Shao, et al., 2021; Lin Jerry, Srivastava, et al., 2021). Due to many different characteristics like scale, dynamically evolving systems, complexity, and self-organization, AIoT is an exciting topic in both the networking and artificial intelligence community (Wu et al., 2020). With the current push in industrial settings, the need for an intelligent mechanism in handling smart

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Expert Systems* published by John Wiley & Sons Ltd.

sensor data became a necessity. One of the main reasons for this is that companies do not analyse their data but learn from the data to extract valuable knowledge and decision in the future.

Although IoT technologies provide tools to deal with smart sensors analysis, the exponentially growing network environment in IoT also brings to light complex and heterogeneous information, which will be a challenge in future industrial settings (Da et al., 2014; Baza et al., 2020; Li et al., 2018; Zhang & Tao, 2021). Without strong intelligent techniques for learning from events as well as data, serious degradation in terms of systems performances may be lead.

1.1 | Motivation

Various artificial intelligence (AI) is widely applied to solve complex problems in IoT applications (Tan et al., 2021). Deep learning, multi-agent systems, knowledge graphs, are just examples of such methods that prove their success in the recent decade. Deep learning (Dai et al., 2019; Dai et al., 2020; Dwivedi et al., 2019; Liu et al., 2019; Luo et al., 2020; Ma et al., 2018; Weng et al., 2019) are used to learn from the previous examples of the industrial environment. For instance, the work in Weng et al. (2019) incorporated both blockchain and reinforcement learning to address the federated learning challenges such as security and privacy. Multi-agents systems (Alqahtani et al., 2020; Casado-Vara et al., 2018; Chen, Srivastava, et al., 2020; Ciatto et al., 2019; Cicirelli et al., 2019; Zhou et al., 2019) are used to address the challenges of the internet of things applications. For instance, the works in Alqahtani et al. (2020) and Casado-Vara et al. (2018) use intelligent agents to improve the security and privacy in the internet of things environments. This is ensured by integrating blockchain technology and certificate authentication mechanisms for data sharing among users of the platform. Besides, the works (Ciatto et al., 2019; Cicirelli et al., 2019) attempt to enhance cognition of the internet of things applications by integrating the main features of the intelligent agents' behaviours in the whole internet of things system. In the context of knowledge graphs, several semantic-based solutions have been proposed in the AIoT (Bellini et al., 2014; Iwendi et al., 2020; Le-Phuoc et al., 2016; Li et al., 2019; Qiu et al., 2018; Qiu et al., 2019). For instance, Qui et al. (2019) combines both word co-occurrences and community detection to build the semantic graph of heterogeneous information of the smart sensors data.

Previous AIoT solutions do not address the industrial challenges where large-scale data are usually collected from distributed and heterogeneous smart sensors. Therefore, new methods are needed to improve the existing AIoT systems in both accuracy and runtime performances.

1.2 | Contributions

The novelty of the present research is to develop an end-to-end solution to overcome the limitations of the AIoT systems, by proving an efficient and intelligent collaborative framework using a multi-agent system, the knowledge graph, and deep learning. The main contributions are fourfold as follows:

1. We propose a new intelligent collaborative framework to handle smart sensor data in the AIoT. The proposed framework uses deep learning and a multi-agent system for creating automated models representing large-scale data collected from heterogeneous and distributed smart sensors.
2. We use both the convolution neural network and the recurrent neural network in the learning phase. The convolution neural network deals with matrix representation such as images, where the recurrent neural network deals with time series representation such as trajectories.
3. We offer a novel intelligent method for enhancing communication among the various actors by combining the knowledge graph process. We also investigate different optimizations such as constraint satisfaction programming, and evolutionary computation to make the whole framework robust and efficient.
4. We evaluate the proposed framework using two industrial use cases, including smart building, and intelligent transportation. The results show that our novel framework can be useful when directly compared to two baseline methods for object detection and two other baseline methods for outlier detection.

The remainder of this work is structured as follows: Related work is studied in Section 2. The proposed framework and designed algorithm are discussed in Section 3. We report our experimental results in Section 4. Section 5 presents the main finding of the application of hybrid collaborative intelligent framework (HCIF) in industrial settings. Section 6 concludes the paper.

2 | LITERATURE REVIEW

Different deep learning methods have been developed in the context of artificial intelligence internet of things applications. Dai et al. (2019) created an architecture based on reinforcement learning that can be used to secure next-generation networks in IIoT. The designed system was

shown to maximize utility as well as accurately cache data sharing throughout the network. Weng et al. (2019) implemented a framework called DeepChain, which is a deep learning-based framework that is also distributed able to navigate federated learning problems. In DeepChain learners could in essence behave incorrectly where updating parameters. Liu et al. (2019) tackle IIoT problems and were able to adopt a learning approachable to give a mechanism used for the evaluation of IIoT systems using in terms of decentralization, scalability, security, throughput, and latency. Dai et al. (2020) again investigate the problem of online offloading but this time made use of a Markov-based decision process. The authors integrated deep learning as well as the genetic algorithm (GA). The process was shown to be able to maximize offloading performance in the long term. Luo et al. (2020) administered a distributed IoT technology that could synchronize many continuous local views using independent SDN (software-defined networking) controllers. The technology could reach a consensus even on global views. The approach itself was shown to reduce computational resources significantly even when considering hidden features of particular controllers as well as resource constraints in the environment used.

To improve cognition, and autonomy several solutions based on multi-agent systems have been proposed. Cicirelli et al. (2019) suggested a multi-agents platform to establish cognition in decentralized smart city based algorithms. It also exploits the intelligent agents for developing internet of things applications and makes it reactive, proactive, and cognitive behaviours. Roberto et al. (2018) developed a multi-agent system to analyse and manage the supply chain process in the internet of things settings. It also provides an efficient mechanism for ensuring security using blockchain technology. Ciatto et al. (2019) developed agentify system which enhances the autonomy of the intelligent agents by exploring sociality and situatedness features of each agent in the system. Alqahtani et al. (2020) developed a multi-agent system for enhancing monitoring and improving the security scheme in the cloud internet of things environment. It creates a middleware layer to manage the communication among users in the platform. Intelligent agents provide security by sharing trust and certificates for authentication. Alsbouei et al. (2020) suggested the use of distributed ledger platform based on mobile agents to improve the scalability of data sharing in tremendous P2P networks. It also provides an intelligence mechanism, where the mobile agents are integrated for node-level communications and data collections.

Communication is a crucial issue in multi-agent systems. A knowledge graph is a powerful artificial intelligence representation that targets the semantic modelling of the different entities in the intelligent systems. Bellini et al. (2014) introduced a system for the management of large-volume data from heterogeneous sources that consider both static and dynamic data in IoT applications. Qui et al. (2018) established a semantic graph-based approach for identifying non-taxonomic connections in IoT settings by integrating semantic graph structure information and context information. Le et al. (2016) provided an uniform consolidated and live display for diverse city data sources. It aggregates and enriches millions of triples for connecting to a graph in real-time each hour by addressing billions of historical and current records. Qui et al. (2019) developed a graph approach for reliably extracting semantic knowledge from heterogeneous data from smart sensors. Because of similarities, data from smart cities are initially calculated using the term co-occurrence. Following that, a semantic network is created based on similarities between smart city data. Finally, a community identification method is used, which may partition a given smart city into independent, disjoint communities, each of which can be defined as a separate concept.

