Scalable AI Frameworks for Real-Time Distributed Monitoring and Decision-Making in Smart Automation

James Robertson
Asst Professor

Abstract

Scalable AI-based systems are outlined that support distributed, real-time monitoring and decision-making in the context of smart automation systems. The architectural principles presented are driven by the following requirements: Distributed sensing and data ingestion; AI methods that support real-time monitoring of large-scale sensor networks; Scalability and performance enhancement; Support for privacy; Governance and data management regulation; Federated capabilities and privacy support preserving data confidentiality. Within this framework, federated adaptive neurons are proposed. These AI nano-scopes provide local information about distributed streams, support anomaly detection and declaration in large sensor networks such as water networks, and can operate in a privacy-preserving environment. Additionally, horizontal and vertical scalability with distributed or high-performance computing is exposed. The proposed principles are supported by experimental cases that illustrate a distributed real-time monitoring capability.

Monitoring of the water quality and surveillance of an aquatic environment is investigated as an application area. A Smart City use case within a European initiative is also proposed. Finally, recent advances regarding governance, security, and privacy of data, especially on personal nature, are examined. These aspects are worthy of special consideration, since data governance regulation is evolving rapidly. The analytics to monitor data that may cross data sovereignty boundaries, such as those originating from users, need to respect end-user privacy preferences.

Keywords: Scalable AI Monitoring Systems, Distributed Real-Time Analytics, Federated Artificial Intelligence, Privacy-Preserving Machine Learning, Federated Adaptive Neurons, Large-Scale Sensor Networks, Anomaly Detection in Water Networks, Smart Automation Architectures, High-Performance Distributed Computing, Data Governance and Sovereignty, Real-Time Stream Processing, Smart City Monitoring Systems, Aquatic Environment Surveillance, Confidentiality-Preserving Data Analytics, Horizontal and Vertical Scalability.

1. Introduction

Within the recent advances in Artificial Intelligence (AI), Deep Learning and other related AI techniques have become increasingly popular for achieving extraordinary success in image and speech recognition, recommender systems, and other applications. Nevertheless, the adoption of these AI techniques in safety-critical applications (e.g., healthcare, aviation, automotive) remains limited because of their lack of inherent support for uncertainty quantification, explainability, and other desirable properties for certification. Indeed, failure of these systems in such applications can result in catastrophic consequences. Consequently, the shortcomings of Deep Learning and related techniques have sparked off research into safer alternatives. These alternatives complement and augment Deep Learning, without attempting to replace it.

Intelligent Power Quality Management (IPQM) in sustainable smart grids utilizing intelligent monitoring, detection and mitigation of power quality abnormalities has attracted increasing research interest. Power Quality (PQ) issues and disturbances, however, remain controversial in the smart grid context. Key power quality disturbances of concern, such as voltage swells and dips, flicker, harmonics, voltage imbalances and

JOURNAL OF INNOVATIVE ACADEMIC RESEARCH
ISSN : 3049-2122
VOLUME: 02 ISSUE: 02
RECEIVED: APRIL 21 REVISED: MAY 04 ACCEPTED: MAY 23 PUBLISHED: JUNE 11

transients, explode into rapidly changing temporal and spatial scales. Consequently, the detection of such abnormalities and disturbances should generally engage a combination of computer vision, Deep Learning and online Data Analytics techniques, rather than a single Deep Learning-based framework.

1.1. Overview of the Study's Framework and Scope

An integrated, evidence-based synthesis addresses the systematic design of scalable artificial intelligence-based systems for distributed real-time monitoring and decision making in smart automation. The population of increasing sensor network deployments, which support high-resolution monitoring of smart environments, gives rise to large data volumes offered for near real-time AI-based monitoring decision-making systems. The investigation identifies the design and implementation challenges concerned with the development of such near real-time AI monitoring and decision-making systems and discusses these around five themes. These relate to the distributed architectures employed for data collection, the AI methods deployed for real-time monitoring, the scalability characteristics of the systems, the security governance and privacy issues encountered, and the design of event dashboards. Discussions centre on the systems' horizontal and vertical scalability, privacy-preserving techniques for federated systems, methods capacity, and the usefulness of event dashboards associated with distributed AI monitoring approaches.

AI-based monitoring systems are being developed for a wide variety of smart automation settings, including urban environments, road networks, transport and logistics processes, health-care activities, buildings, commercial and industrial processes, agriculture and food production, and many others. Such systems inherit large data volumes sourced from dense sensor deployments capable of supporting high-frequency sampling across the surroundings of the smart environment or process of interest. For these reasons, the population of AI-based monitoring systems is likely to continue increasing rapidly. The concerns of the study directly relate to those authors who seek to design and implement such monitoring systems.