Through this short review of the literature, one can see that there is a lot of exciting research that has been implemented with connections to AIIoT. Many research items look into learning mechanisms for the creation of automated models, while some methods also explore multi-agent systems that are capable of dealing with smart sensor data from distributed environments. Moreover, other methods dive into semantic modelling for helping data sharing in IoT environments. Nonetheless, these methods have not been shown to be successful in industrial settings, where there has not been an abundance of recent research. In a different direction altogether, in this research work, we investigate to our knowledge the first dedicated complete framework that combines knowledge graphs, multi-agent systems, and deep learning to envision well-used solutions to solve smart sensor data for applicability in AIIoT.

3 | HYBRID COLLABORATIVE INTELLIGENT FRAMEWORK

3.1 | Principle

Let us begin by describing the key elements of the HCIF. As shown in Figure 1. The input data is collected from smart sensors in the internet of things settings. The data is entered into the HCIF framework, for analysis and pattern learning. The output of the HCIF will be a set of guidelines to the user for helping the decision-making process. Note that the proposed HCIF framework is a generic one, which may be applied for solving any decision-making tasks including, prediction, outlier analysis, regression, and so on. Our framework builds upon several intelligent components such as deep learning, multi-agent systems, knowledge graphs, constraint satisfaction, and evolutionary computation for dealing with smart sensor data. In particular, we use deep learning to create automated models from the smart sensors data. As HCIF is a generic framework, different deep learning architectures are used, the convolution neural network is used for handling images and videos, where the recurrent neural network is used for handling time series data. The multi-agents system is then used to deal with distributed and heterogeneous smart sensor data, where the communication is realized using the knowledge graph techniques. Different optimization has been integrated, the first one is the constraint satisfaction programming, where the second one is the evolutionary computation process. Each agent first performs the deep learning architectures to

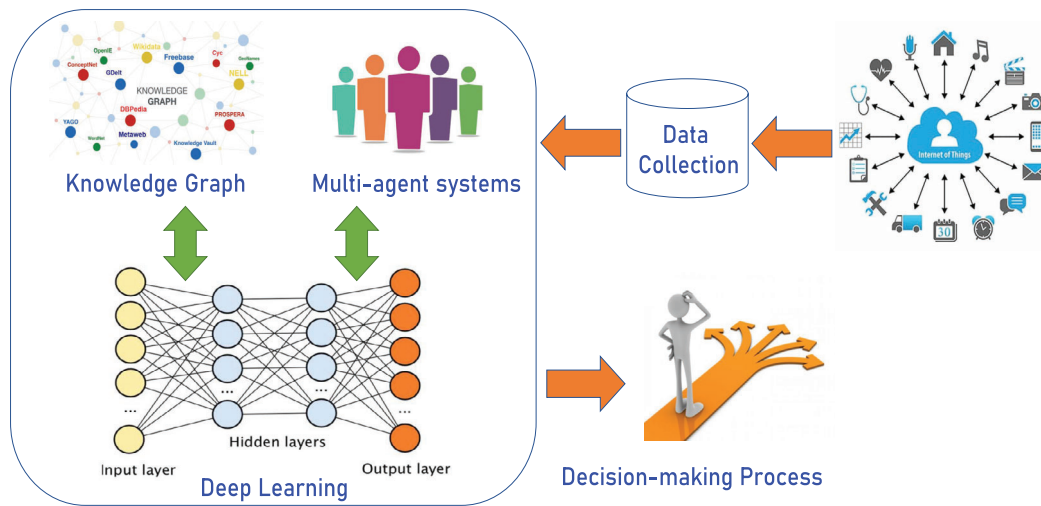


FIGURE 1 Hybrid collaborative intelligent framework

build intelligent models. In the inference step, each agent uses ontology matching to interpret its output to the other agents in the system. The security privacy issues are modelled by the set of constraints, and solved by two strategies; the first one is an exact strategy, where the complete instantiation is generated, and the second one is the approximate-based strategy, where the aim is to find an instantiation converged to the optimal one. PSO is also made use of for the optimization of processing runtime for the solver as developed. HCIF components are detailed and described in the rest of this section.

3.2 | Deep learning

Here, we made use of two DL architectures for use in the training of smart sensor data.

1. *Convolution neural network*: Convolution Neural Network (CNN) (Mishkin et al., 2017) is a group of deep frameworks that can be applied often into computer vision (CV) problems like visual recognition as well as object detection. CNN has also recently been applied to many other data representation types like time series as well as textual data. The primary purpose of CNN is in feature extraction directly from matrix data making use of convolutional filters. Here, we can define convolutional filters as the set of weights applied directly to each and every element within the matrix data pixels. Weights are adjusted as well as learned making use of traditional backpropagation methods. In our research, we used VGG16, the well-known CNNs, which is known for being able to achieve 92.8% top six test accuracy on Image Net, with close to 16 million image examples in close to 2000 classes.()
2. *Recurrent neural network*: Recurrent neural network (RNN) architecture is used (De Mulder et al., 2015). The input consists of heterogeneous smart sensor data. A multilayer-feed-forward network is applied made up of many neurons that are formed into layers. For any given layer, a neuron gets linked to all other neurons using a combination of different weight values. Input layer neurons get associated with every input data. Output layer neurons get connected to each output. The overall aim is to be able to minimize the error between the data that is output from the network and its corresponding ground truth data.

The m th neuron output from layer l , defined as s_l^m is shown in Equation 1. We note here that the output sums for all neurons for a given layer set to 1, as illustrated in Equation 2.

$$s_l^m = \sigma \left(\sum_{j=1}^{|l-1|} s_{l-1}^j \cdot \omega_{l-1}^{mj} + b_l^m \right), \quad (1)$$

with

$$\sum_{m=1}^{|l|} s_l^m = 1, \quad (2)$$

where b_l^m is the bias value for neuron s_l^m , $|I|$ is the number of neurons in layer l , s_{l-1}^j is the output of j th neuron in $l-1$ layer, ω_{l-1}^{mj} is the weight value connects neurons s_l^m to s_{l-1}^j , and $\sigma(\cdot)$: activation function.

The updating weight rule is given in each iteration i by multiplying the previous weights with the i th smart sensor data D_i , the learning rate μ , and the error rate E_i , as illustrated in Equation 3:

$$\omega_{l-1}^{mj}(i) = \omega_{l-1}^{mj}(i-1) - \mu \times D_i \times 2 \times E_i, \quad (3)$$

where the error rate E_i is described in Equation 4 as follows:

$$E_i = \sum_{j=1}^k (D_{ij} - \hat{D}_{ij})^2 \quad (4)$$

The algorithm starts by initializing weight values. At iteration i , neurons of the input layer get input data that consists of smart-sensor data. Input gets explored through neurons in the network making use of Equation 1. Network output gets compared to ground truth, where any error is found using Equations 4. Error-values are propagated throughout the network for updating weight value as shown using Equation 3. The process is repeated for all input.

3.3 | Multi-agents systems

Collaboration is an important factor in the AIoT. To simulate the interaction among the different entities in the AIoT, a multi-agent system is integrated into the HCIF framework. To increase the smart behaviour of the agents in HCIF, reinforcement learning is developed. Therefore, we define a multi-agents system by a tuple $\langle \mathcal{A}, \mathcal{S}, \mathcal{U}, \mathcal{R} \rangle$. \mathcal{A} is the set of agents, each can be shown to be a decision process based on Markov, where \mathcal{S} is defined as a finite set of environment states, where \mathcal{U} is defined as a set of actions and also where \mathcal{R} can be defined as the reward function. Each agent's behaviour as it pertains to \mathcal{A} can be represented by a policy, which specifies how every agent can choose actions for a given state. Each agent's purpose is to maximize a given objective function based on its policy. For example, in the case of prediction, a policy of each agent is to maximize the number of predicted objects. The following describes the different components of the multi-agents system of HCIF:

1. State: Each agent's future action is determined by the preceding states' decisions. As a result, each agent's state is made up of two parts: the set of prior actions and the current data to be handled. The size of \mathcal{S} is determined by observations in a particular database.
2. Action: It is the assignment of each observation's decision-making behaviour in the database. For instance, in a prediction job, it is the assignment of each object's class.
3. Reward: Determining a suitable reward function is critical. It enables each agent in \mathcal{A} to learn more effectively. We used data that contained ground truth to provide rewards to the agent's activities. The following is the definition of the reward function:

$$R(\mathcal{A}_i, \mathcal{U}_j) = \begin{cases} 1, & \text{if } \mathcal{A}_i(\mathcal{U}_j, O_j) = \mathcal{G}(O_j) \\ 0, & \text{otherwise,} \end{cases}$$

where $\mathcal{A}_i(\mathcal{U}_j, O_j)$ is the decision of the agent \mathcal{A}_i , whether the observation O_j has correct action or not. $\mathcal{G}(O_j)$ is the ground-truth of the observation O_j .

4. Environment: The environment is a collection of databases that store a huge amount of data from smart sensor devices. This enables the environment to produce specific situations for the agent's training and to estimate the optimal actions to perform.

Each and every agent \mathcal{A}_i begins by scanning all of the observations in the i th smart sensor, agent \mathcal{A}_i then calculates the 1st observation, as well as remaining observations for the i th sensor. The given reward function $R(\mathcal{A}_i, \mathcal{U}_j)$ is calculated on each decision using ground truth of 1st observation. As detailed, the process given is repeated then for the remainder of observations of i th sensor. The resulting local decision set LD _{i} can be derived for each and every agent \mathcal{A}_i .

3.4 | Knowledge graphs

Intelligent communication may be used to integrate the local choices of the agents in \mathcal{A} by utilizing a knowledge graph. Ontology matching is one of the robust techniques, which allows representing the knowledge graphs. In the following, we present an approach that allows matching between the different ontologies of the agents in \mathcal{A} .

Definition 1. Consider the set of l ontologies $\mathcal{O} = \{\mathcal{O}_1, \dots, \mathcal{O}_l\}$, each ontology \mathcal{O}_i shows the set of m_i instances of the agent \mathcal{A}_i such that $\mathcal{I}^i = \{\mathcal{I}_1^i, \dots, \mathcal{I}_{m_i}^i\}$, and n_i properties or attributes $\mathcal{P} = \{\mathcal{P}_1^i, \dots, \mathcal{P}_{n_i}^i\}$. The problem statement's purpose relating to the ontology-matching problem among agents that are $\in \mathcal{A}$ can be said to be for the determination of common properties that exist between ontologies. In other words, for the determination of the function \mathcal{M} as is given in Equation 5.

$$\mathcal{M}(\mathcal{O}_i, \mathcal{O}_j) = \left\{ \left\{ \bigcup_{1 \leq m, s \leq m_j} \mathcal{M}_{ij}(\mathcal{I}_m^i, \mathcal{I}_s^j) \right\} \right\}, \quad (5)$$

where

$$\mathcal{M}_{ij}(\mathcal{I}_m^i, \mathcal{I}_s^j) = \{p \mid p \in \mathcal{I}_m^i \wedge p \in \mathcal{I}_s^j\}. \quad (6)$$

The naïve approach to the ontology-matching problem is to compare all values of instances across ontologies. The outcome of the alignment is determined by the matching procedure. Each match may result in a unique alignment instance. Each alignment result is then reviewed and compared to the reference alignment. The reference alignment is one that has been recommended by a user or domain expert. The reference alignment comprises all instances of the common ontology. Any sort of naïve approach can be highly costly from a time perspective especially for smart sensors that are known to usually collect vast amounts of data. As a solution to this problem, in this paper we suggest and propose a novel algorithm that can be used directly with ontology matching in any smart sensor environment. The algorithm first breaks up the entire set of instances of each and every ontology into many clusters that still have dependencies. Through this process, each cluster still has some correlated instances that will need to be processed later. Second, the algorithm explores all of the instances in a given cluster to try and locate common features. The algorithm primarily uses the matching and clustering processes. With the clustering process, the set of instances are split up into many collections of what are known as sub-instances (sub-clusters) implementing data mining methodologies (Fournier-Viger et al., 2019; Gan et al., 2017; Lin, Djenouri, & Srivastava, 2021; Lin, Djenouri, Srivastava, Yun, & Fournier-Viger, 2021). So far the steps can all be considered as pre-processing steps. Next, the instance sets are further grouped into separate clusters with a smaller number of instances per cluster. Common properties are shared between cluster instances to some threshold value. This threshold value as such allows each cluster to remain highly correlated. Through a matching process, the algorithm investigates all of the instances in a given cluster to look for alignments. Alignments between instances of two ontologies and their clusters are performed instead of looking to perform alignments between instances of ontologies one at a time.

3.5 | Optimization

To improve privacy, we employ an optimization solution that is based on a constraint satisfaction problem. The limitations of the whole smart sensor data set in IoT settings are divided into various groups. The clusters are investigated in order to identify the optimal configuration that secures privacy. We design two techniques for investigating constraint clusters:

Approximation-based strategy. Clusters are processed separately not considering shared data between sites. Local instantiation of data is carried out on each cluster using a search process. A merging function is used to create a global instantiation. This function combines all local instantiations of data. Such an approach returns a subset of all data from the whole set of constraints. This is because the shared data were not taken into account in the search process.

Proposition 1. An upper/lower bound of constraints satisfied by the approximation-based strategy, noted $|\mathcal{A}|$, is $|C| - \left(\sum_{i=1}^k S(G_i) \right)$, where C is the clusters of the constraints, k is the number of clusters, $S(G_i)$ is the number of conflicts constraints of the cluster G_i , and we note $|C| - \left(\sum_{i=1}^k S(G_i) \right) \leq |\mathcal{A}| \leq |C|$.

Proof. In the worst case, the number of non-satisfied constraints of the given approximation-based strategy is equal to the number of constraints from the shared variables between all clusters. This is maybe realized, where the shared variables appeared in the constraints of the clusters. In this case, the number of satisfied constraints of the approximation-based strategy is equal to $|C|$ minus all the number of constraints, containing the shared variables of all clusters, equal to $\sum_{i=1}^k S(G_i)$. In the case of the conflict, constraints do not use the shared variables during the search process, the number of satisfied constraints of the approximation-based strategy is $|C|$. \square

According to this proposition, the quality of the approximation-based method is largely dependent on the quantity of common data between all clusters. The approximation-based method can meet all requirements if the number of shared data is kept to a minimum. This will be fixed by carefully selecting the number of clusters.