Fig 1: Scaling Intelligence: A Multidimensional Framework for Distributed Real-Time AI Monitoring and Decision-Making in Smart Automation

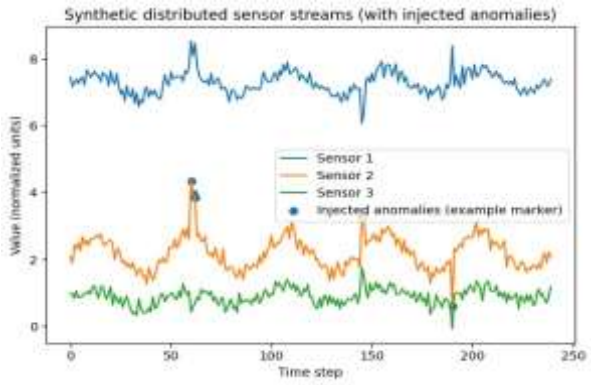
2. Background and Motivation

Increasingly, large-scale distributed smart environments are being explored as innovative embedded solutions in automation, automotive, and logistics applications. Their design and deployment require a scalable, end-to-end data dissemination framework for real-time decision-making and tailoring. Smart manufacturing environments leverage a collection of cooperative, process-embedded agents that support semi-free flow of materials and decision-making knowledge based on real-time data. AI methods are proposed to enable semi-autonomous decision making for real-time monitoring of safety-critical environments. The AI techniques handle key aspects of automation decision making, such

JOURNAL OF INNOVATIVE ACADEMIC RESEARCH
ISSN : 3049-2122
VOLUME: 02 ISSUE: 02
RECEIVED: APRIL 21 REVISED: MAY 04 ACCEPTED: MAY 23 PUBLISHED: JUNE 11

as detecting or predicting abnormal events originating from sensors and other corroborative sources.

For safety-critical applications in smart automation, monitoring must be deep and observe sensor data from possible abnormal conditions at the same time—single or combinations of faults. Deep-learning-based anomaly detection typically operates with fresh batches of incoming data ready for latest deep model-driven inferences. The current architecture doesn't support horizontal or vertical expansion with fresh batches except for a standalone solution using GPUs. A new anomaly detection framework has been created that abstracts deep-learning-based anomaly detection as a service with federated-style learning support. The monitoring architecture uses an agent-based approach with decentralized operations, allowing neighboring clusters on multiple failover systems to share data even during closed working hours in a group.



Equation 1) Distributed sensing → streaming data model (foundation)

1.1 Sensor streams

Assume M sensors/nodes. At time t, node i produces a feature vector:

$$\mathbf{x}_i(t) \in \mathbb{R}^d$$

Stack all sensors at time t into a global "state" vector:

$$\mathbf{X}(t) = [\mathbf{x}_1(t)^\top, \mathbf{x}_2(t)^\top, \dots, \mathbf{x}_M(t)^\top]^\top \in \mathbb{R}^{Md}$$

1.2 Sliding window for online monitoring

Real-time detectors often run on a rolling window of the last W samples:

$$\mathcal{W}_i(t) = \{\mathbf{x}_i(t-W+1), \dots, \mathbf{x}_i(t)\}$$

This enables **online adaptation** (mentioned for LOF in the manuscript).

1.3 Generic anomaly decision rule

Most anomaly detectors ultimately compute a score $s_i(t)$ and threshold it:

$$s_i(t) = f(\mathcal{W}_i(t)) \qquad \text{Declare anomaly if } s_i(t) > \tau_i$$

2.1. Rationale and Objectives of the Study

Owing to the growing economic burden of system downtime, maintenance errors, and safety incidents, considerable research effort has focused on real-time monitoring and predictive maintenance of industry systems, buildings, and other infrastructure. Within the smart automation domain, where such monitoring capability is particularly important, the implementation of monitoring solutions is limited because current AI techniques for real-time monitoring—especially end-to-end learning-based supervised models—require considerable labeled data and are thus expensive to develop. The study proposes scalable AI-based systems suitable for distributed data gathering, monitoring, and decision making, considered particularly in the context of monitoring tasks requiring either labeled data and long training times or complex intelligent algorithms. Both monitoring tasks integrate monitoring-related data from geographically distributed sensors. In the first task, these data are employed to detect anomalies across the entire monitoring domain; anomaly occurrences are subsequently detected at individual sensors.

JOURNAL OF INNOVATIVE ACADEMIC RESEARCH
ISSN : 3049-2122
VOLUME: 02 ISSUE: 02
RECEIVED: APRIL 21 REVISED: MAY 04 ACCEPTED: MAY 23 PUBLISHED: JUNE 11

The monitoring models for the second task deploy AI techniques for predictive maintenance. These models are trained using historical data, allowing with online detection of remaining useful life in a monitored component. Models for both tasks are developed within a horizontal-gain cluster-based architecture supporting model completion of sensor nodes of different hardware performance. Horizontal-scalability requirements are defined based on characteristics of the used AI techniques and monitoring tasks, and a virtual-device mechanism enabling efficient completion of inference requests is proposed. The monitoring systems are realized as service applications within a smart automation platform developed in the Cloud-Edge-IoT Computing framework. Practical feasibility, gain characteristics, and horizontal-scalability requirements of the proposed systems are validated through two real-life applications.

3. Architectural Principles

Advanced and Intelligent Systems (AIS) that enable distributed real-time monitoring and decision-making operate according to three basic architectural principles. The first consists of a distributed and integrated sensing infrastructure for data collection that enables process knowledge creation through different AI techniques. The second supports an AI monitoring engine that, through federated machine-learning approaches, allows for horizontal scalability—enabling a large variety of sensors to be integrated, from low-cost devices to more expensive ones. The third vertical performance-capacity principle enables greater levels of monitoring at different time horizons according to the importance of the process and associated risk.

A critical AI-based monitoring capability encompasses the detection of anomalous behaviour in complex physical processes proposed according to concepts of local global monitoring, fault detection and diagnosis fusion levels, and vertical scalability between low-dimensional models suitable for real-time applications and high-dimensional models that allow better representation of normal behaviour. Monitoring is performed through the distributed anomaly detection characteristics of the Local Outlier Factor technique with online adaptation to the evolving system conditions.

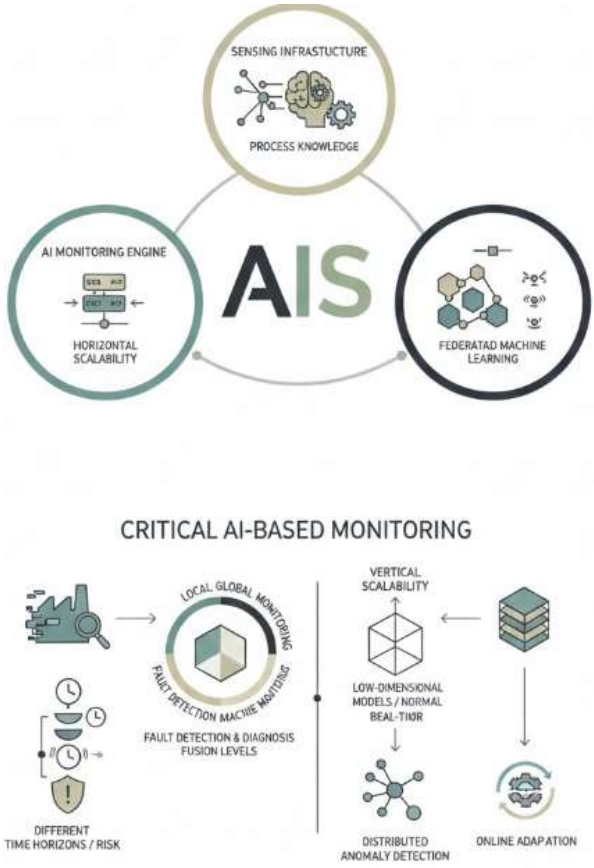


Fig 2: Federated Intelligence Frameworks: Scalable Anomaly Detection and Multi-Dimensional Monitoring in Advanced Intelligent Systems

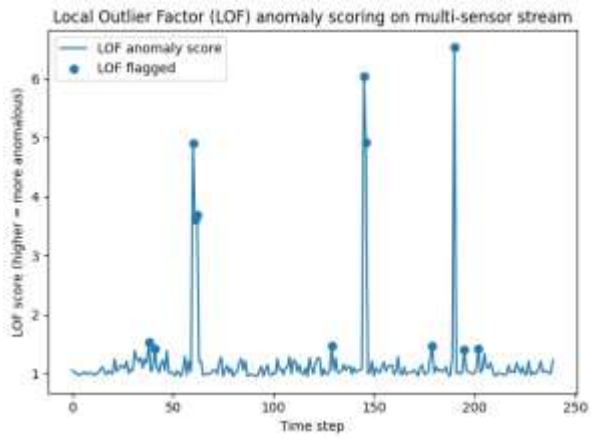
3.1. Distributed Sensing and Data Ingestion

Distributed sensing is a key enabler for smart automation systems. Industrial and other systems generate copious amounts of data though often from a

JOURNAL OF INNOVATIVE ACADEMIC RESEARCH
ISSN : 3049-2122
VOLUME: 02 ISSUE: 02
RECEIVED: APRIL 21 REVISED: MAY 04 ACCEPTED: MAY 23 PUBLISHED: JUNE 11

relatively fewer number of data sources such as industrial machines or smart buildings. Collecting data from a few geo-distributed sources via less accurate but lower-cost sensors is more economical than deploying and using highly accurate data sources everywhere. These observations motivate distributed sensing composed of one or more geo-distributed data sources for a data indication or data-sharing task.

Fusing analytic techniques like analytics and communication technology capabilities such as locality-aware content delivery and massive machine-type communication can reduce the response-time requirements of such AI-based systems. Characteristics of the data and the data-processing operations also help categorize the supported AI-based techniques into real-time, near-real-time, and latency-insensitive processing. Such a Distributed Sensing-Based AI-Pipeline (DS-AIP) framework aids in monitoring, securing, and controlling smart automation systems. Incorporating the salient principles of distributed sensing improves data governance. The geo-distributed data sources combined with a proper privacy-preservation mechanism provide multiple levels of privacy and security against both unintentional and malicious data exposure.



Equation 2) Local Outlier Factor (LOF) for distributed anomaly detection (derived fully)

Let the current data window (at some node, or aggregated features) be a set:

$$\mathcal{D} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\},\ \mathbf{x}_j \in \mathbb{R}^d$$

Choose k (neighbors). LOF is density-based: "How isolated is **x** relative to its neighbors?"

Step 1 — k-distance and k-neighborhood

Define the distance function $d(\mathbf{x}, \mathbf{y})$ (often Euclidean).

k-distance of a point **x**:

$\text{kdist}(\mathbf{x}) =$ distance from **x** to its k-th nearest neighbor

k-neighborhood:

$$N_k(\mathbf{x}) = \{\mathbf{o} \in \mathcal{D}\backslash\{\mathbf{x}\} :\ d(\mathbf{x}, \mathbf{o}) \leq \text{kdist}(\mathbf{x})\}$$

Step 2 — reachability distance (stabilizes very close points)

$$\text{reach\_dist}_k(\mathbf{x}, \mathbf{o}) = \max\{\text{kdist}(\mathbf{o}),\ d(\mathbf{x}, \mathbf{o})\}$$

Why "max"?
If **o** lies in a very dense region, its k-distance is small; using the max prevents unrealistically tiny distances from exploding density estimates.