Exact strategy: In the search process, this method takes into account both shared data and clusters. This enables us to identify the best potential instantiation of smart sensor data to meet all restrictions. To generate the local instantiation of data, the search method is first performed to each cluster of constraints. The local instantiation of data is concatenated with the instantiation of shared data to get the global instantiation of all constraints.

In our previous works (Belhadi, Djenouri, Diaz, et al., 2021; Djenouri et al., 2020), we tried GA (Genetic Algorithm) and PSO (Particle Swarm Optimization) in solving the hyper-parameters issues, however, we find PSO provides better intensification and diversification compared to the GA. PSO (Darwish et al., 2020) can be used for solver optimization in terms of runtime. PSO was chosen for this assignment because of its well-known balance of intensification and diversity, both of which are crucial in this context (constraint satisfaction). Next, the key idea of PSO is given for solving our constraint problem.

Initialization of Population: We randomly generate the initial population for particles making use of the solution space.

Particle Updating: Let us consider the set of particles P , then there exists a vector X for position such that $X_{ti} = (x_{i1}x_{i2}x_{i3}\dots x_{in})^T$ and we say that the velocity vector is $V_{ti} = (v_{i1}v_{i2}v_{i3}\dots v_{in})^T$ at iteration t for each particle i that it is composed of. The particles themselves are able to update their own positions using velocity formula given in Equations 7 and 8.

$$V_i^{t+1} = w \times V_i^t + c1 \times (p^t - X_i^t) + c2 \times (p^* - X_i^t), \quad (7)$$

and

$$X_i^{t+1} = X_i^t + V_i^{t+1}, \quad (8)$$

where $i = 1, 2, \dots, |P|$.

Here we can say that individual particles may be able to update positions making use of Equation 7, where two factors are given, c_1 as well as c_2 , that can respectively be shown to have contribution in particle movement for each and every iteration. Here, let p^t be the position of the best particle iteration t . Furthermore, let us define p^* as the position of the best particle overall iteration. Equation 8 may be used in the creation of updates in particle position. Let the parameter w be positive as well as constant, and say that w can be used for balancing the global search, which is known as exploration, as well as the local search, known simply as exploitation.

Fitness computing: It is used to evaluate the solution S , where S is the number of constraints met. If a constraint does not contradict the data exchanged between the sensors, it is met. The objective is to maximize this value as much as possible. The following is the fitness function:

$$\text{Fitness}(S) = \sum_{C_i \in C} \text{satisfied}(C_i), \quad (9)$$

where $\text{satisfied}(C_i) = 1$ if C_i is satisfied, 0, otherwise.

4 | PERFORMANCE EVALUATION

In this section, we compare the proposed HCIF and compare it with state-of-the-art solutions. In particular, the following two case studies have been analysed:

1. *Smart building:* We used the NESTA162-bus Data (Thams et al., 2019) which is a set of data that contains $N - 1$ possible contingencies that represent all plausible operating points for any given energy demand profile. This set contains over 1 million points. In this dataset, topology changes are also included for $N - 1$ investigations. The aim is to identify fault diagnosis in electric power systems for smart building applications using object detection. For evaluating the objects detected, accuracy represented by mean average precision (mAP) and computational runtime are used. mAP can be considered a well-known metric for the evaluation of object detection systems, which is the sum of precision of all objects detected over the number of all objects detected, as shown in Equation 10:

$$\text{mAP} = \frac{\sum_{i=0}^n \text{Precision}(i)}{n}, \quad (10)$$

where $\text{Precision}(i)$ is defined as precision at rank i , and n is defined as all detected objects. In other words, the first i ranked objects can be considered while all other remaining objects can be ignored. We compare HCIF with the clustering particle swarm optimization for object

detection (CPOD) (Djenouri et al., 2020), and fast region convolution neural network (FastRCNN) (Girshick, 2015) which are state of the art detection solutions.

2. *Intelligent transportation*: For ITS, we make use of data from *ECML PKDD 2015* available at Alqahtani et al., 2020. This database is time-series based containing actual trajectories from over 400 taxis in Porto, Portugal in the span of 2013–2014. The CSV file contains over 3 GB. Each row defined a singular trip, which includes *TripID*, *TaxiID*, *CallType*, as well as timestamps and GPS information. The CSV file also contains a singular coordinates pair for every trip for every 15 s of that particular trip. The last item in each row is the logged destination for the journey, and the first item logged is the beginning of the journey. The aim of this dataset was originally to find local as well as global taxi trajectory outliers. As a generality, a well-known common issue in anomaly detection is in the actual evaluation procedure itself. This will be the case as well in new IoT environments such as AIoT, where the baseline is usually unknown due to the newness of the environment. To help with quantitative evaluation, in this paper, we make use of the methodology implemented with (Belhadi et al., 2020) where synthetic anomalous patterns are injected.

Injecting local outliers: local outliers are generated by adding noise several times with a certain probability $p \sim \mathcal{U}(0.8, 1.0)$ and a given threshold μ ;

Injecting global outliers: From the local outliers, we again add noise but now only a few times with a certain probability $p \sim \mathcal{U}(0.0, 1.0)$ and a given μ .

For both injections, each data d_i in each dataset is changed as follows:

$$d_i = \begin{cases} d_i + n \sim \mathcal{N}(0, 1) & \text{if } d \geq \mu \\ d_i & \text{otherwise.} \end{cases} \quad (11)$$

The ratio of corrected returned outliers to all outliers is used to do the evaluation. This number is between 0 and 1, with a larger value indicating greater accuracy. We compare HCIF with the collective abnormal of human behaviours (CaHB) (Belhadi, Djenouri, Srivastava, et al., 2021), and group trajectory outliers (GTO) (Belhadi et al., 2020) which are the state-of-the-art anomaly detection solutions.

Our experimental evaluation has been performed on a personal computer with a 64-bit core i7 processor running Windows 10 and 16 GB of RAM. The host CPU is a 64-bit quad-core Intel Xeon E5520 running at 2.27 GHz. The GPU is an Nvidia Tesla C2075 with 448 CUDA cores for 448 (14 multiprocessors with 32 cores each) and a clock speed of 1.15 GHz. It has a global memory capacity of 2.8 GB, a shared memory capacity of 49.15 KB, and a warp size of 32. Single precision is utilized on both the CPU and GPU. In our implementation scenario, we employed GPU blocks to replicate the environment of a multi-agent system. Each agent is assigned to a single GPU block, with the associated agent sharing the block's shared memory. The communication between agents is accomplished through the use of the GPU host's global and constant memory.

4.1 | Smart building

Our experimental evaluation in this subsection aims to validate the usability of HCIF in smart building applications. Figure 2 presents both runtime and accuracy of HCIF compared to the two baseline algorithms FastRCNN (Girshick, 2015), and CPOD (Djenouri et al., 2020) using smart building dataset while varying the number of points and images. As shown in the first and third plots of Figure 1, the accuracy of the HCIF on NESTA162-bus database is compared with FastRCNN (Girshick, 2015), and CPOD (Djenouri et al., 2020). Making adjustments to the points as well as images for input, HCIF can be seen to outperform both baseline algorithms all three comparatives (accuracy, runtime, and mAP). This is explained by the fact that HCIF uses efficient strategies to identify objects by incorporating both the convolution neural network and the multi-agents system, the learning is highly optimized using particle swarm optimization and constraint satisfaction programming. The second, and the fourth plots of the Figure 2 the runtime of the HCIF on NESTA162-bus database is compared with FastRCNN (Girshick, 2015), and CPOD (Djenouri et al., 2020). Causing variations in image numbers from 2000 to 7000, HCIF can outperform both baseline algorithms in just runtime. We can give meaning to this by stating that HCIF prunes the bounding boxes by using intelligent communication among the different agents in the framework efficiently.

4.2 | Intelligent transportation

Our experimental evaluation in this subsection aims to validate the usability of HCIF in intelligent transportation applications. Figure 3 presents both runtime and accuracy of HCIF compared to the two baseline algorithms CaHB (Belhadi, Djenouri, Srivastava, et al., 2021), and GTO (Belhadi et al., 2020) using intelligent transportation dataset while varying the number of points and trajectories. As shown in the first and the third plots of Figure 3, the accuracy of the HCIF on ECML PKDD database, is compared with CaHB (Belhadi, Djenouri, Srivastava, et al., 2021), and GTO

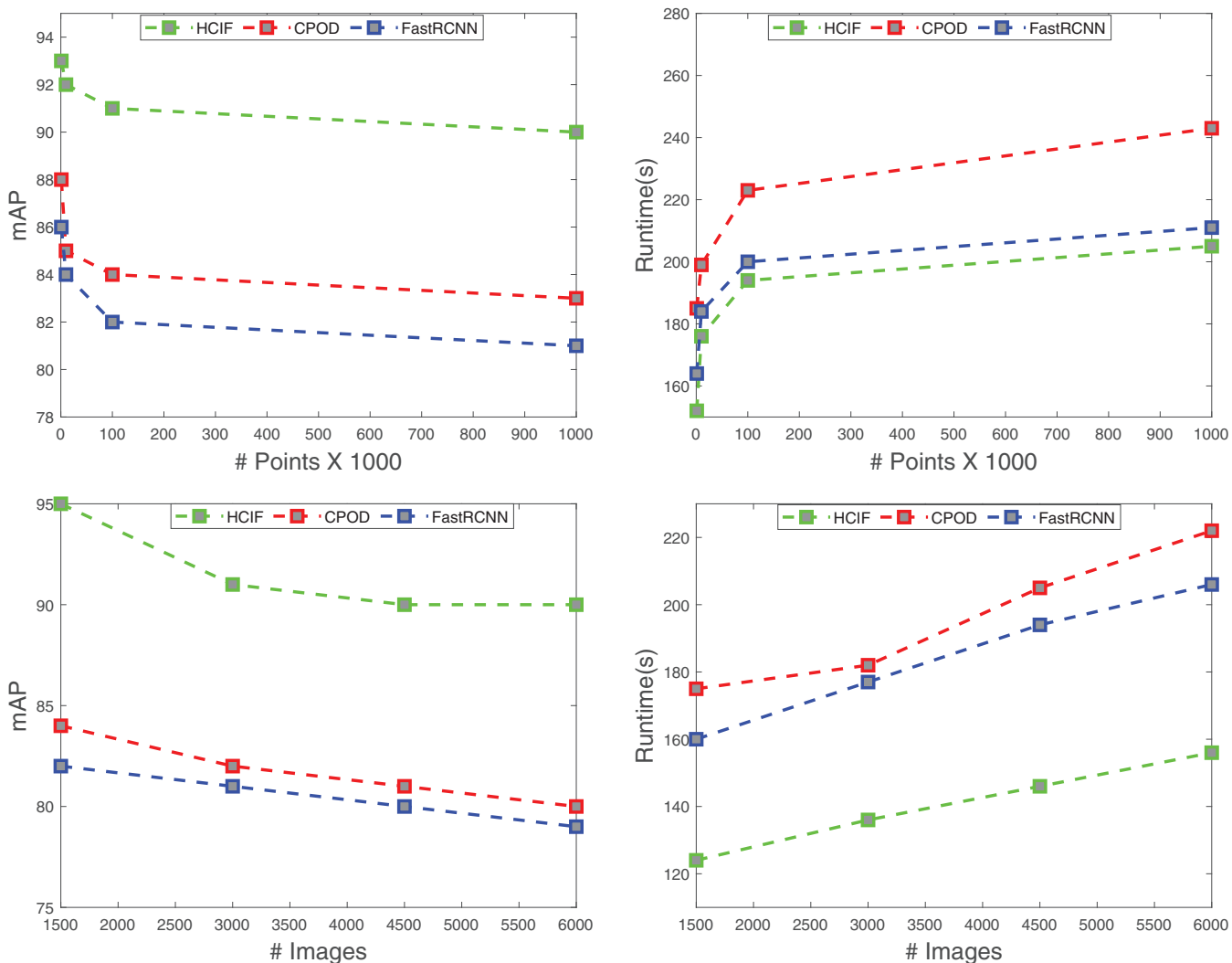


FIGURE 2 Performance comparison of the hybrid collaborative intelligent framework and the state-of-the-art object detection solutions in the context of smart building application by varying the number of points, and the number of images

(Belhadi et al., 2020). By varying with the number of points and trajectories used as input, HCIF can be seen to outperform both baseline algorithms in just accuracy, shown using the ratio of the corrected outliers. This is explained by the fact that HCIF uses efficient strategies to detect both local and global anomalous by incorporating both the recurrent neural network and the multi-agents system. The second, and the fourth plots of the Figure 3, the runtime of HCIF on ECML PKDD database is compared with CaHB (Belhadi, Djenouri, Srivastava, et al., 2021), and GTO (Belhadi et al., 2020). Giving variations in trajectory numbers from 1500 to 5000, HCIF can be seen to outperform both baseline algorithms in just runtime. We can give meaning to this by stating that HCIF efficiently explores the solution space of both local and global outliers. Indeed, the local outlier is determined on each agent, where robust communication among agents allows the identification of global outliers.

5 | DISCUSSIONS AND FUTURE PERSPECTIVES

The main results of using a hybrid framework to handle smart sensor data from AIoT settings are discussed in this section. First off, heterogeneous data can easily be handled by the proposed framework by incorporating various deep learning architectures, and knowledge graphs. The deep learning architectures allow solving different data representations collected from different kinds of smart sensors data. Convolution neural networks allow dealing with matrix data such as images, and transactions, where the recurrent neural networks perform very well on time series data such as energy consumption, traffic flow, and signal processing. This differs from previous intelligent approaches that only focused on single data representation. From an optimization research standpoint, the proposed framework shows a differing combination of optimization techniques, to accurately improve the performance of modern industrial systems. In our specific context, constraint satisfaction and evolutionary