Step 3 — local reachability density (LRD)

Average reachability distance from **x** to its neighbors:

$$\overline{r}_k(\mathbf{x}) = \frac{1}{|N_k(\mathbf{x})|} \sum_{\mathbf{o} \in N_k(\mathbf{x})} \text{reach\_dist}_k\,(\mathbf{x}, \mathbf{o})$$

Invert to get a density-like quantity:

JOURNAL OF INNOVATIVE ACADEMIC RESEARCH
ISSN : 3049-2122
VOLUME: 02 ISSUE: 02
RECEIVED: APRIL 21 REVISED: MAY 04 ACCEPTED: MAY 23 PUBLISHED: JUNE 11

$$\text{lrd}_k(\mathbf{x}) = \frac{1}{\bar{r}_k(\mathbf{x})}$$

$$= \left(\frac{1}{|N_k(\mathbf{x})|} \sum_{\mathbf{o} \in N_k(\mathbf{x})} \text{reach\_dist}_k(\mathbf{x}, \mathbf{o}) \right)^{-1}$$

Step 4 — LOF score

Compare **x**'s density to its neighbors' densities:

$$\text{LOF}_k(\mathbf{x}) = \frac{1}{|N_k(\mathbf{x})|} \sum_{\mathbf{o} \in N_k(\mathbf{x})} \frac{\text{lrd}_k(\mathbf{o})}{\text{lrd}_k(\mathbf{x})}$$

Interpretation:

- LOF ≈ 1: similar density as neighbors → normal
- LOF > 1: lower density than neighbors → outlier/anomaly

Step 5 — Online adaptation (streaming version)

To adapt online, compute LOF on a rolling window $\mathcal{W}(t)$ and update neighbors/densities as points enter/leave:

At time t:

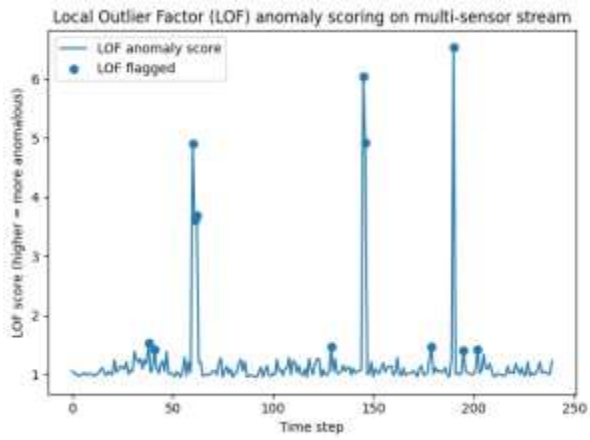
$$\mathcal{D}(t) = \{\mathbf{x}(t - W + 1), \ldots, \mathbf{x}(t)\}$$

Compute $\text{LOF}_k(\mathbf{x}(t))$ against $\mathcal{D}(t)$.

4. AI Methods for Real-Time Monitoring

Real-time analysis of continuously growing data streams can greatly support decision-making. Herein, AI methods efficiently enabling rapid detection of anomalies across data streams originating from various geographical regions are presented: Multi-Stream Temporal Convolutional Network (MSTCN) carries out anomaly detection on multiple data streams in parallel; a distribution-free ensemble learning of anomaly detectors with heterogeneous input data queues, EC3*, is proposed; a generalized framework integrating a sampling-based data-filtering technique with various classes of detectors for imbalanced data-stream environments is introduced; and a decentralized adaptation of MSTCN is developed for detecting anomalies in geographical areas with a limited number of monitoring sensors.

The role of smart AI-based systems for anomaly detection in monitoring automation systems is discussed. Such systems compete for an unknown and unpredicted supply–demand mismatch by analyzing anomalies in a large number of geographical regions concentrated in a distribution center, providing valuable information for reallocation decisions. Unsenseable demand for detection in remote areas reduces inefficiencies, and insufficient supply in a distribution cluster warns of shortages in a small geographical region. Failed TAC (Target and Compare) strategies relate to unpareto-optimal systemic conditions. Applications consider detection of anomalies in eBay sales statistics, the Germany electricity process, and NYSE trading volumes.



JOURNAL OF INNOVATIVE ACADEMIC RESEARCH
ISSN : 3049-2122
VOLUME: 02 ISSUE: 02
RECEIVED: APRIL 21 REVISED: MAY 04 ACCEPTED: MAY 23 PUBLISHED: JUNE 11

Equation 3) Gaussian Process (GP) + PCA hybrid anomaly detection (derived fully)

3.1 GP time-series model (single sensor)

Assume we observe time series pairs (t_n, y_n), $n = 1..N$.

Step 1 — Prior over functions

A Gaussian Process prior:

$$f(t) \sim \mathcal{GP}\big(m(t), k(t,t')\big)$$

Often $m(t) = 0$. The kernel k encodes smoothness/seasonality.

Observation model (noise):

$$y_n = f(t_n) + \epsilon_n, \quad \epsilon_n \sim \mathcal{N}(0, \sigma^2)$$

Step 2 — Build covariance matrices

Let $\mathbf{t} = [t_1, \dots, t_N]^\top$. Define:

$$\mathbf{K} = K(\mathbf{t}, \mathbf{t}) \in \mathbb{R}^{N \times N}, \quad K_{ij} = k\big(t_i, t_j\big)$$

Noisy covariance:

$$\mathbf{K}_y = \mathbf{K} + \sigma^2 \mathbf{I}$$

Step 3 — Posterior predictive distribution at a new time t_*

Compute:

$\mathbf{k}_* = K(\mathbf{t}, t_*) \in \mathbb{R}^N \; k_{**} = k(t_*, t_*)$

Predictive mean:

$$\mu_* = m(t_*) + \mathbf{k}_*^\top \mathbf{K}_y^{-1}\big(\mathbf{y} - m(\mathbf{t})\big)$$

Predictive variance:

$$\sigma_*^2 = k_{**} - \mathbf{k}_*^\top \mathbf{K}_y^{-1} \mathbf{k}_*$$

Step 4 — GP anomaly score from residual

Residual:

$$e(t_*) = y_* - \mu_*$$

A common standardized score (z-score):

$$z(t_*) = \frac{|e(t_*)|}{\sigma_*}$$

Flag anomaly if $z(t_*) > z_\tau$ (e.g., $z_\tau = 2$ or 3).

4.1. Anomaly Detection in Distributed Environments

Novel methods for real-time anomaly detection in distributed systems are also essential. Distributed and federated approaches have been proposed, primarily for binary classification, based on subspace learning concepts for both divergence and similarity measures. To break the limitation imposed by subspace learning, a multi-label and hierarchical approach was developed for anomaly detection, enabling not only distributed but also horizontal scaling. The framework combines hierarchical clustering using different algorithms for different clusters, multilabel prediction based on local classifiers, and a random forest classifier for refining the prediction using local predictions as additional information. The approach was validated on both synthetic and real data sets.

Anomaly detection in multidimensional continuous time series is also important for the proposed smart automation architecture, mainly for use cases that need real-time alerts. A performance-efficient approach employs Gaussian processes (GPs) to detect seasonality and predict future normal behaviour, combining these capabilities with a PCA-based global spatial model to identify global anomalies during the prediction horizon. The primary advantages of the proposed hybrid approach are its linear complexity and ability to capture time-shifting anomalies.

JOURNAL OF INNOVATIVE ACADEMIC RESEARCH
ISSN : 3049-2122
VOLUME: 02 ISSUE: 02
RECEIVED: APRIL 21 REVISED: MAY 04 ACCEPTED: MAY 23 PUBLISHED: JUNE 11

Method	Core idea	Why it fits distributed real-time monitoring	Main limitations
LOF (online)	Local density ratio outlier score	Fast, local, works with streaming windows	Sensitive to k/distance; needs drift handling
GP + PCA hybrid	Forecast + global residuals	Uncertainty + seasonality; can catch time-shift anomalies	Kernel choice; GP scalability needs approximations
Federated averaging	Aggregate model updates	Keeps raw data local; supports cross-site learning	Non-IID data; comms cost; governance overhead

5. Scalability and Performance Considerations

The proposed distributed, horizontal architecture for sensing, monitoring and decision support systems is scalable by design. Hardware and software resources can be added incrementally in an elastic manner to handle increasing loads. A vertical architecture for a monitoring and decision support tool can be derived from the distributed horizontal one. All sensing sources send their collected data to a central data lake, and a powerful computing unit in the cloud serves as the big data analytic engine.

Scalability can be considered in two ways: by adding more resources horizontally (horizontal scaling) or by replacing existing resources with stronger and faster ones vertically. A distributed decision-support tool comprised of interconnected layers can also be built. In this case, front-end monitoring agents collect and process local data. Local anomalies are detected, and notifications are issued in real time to the clients connected to that layer of the tool. Data are also ingested by the second layer of the tool, where a more powerful entity validates local anomalies and detects anomalies at a higher level of abstraction using the aggregated data coming from the local layers.

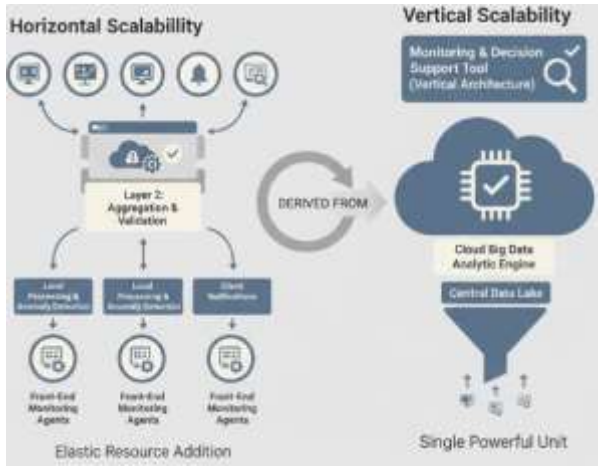


Fig 3: Elastic Intelligence: A Multi-Layered Framework for Scalable Distributed Monitoring and Hierarchical Decision Support Systems

5.1. Horizontal and Vertical Scalability

The scalability of AI methods for real-time monitoring in distributed systems will be investigated, with a focus on two main dimensions: horizontal and vertical scalability. Horizontal scalability refers to the ability to distribute training and/or inference over a set of independent nodes, so that the overall workload perceived by any single node can be reduced by distributing it to a larger number of nodes. The ability to use special-purpose computational resources, e.g. accelerators like GPUs, to increase the throughput in a