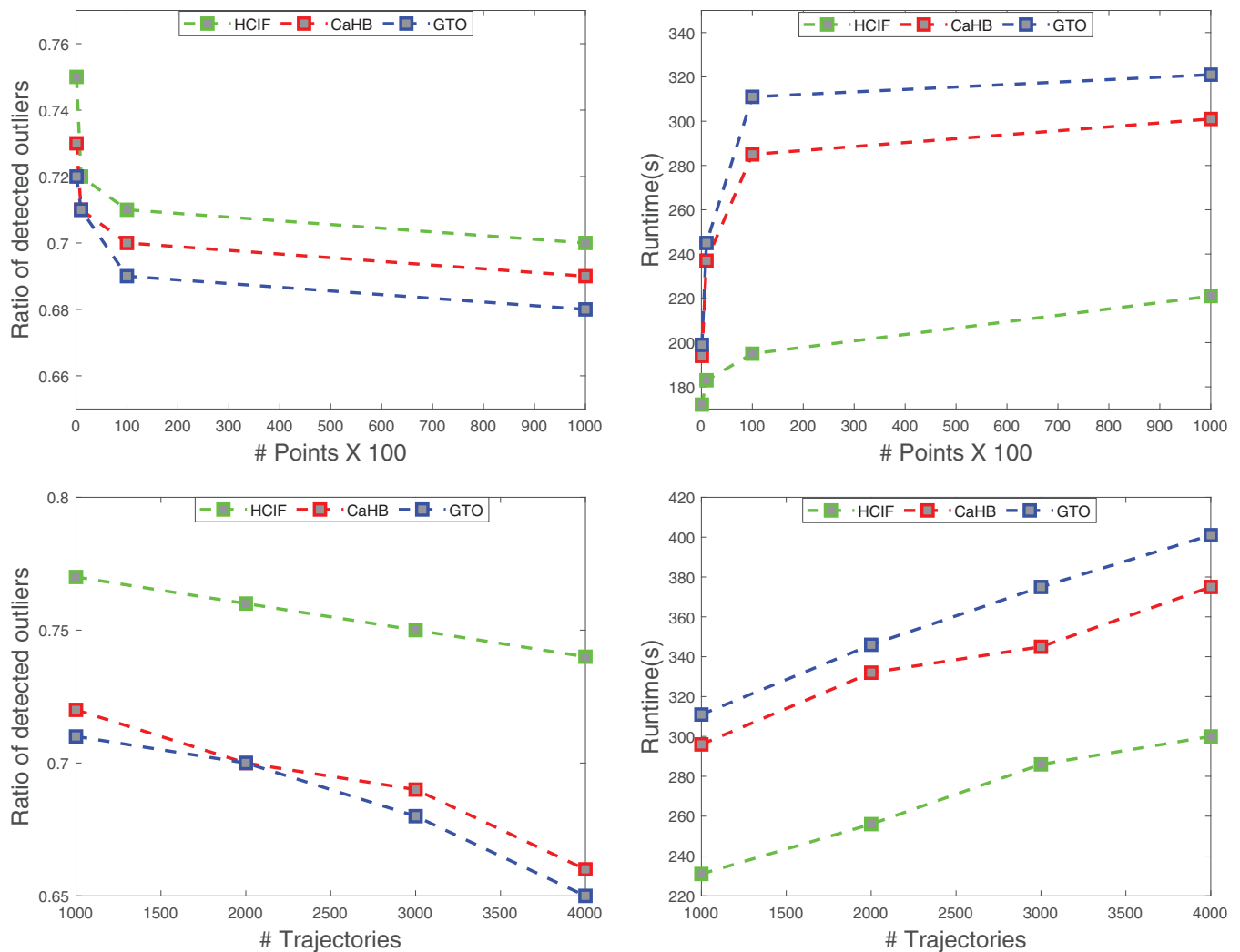


FIGURE 3 Performance comparison of the hybrid collaborative intelligent framework and the state-of-the-art outlier detection solutions in the context of intelligent transportation application by varying the number of points, and the number of trajectories

computation meet the artificial intelligence internet of things for handling complex smart sensors data. Further findings of this work are that the multi-agents system benefits from the ontology matching process. Thus, each agent can communicate with the other agents through the set of common instances and properties derived by the ontology matching process. Finally, our framework is designed generically and as such can easily be applied to differing AIoT applications, which requires intelligent collaboration among different complex entities. This scenario reflects well the current needs in the industrial environments. Instead of the positive outcomes of this paper, future work may include:

1. *Learning step improvements.* Convolution neural networks and recurrent neural networks have been used as deep learning architectures. Additional deep learning models may be investigated for reducing the error-learning rate. Therefore, a topic of interest is the integrating of other deep learning models into the proposed framework in the future, such as transfer learning (Lin et al., 2020), active learning (Deng et al., 2020), and reinforcement learning (Chen, Cui, et al., 2020). We also could potentially look at mechanisms for the automation of the tuning procedure for hyper parameters in deep-learning models. Forcing the framework to take several iterations to find the best hyper-parameters seems wasteful. Using some sort of knowledge base that can then be studied based on correlative effects would be an interesting future direction.
2. *Runtime processing improvements.* For a fast framework performance boost as well as allow its application to larger problems through the exploitation of HP-computing (high-performance computing) tools like supercomputers (Chen et al., 2018) and techniques like cluster computing (Zhang et al., 2017). With this in mind, our aim as researchers is to identify an independent job in every cluster of sensor data by keeping in mind the HP computing challenges that may occur which include communication, load balancing, synchronization, as well as memory management. Furthermore, making use of methodologies to handle load balancing becomes very important. A known way to handle this well-known

issue is through the development of decomposition strategies that can allow the location of equitable clusters concerning the number of smart sensors identified per cluster.

3. *Use cases.* This paper already presented four use cases of the proposed framework, namely smart building, and intelligent transportation. Taking into account the positive results, further work in domain-specific problems that present big data may be of interest. Some examples include business intelligence applications or the studying of business-specific financial data. For automated trading, runtime performance would be critical since trading can be very time-dependent.

6 | CONCLUSION

This paper proposes a novel intelligent framework for the collaborative IoT by combining different components. Deep learning architectures such as convolution neural networks and recurrent neural networks are used to learn from the different smart sensors' data. A multi-agent system is used to simulate the collective behaviours of the smart sensors in the internet of things environment. The communication among the different agents is assured by developing a new strategy for the ontology matching process. Different optimizations based on constraint satisfaction and meta-heuristics are also investigated. Our proposed framework is directly compared to existing artificial intelligence internet of things technologies. Our results are strong and show clearly that our proposed framework has good performance compared to the baseline solutions. With respect to future perspectives, we plan to improve the learning process by adding additional deep learning models by using transfer learning, and active learning. Another perspective of this research work is the ability to handle big data using high-performance computing. Exploring other use cases such as business intelligence is also on our future agenda.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in NESTA162-bus Data at <https://arxiv.org/abs/1806.01074>, reference number 42.