JOURNAL OF INNOVATIVE ACADEMIC RESEARCH
ISSN : 3049-2122
VOLUME: 02 ISSUE: 02
RECEIVED: APRIL 21 REVISED: MAY 04 ACCEPTED: MAY 23 PUBLISHED: JUNE 11

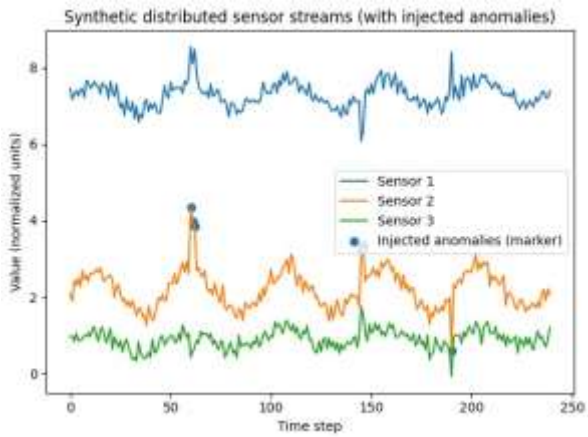
fast and cost-effective way is a particular case of horizontal scalability.

Detecting different types of anomalies in a distributed sensing environment can be formalized by means of an extreme multi-label classification problem. Many current solutions are based on locally trained classifiers. Multiple classifiers can be trained in parallel on different nodes using features from the same node, as the data and label dimensions are large and typically sparse with respect to the label dimension. Training when GPU nodes are available is faster and cheaper when the workload of the non-GPU nodes is smaller. So, as more trained classifiers are deployed, these benefits increase. Inference can be done through a single classifier or multiple classifiers in parallel, depending on the type of anomaly. An ensemble-based framework that allows combining these different inference approaches is being explored.

6. Data Governance, Security, and Privacy

Data governance frameworks for large-scale smart sensing environments ideally include security, privacy, and usability as key goals. The terms data governance and management are often used interchangeably, but management is typically used to refer to the more technical aspects of storage provisioning, data storage and retrieval. A more detailed definition of data governance indicates a framework for decision making on data management so that the data remains usable, relevant, and secure at all its life cycle phases.

Federated learning is a distributed method for training Machine Learning models across decentralized devices that hold local data samples without exchanging them. Cross-silo federated learning refers to a scenario with a small number of participants, such as institutions or companies, and where communication between parties and central server is usually secured and bounded. In this scenario, differential privacy mechanisms can be applied to the local model updates before being sent to the central server.



Equation 4) Federated learning + differential privacy (derived fully)

4.1 Local training objective

Client/node i has dataset D_i. Model parameters are $\mathbf{w}$.

Local empirical risk:

$$F_i(\mathbf{w}) = \frac{1}{|D_i|} \sum_{(x,y)\in D_i} \ell\,(\mathbf{w}; x, y)$$

Global objective (weighted by data size):

$$F(\mathbf{w}) = \sum_{i=1}^{K} \frac{|D_i|}{\sum_j |D_j|} F_i(\mathbf{w})$$

4.2 Federated Averaging (FedAvg) update rule

At round t, server broadcasts $\mathbf{w}^{(t)}$.
Each client runs E steps of SGD:

Client update (one SGD step):

JOURNAL OF INNOVATIVE ACADEMIC RESEARCH
ISSN : 3049-2122
VOLUME: 02 ISSUE: 02
RECEIVED: APRIL 21 REVISED: MAY 04 ACCEPTED: MAY 23 PUBLISHED: JUNE 11

$$\mathbf{w} \leftarrow \mathbf{w} - \eta \nabla F_i(\mathbf{w})$$

After local training, client sends either $\mathbf{w}_i^{(t+1)}$ or the delta:

$$\Delta_i^{(t)} = \mathbf{w}_i^{(t+1)} - \mathbf{w}^{(t)}$$

Server aggregates (weighted average):

$$\mathbf{w}^{(t+1)} = \sum_{i=1}^{K} \frac{|D_i|}{\sum_j |D_j|} \mathbf{w}_i^{(t+1)}$$

or equivalently

$$\mathbf{w}^{(t+1)} = \mathbf{w}^{(t)} + \sum_{i=1}^{K} \frac{|D_i|}{\sum_j |D_j|} \Delta_i^{(t)}$$

6.1. Federated and Privacy-Preserving Analytics

Machine learning and data mining often require centralized data lakes that collect the data from their premises and transfer it to a centralized location for analysis. However, a new paradigm of federated learning-based analytics has recently emerged that allows organizations to perform machine learning on models located on the respective premises without accessing their data, significantly mitigating the risks of exposing sensitive data. Essentially, a detailed model is shared within a federated group of participants where the cores of the models are exchanged while the layers responsible for providing identification features remain hidden. Federated learning analysis models also enable participants to extract knowledge from their datasets without the risk of exposing sensitive information. Federated transfer learning can combine data from different premises while preserving the privacy of sensitive data.

The concern for data privacy can further be integrated into these concepts by relying on secure multiparty computation techniques. In this secure multiparty computation model, all data at all participant locations is protected by secure multiparty computation in such a way that the information shared with the party providing the core model will be inaccessible to them and will have to be decrypted by all other premises at the end of the process in order to finalize the training of the core model. This is a complex and time-consuming process when the number of participants is large, though, as it must be executed by every premise. Moreover, federated analytics are usually limited to a particular analytical task. Privacy-preserving data mining allows analytics to be executed on sensitive data without exposing the data to each other.

7. Conclusion

Scalable AI-Based Systems for Distributed Real-Time Monitoring and Decision-Making in Smart Automation

Increasingly distributed operations, sensor proliferation and the requirement to monitor an evergrowing number of entities have created demands that typical centralized AI-based solutions cannot meet. These and additional drivers of scalability, such as the sheer volume of data, the need to serve an ever-growing number of users, and the need to provide data locally while observing privacy and data protection regulations, must be recognized and addressed. Once these drivers have been identified, it becomes possible to assess whether horizontal or vertical scaling is the best answer. A complementary approach is to move from generalization to specialization by analyzing distribution and segregation at lower hierarchical levels.

The synthesis therefore highlights scalable AI-based solutions that address distributed real-time monitoring and decision-making requirements for smart automation including IIoT, smart cities, smart buildings, smart grids and smart healthcare. AI-based methods, services and operators explicitly designed to work in a distributed or decentralized manner are targeted. To address the fact that security, privacy and guaranteed data access are becoming increasingly important, and on ensuring compliance with legal and ethical regulations, the focus

JOURNAL OF INNOVATIVE ACADEMIC RESEARCH
ISSN : 3049-2122
VOLUME: 02 ISSUE: 02
RECEIVED: APRIL 21 REVISED: MAY 04 ACCEPTED: MAY 23 PUBLISHED: JUNE 11

is on data-governance, security and privacy-triggered distribution. Scalability, and in particular horizontal scalability, is a first-order design concern and, where appropriate, the trade-off with vertical scalability is explored.

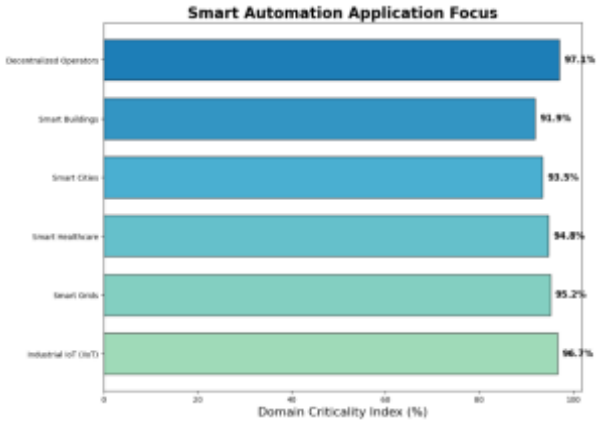


Fig 4: Smart Automation Application Focus

7.1. Summary of Key Insights and Future Directions

In summary, the evidence-based synthesis presented in this study emphasizes scalable AI-based systems for distributed real-time monitoring and decision making in smart automation. The central argument states that emerging concepts, standards, and technology frameworks make it feasible to extend both monitoring and actuation components of smart automation systems to broader, distributed environments. By given attention to monitoring of various types, including process, quality, status, cyber-security, performance, and fault, as well as anomaly detection in large numbers of connected devices, robots, and production, transportation, or service systems, these systems can provide the timely and collaborative information needed for smarter decision making, including in-situ or in-cycle modification of normal operational parameters.

While the recent results are very encouraging, several challenges and research directions remain. The rest of the chapter therefore outlines some pointers and ideas for possible future exploration and development. Scalability of both sensing and data ingestion as well as the monitoring algorithms and methods themselves is a key consideration. Horizontal scalability covers the growing number of connected devices, robots, and component systems — such as physical objects municipalities or smart cities, sensors or actuators integrated on smart physical objects, pedestrians — and the gradual development of larger-scale sensor-actuator networks in control applications. At the same time, there is also demand for vertical scalability of the architecture to allow its potential deployment in local, peripheral, or edge computing environments, with minimal or no dependence on powerful central-resource systems capable of immediately training larger detection models.

8. References

1. Nguyen, D. C., Ding, M., Pathirana, P. N., Seneviratne, A., Li, J., & Poor, H. V. (2021). Federated learning for Internet of Things: A comprehensive survey. IEEE Communications Surveys & Tutorials, 23(3), 1622–1658.
2. Kherbache, M., Maimour, M., & Rondeau, E. (2021). When digital twin meets network softwarization in the Industrial IoT: Real-time requirements case study. Sensors, 21(24), 8194.
3. Kolla, S. K., & Mangalampalli, B. M. (2024). Edge-Based Deep Learning Systems for Point-of-Care Diagnostic Intelligence. Journal of Neonatal Surgery, 13(1), 2387-2399.
4. Ray, P. P. (2022). A review on TinyML: State-of-the-art and prospects. Journal of King Saud University—Computer and Information Sciences, 34(4), 1595–1623.
5. Jiang, Z., Guo, Y., & Wang, Z. (2021). Digital twin to improve the virtual-real integration of industrial IoT. Journal of Industrial Information Integration, 22, 100196.
6. Radanliev, P., De Roure, D., Nicolescu, R., Huth, M., Santos, O., & Montalvo, R. M. (2022). Digital