ORCID

Gautam Srivastava  <https://orcid.org/0000-0001-9851-4103>

Essam H. Houssein  <https://orcid.org/0000-0002-8127-7233>

Jerry Chun-Wei Lin  <https://orcid.org/0000-0001-8768-9709>

REFERENCES

- Alqahtani, F., Al-Makhadmeh, Z., Tolba, A., & Said, O. (2020). TBM: A trust-based monitoring security scheme to improve the service authentication in the internet of things communications. *Computer Communications*, 150, 216–225.
- Alsoubi, T., Qin, Y., Hill, R., & Al-Aqrabi, H. (2020). Enabling distributed intelligence for the internet of things with IOTA and mobile agents. *Computing*, 102(6), 1345–1363.
- Arafah, M., Ceravolo, P., Mourad, A., Damiani, E., & Bellini, E. (2021). Ontology based recommender system using social network data. *Future Generation Computer Systems*, 115, 769–779.
- Baza M., Mahmoud M., Srivastava G., Alasmay W., Younis M.. A light blockchain-powered privacy-preserving organization scheme for ride sharing services. *Proceedings of the IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, 2020.
- Belhadi, A., Djenouri, Y., Diaz, V. G., Houssein, E. H., & Lin, J. C. (2021). Hybrid intelligent framework for automated medical learning. *Expert Systems*.
- Belhadi, A., Djenouri, Y., Djenouri, D., Michalak, T., & Lin, J. C. (2020). Deep learning versus traditional solutions for group trajectory outliers. *IEEE Transactions on Cybernetics*.
- Belhadi, A., Djenouri, Y., Srivastava, G., Djenouri, D., Lin, J. C., & Fortino, G. (2021). Deep learning for pedestrian collective behavior analysis in smart cities: A model of group trajectory outlier detection. *Information Fusion*, 65, 13–20.
- Bellini, P., Benigni, M., Billero, R., Nesi, P., & Rauch, N. (2014). Km4City ontology building vs data harvesting and cleaning for smart-city services. *Journal of Visual Languages and Computing*, 25(6), 827–839.
- Casado-Vara, R., Prieto, J., Prieta, F., & Corchado, J. M. (2018). How blockchain improves the supply chain: Case study alimentary supply chain. *Procedia Computer Science*, 134, 393–398.
- Chen, C., Cui, M., Li, F. F., Yin, S., & Wang, X. (2020). Model-free emergency frequency control based on reinforcement learning. *IEEE Transactions on Industrial Informatics*, 17(9), 2336–2346.
- Chen, C.-H., Lin, M.-Y., & Liu, C.-C. (2018). Edge computing gateway of the industrial internet of things using multiple collaborative microcontrollers. *IEEE Network*, 32(1), 24–32.
- Chen, Q., Srivastava, G., Parizi, R. M., Aloqaily, M., & Al, R. I. (2020). An incentive-aware blockchain-based solution for internet of fake media things. *Information Processing & Management*, 57(6), 102370.
- Ciatto G., Maffi A., Mariani S., Omicini A.. Towards agent-oriented blockchains: Autonomous smart contracts. *Proceedings of the International Conference on Practical Applications of Agents and Multi-Agent Systems*, pp. 29–41, 2019.
- Cicirelli F., Guerrieri A., Spezzano G., Vinci A.. A cognitive enabled, edge-computing architecture for future generation IoT environments. *Proceedings of the IEEE 5th World Forum on Internet of Things (WF-IoT)*, pp. 35–40, 2019.
- Da, X. L., He, W., & Li, S. (2014). Internet of things in industries: A survey. *IEEE Transactions on Industrial Informatics*, 10(4), 2233–2243.

- Dai, Y., Xu, D., Maharjan, S., Chen, Z., He, Q., & Zhang, Y. (2019). Blockchain and deep reinforcement learning empowered intelligent 5G beyond. *IEEE Network*, 33(3), 10–17.
- Dai, Y., Xu, D., Zhang, K., Maharjan, S., & Zhang, Y. (2020). Deep reinforcement learning and permissioned blockchain for content caching in vehicular edge computing and networks. *IEEE Transactions on Vehicular Technology*, 69(4), 4312–4324.
- Darwish, A., Ezzat, D., & Hassani, A. E. (2020). An optimized model based on convolutional neural networks and orthogonal learning particle swarm optimization algorithm for plant diseases diagnosis. *Swarm and Evolutionary Computation*, 52, 100616.
- De Mulder, W., Bethard, S., & Moens, M.-F. (2015). A survey on the application of recurrent neural networks to statistical language modeling. *Computer Speech & Language*, 30(1), 61–98.
- Deng, H., Yang, K., Liu, Y., Zhang, S., & Yao, Y. (2020). Actively exploring informative data for smart modeling of industrial multiphase flow processes. *IEEE Transactions on Industrial Informatics*, 17(2), 8357–8366.
- Djenouri, Y., & Comuzzi, M. (2017). Combining Apriori heuristic and bio-inspired algorithms for solving the frequent itemsets mining problem. *Information Sciences*, 420, 1–15.
- Djenouri, Y., Srivastava, G., & Lin, J. C. (2020). Fast and accurate convolution neural network for detecting manufacturing data. *IEEE Transactions on Industrial Informatics*, 17(4), 2947–2955.
- Dwivedi, A. D., Srivastava, G., Dhar, S., & Singh, R. (2019). A decentralized privacy-preserving healthcare blockchain for IoT. *Sensors*, 19(2), 326.
- Fournier-Viger P., Lin JCW, Truong-Chi T., Nkambou R. A survey of high utility itemset mining. *High-Utility Pattern Mining*. 2019: 1–45.
- Gan, W., Lin, H. C., & Zhan, J. (2017). Data mining in distributed environment: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 7(6), e1216.
- Girshick R. Fast r-cnn. *Proceedings of the IEEE International Conference on Computer Vision*, 2015.
- Goodwin, J. C., & Russomanno, D. J. (2009). Ontology integration within a service-oriented architecture for expert system applications using sensor networks. *Expert Systems*, 26(5), 409–432.
- Iwendi, C., Moqurrab, S. A., Anjum, A., Khan, S., Mohan, S., & Srivastava, G. (2020). N-sanitization: A semantic privacy-preserving framework for unstructured medical datasets. *Computer Communications*, 161, 160–171.
- Le-Phuoc, D., Quoc, H. N. M., Quoc, H. N., Nhat, T. T., & Hauswirth, M. (2016). The graph of things: A step towards the live knowledge graph of connected things. *Journal of Web Semantics*, 37, 25–35.
- Li, G., Wu, J., Li, J., Wang, K., & Ye, T. (2018). Service popularity-based smart resources partitioning for fog computing-enabled industrial internet of things. *IEEE Transactions on Industrial Informatics*, 14(10), 4702–4711.
- Li, J., Zhou, Z., Wu, J., Li, J., Mumtaz, S., Lin, X., Gacanin, H., & Alotaibi, S. (2019). Decentralized on-demand energy supply for blockchain in internet of things: A microgrids approach. *IEEE Transactions on Computational Social Systems*, 6(6), 1395–1406.
- Lin, H., Hu, J., Xiaoding, W., Alhamid, M. F., & Piran, M. J. (2020). Towards secure data fusion in industrial IoT using transfer learning. *Proceedings of the IEEE Transactions on Industrial Informatics*.
- Lin, J. C. W., Djenouri, Y., & Srivastava, G. (2021). Efficient closed high-utility pattern fusion model in large-scale databases. *Information Fusion*, 76, 122–132.
- Lin, J. C. W., Djenouri, Y., Srivastava, G., Yun, U., & Fournier-Viger, P. (2021). A predictive GA-based model for closed high-utility itemset mining. *Applied Soft Computing*, 108, 107422.
- Lin Jerry, C. W., Shao, Y., Djenouri, Y., & Yun, U. (2021). ASRNN: A recurrent neural network with an attention model for sequence labeling. *Knowledge-Based Systems*, 212(5), 106548.
- Lin Jerry, C. W., Srivastava, G., Zhang, Y., Djenouri, Y., & Aloqaily, M. (2021). Privacy preserving multi-objective sanitization model in 6G IoT environments. *IEEE Internet of Things Journal*, 8(7), 5340–5349.
- Liu, M., Yu, F. R., Teng, Y., Leung, V. C. M., & Song, M. (2019). Performance optimization for blockchain-enabled industrial internet of things (IIoT) systems: A deep reinforcement learning approach. *IEEE Transactions on Industrial Informatics*, 15(6), 3559–3570.
- Lom, M., & Pribyl, O. (2021). Smart city model based on systems theory. *International Journal of Information Management*, 56, 102092.
- Luo J, Chen Q, Richard Yu F, Tang L. Blockchain-enabled software-defined industrial internet of things with deep reinforcement learning. . 2020
- Ma, Y., Wu, Y., Ge, J., & Jun, L. (2018). An architecture for accountable anonymous access in the internet-of-things network. *IEEE Access*, 6, 14451–14461.
- Mishkin, D., Sergievskiy, N., & Matas, J. (2017). Systematic evaluation of convolution neural network advances on the imagenet. *Computer Vision and Image Understanding*, 161, 11–19.
- Qiu, J., Chai, Y., Liu, Y., Gu, Z., Li, S., & Tian, Z. (2018). Automatic non-taxonomic relation extraction from big data in smart city. *IEEE Access*, 6, 74854–74864.
- Qiu, J., Chai, Y., Tian, Z., Du, X., & Guizani, M. (2019). Automatic concept extraction based on semantic graphs from big data in Smart City. *IEEE Transactions on Computational Social Systems*, 7(1), 225–233.
- Rose, G., Raghuram, P., Watson, S., & Wigley, E. (2021). Platform urbanism, smartphone applications and valuing data in a smart city. *Transactions of the Institute of British Geographers*, 46(1), 59–72.
- Tan L., Yu K., Ming F., Chen X., Srivastava G.. Secure and resilient artificial intelligence of things: A honeynet approach for threat detection and situational awareness. *IEEE Consumer Electronics Magazine*, 2021
- Tange, K., De Donno, M., Fafoutis, X., & Dragoni, N. (2020). A systematic survey of industrial internet of things security: Requirements and fog computing opportunities. *IEEE Communication Surveys and Tutorials*, 22(4), 2489–2520.
- Thams, F., Venzke, A., Eriksson, R., & Chatzivasileiadis, S. (2019). Efficient database generation for data-driven security assessment of power systems. *IEEE Transactions on Power Systems*, 35(1), 30–41.
- Weng, J., Weng, J., Zhang, J., Li, M., Zhang, Y., & Luo, W. (2019). Deepchain: Auditable and privacy-preserving deep learning with blockchain-based incentive. *IEEE Transactions on Dependable and Secure Computing*, 18(5), 2438–2455.
- Wu, H., Han, H., Wang, X., & Sun, S. (2020). Research on artificial intelligence enhancing internet of things security: A survey. *IEEE Access*, 8, 153826–153848.
- Zhang, J., & Tao, D. (2021). Empowering things with intelligence: A survey of the progress, challenges, and opportunities in artificial intelligence of things. *IEEE Internet of Things Journal*, 8(10), 7789–7817.

- Zhang, Q., Zhu, C., Yang, L. T., Chen, Z., Zhao, L., & Li, P. (2017). An incremental CFS algorithm for clustering large data in industrial internet of things. *IEEE Transactions on Industrial Informatics*, 13(3), 1193–1201.
- Zhou, R., Zhang, X., Wang, X., Yang, G., Wang, H., & Wu, Y. (2019). Privacy-preserving data search with fine-grained dynamic search right management in fog-assisted internet of things. *Information Sciences*, 491, 251–264.

AUTHOR BIOGRAPHIES

Youcef Djenouri obtained the PhD in Computer Engineering from the University of Science and Technology USTHB, Algiers, Algeria, in 2014. He is currently a research scientist at SINTEF Digital in Oslo, Norway. He is working on topics related to artificial intelligence and data mining, with focus on association rules mining, frequent itemsets mining, parallel computing, swarm and evolutionary algorithms and pruning association rules. He has published more than 100 refereed research papers, in the areas of data mining, parallel computing and artificial intelligence.

Asma Belhadi obtained her PhD in Computer Engineering from the University of Science and Technology USTHB Algiers, Algeria, in 2016. She is a postdoctoral researcher at Kristiania University College in Oslo, Norway. She is working on topics related to artificial intelligence and data mining, with focus on smart city applications. She has published more than 50 refereed research articles in the areas of artificial intelligence, and smart city applications.

Gautam Srivastava was awarded his B.Sc. degree from Briar Cliff University in the U.S.A. in the year 2004, followed by his M.Sc. and Ph.D. degrees from the University of Victoria in Victoria, British Columbia, Canada in the years 2006 and 2012, respectively. He then taught for 3 years at the University of Victoria in the Department of Computer Science, where he was regarded as one of the top undergraduate professors in the Computer Science Course Instruction at the University. From there in the year 2014, he joined a tenure-track position at Brandon University in Brandon, Manitoba, Canada, where he currently is active in various professional and scholarly activities. He was promoted to the rank of Associate Professor in January 2018. Dr. G, as he is popularly known, is active in research in the field of cryptography, data mining, security and privacy, and blockchain technology. In his 5 years as a research academic, he has published a total of 70 papers in high-impact conferences in many countries and in high status journals (SCI, SCIE) and has also delivered invited guest lectures on big data, cloud computing, internet of things, and cryptography at many universities worldwide. He is an Editor of several SCI/SCIE journals. He currently has active research projects with other academics in Taiwan, Singapore, Canada, Czech Republic, Poland, and the U.S.A. He is an IEEE Senior Member and also an Associate editor of the world renowned IEEE Access journal.

Essam H. Houssein received his PhD degree in Computer Science in 2012. He is an Associate Professor at the Faculty of Computers and Information, Minia University, Egypt. He is the founder and chair of the Computing and Artificial Intelligence Research Group (CAIRG) in Egypt. He has more than 100 scientific research papers published in prestigious international journals in the topics for instance meta-heuristics optimization, artificial intelligence, image processing, IoT and its applications. He also serves as a reviewer of more than 40 journals (Elsevier, Springer, IEEE, etc.). His research interests include WSNs, IoT, AI, Bioinformatics and Biomedical, Image processing, Data mining, and Meta-heuristics Optimization techniques.

Jerry Chun-Wei Lin received his Ph.D. in Computer Science and Information Engineering from National Cheng Kung University, Tainan, Taiwan, in 2010. He is now working as an Associate Professor at Department of Computing, Mathematics, and Physics, Western Norway University of Applied Sciences (HVL), Bergen, Norway. His research interests include data mining, privacy-preserving and security, big data analytics, and social networks. He has published more than 200 research papers in peer-reviewed international conferences and journals, which have received more than 1900 citations. He is the co-leader of the popular SPMF open-source data-mining library and the Editor-in-Chief of Data Mining and Pattern Recognition (DSPR), Associate Editor of Journal of Internet Technology and IEEE Access, and a member of the Editorial Board of Intelligent Data Analysis.

How to cite this article: Djenouri, Y., Belhadi, A., Srivastava, G., Houssein, E. H., & Lin, J. C.-W. (2021). Sensor data fusion for the industrial artificial intelligence of things. *Expert Systems*, e12875. <https://doi.org/10.1111/exsy.12875>