JOURNAL OF INNOVATIVE ACADEMIC RESEARCH
ISSN : 3049-2122
VOLUME: 02 ISSUE: 02
RECEIVED: APRIL 21 REVISED: MAY 04 ACCEPTED: MAY 23 PUBLISHED: JUNE 11

twins: Artificial intelligence and the IoT cyber-physical systems in Industry 4.0. International Journal of Intelligent Robotics and Applications, 6, 171–185.
7. Inala, R., & Somu, B. (2024). Agentic ai in retail banking: Redefining customer service and financial decision-making. Journal of Artificial Intelligence and Big Data Disciplines, 1(1), 1-19.
8. Khan, L. U., Pandey, S. R., Tran, N. H., Saad, W., Han, Z., & Hong, C. S. (2021). Federated learning for edge networks: Resource optimization and incentive mechanism. IEEE Communications Magazine, 59(11), 88–93.
9. Abreha, H. G., Hayajneh, M., & Serhani, M. A. (2022). Federated learning in edge computing: A systematic survey. Sensors, 22(2), 450.
10. Gottimukkala, V. R. R. (2024). Federated Learning Approaches for Fraud Detection in International Payment Systems. https://www. jisem-journal. com/download/118_JISEM. pdf.
11. Duan, Q., Hu, S., Deng, R., & Lu, Z. (2022). Combined federated and split learning in edge computing for ubiquitous intelligence in Internet of Things: State-of-the-art and future directions. Sensors, 22(16), 5983.
12. Schizas, N., Karras, A., Karras, C., & Sioutas, S. (2022). TinyML for ultra-low power AI and large scale IoT deployments: A systematic review. Future Internet, 14(12), 363.
13. Singh, R., & Gill, S. S. (2023). Edge AI: A survey. Internet of Things and Cyber-Physical Systems, 3, 71–92.
14. Inala, R. (2023). AI-powered investment decision support systems: Building smart data products with embedded governance controls. Journal for ReAttach Therapy and Developmental Diversities, 6(10), 2251-2266.
15. Bourechak, A., Zedadra, O., Kouahla, M. N., Guerrieri, A., Seridi, H., & Fortino, G. (2023). At the confluence of artificial intelligence and edge computing in IoT-based applications: A review and new perspectives. Sensors, 23(3), 1639.
16. Silva, L. G. F. da, Sadok, D. F. H., & Endo, P. T. (2023). Resource optimizing federated learning for use with IoT: A systematic review. Journal of Parallel and Distributed Computing, 175, 92–108.
17. Kolla, T. (2024). Graph Neural Networks for HCC Risk Adjustment and Interoperability. International Journal of Science, Research and Technology, 7(6), 13244-13255.
18. Campolo, C., Genovese, G., Singh, G., & Molinaro, A. (2023). Scalable and interoperable edge-based federated learning in IoT contexts. Computer Networks, 223, 109576.
19. Ahmad, S., Shakeel, I., Mehfuz, S., & Ahmad, J. (2023). Deep learning models for cloud, edge, fog, and IoT computing paradigms: Survey, recent advances, and future directions. Computer Science Review, 49, 100568.
20. Singh, R., & Gill, S. S. (2023). Edge AI: A survey. Internet of Things and Cyber-Physical Systems, 3, 71–92.
21. Davuluri, P. S. L. (2023). AI-Augmented Sanctions Screening: Enhancing Accuracy and Latency in Real Time Compliance Systems. AI-Augmented Sanctions Screening: Enhancing Accuracy and Latency in Real Time Compliance Systems (December 15, 2023).
22. Li, H., Ge, L., & Tian, L. (2024). Survey: Federated learning data security and privacy-preserving in edge-Internet of Things. Artificial Intelligence Review, 57, 130.
23. Savaglio, C., Mazzei, P., & Fortino, G. (2024). Edge intelligence for Industrial IoT: Opportunities and limitations. Procedia Computer Science, 232, 397–405.
24. Ning, Z., Tolba, A., & Kong, X. (2022). Edge intelligence for 6G-enabled Industrial Internet of Things. Wireless Communications and Mobile Computing, 2022.
25. Sharma, A., Kosasih, E., Zhang, J., Brintrup, A., & Calinescu, A. (2022). Digital twins: State of the art theory and practice, challenges, and open research questions. Journal of Industrial Information Integration, 30, 100383.

JOURNAL OF INNOVATIVE ACADEMIC RESEARCH
ISSN : 3049-2122
VOLUME: 02 ISSUE: 02
RECEIVED: APRIL 21 REVISED: MAY 04 ACCEPTED: MAY 23 PUBLISHED: JUNE 11

26. Inala, R. Designing Scalable Technology Architectures for Customer Data in Group Insurance and Investment Platforms.
27. Ning, Z., Dong, P., Wang, X., Hu, X., Lv, J., Guo, Y., & Wang, J. (2022). When deep reinforcement learning meets 5G-enabled edge computing. IEEE Network, 36(2), 46–52.
28. Zhao, Y., Zhao, J., Jiang, L., Wang, R., & Tan, H. (2022). Federated learning for edge intelligence in Internet of Things: A survey. IEEE Internet of Things Journal, 9(18), 17195–17214.
29. Zaidi, S. A. R., Hayajneh, A. M., Hafeez, M., et al. (2022). Unlocking edge intelligence through tiny machine learning (TinyML). IEEE Access, 10, 100867–100882.
30. Ray, P. P. (2022). A review on TinyML: State-of-the-art and prospects. Journal of King Saud University—Computer and Information Sciences, 34(4), 1595–1623.
31. Singh, R., Gill, S. S., & others. (2023). AI-based fog and edge computing: A systematic review, taxonomy and future directions. Internet of Things, 21, 100674.
32. Han, C., Li, T., Chen, Q., Wu, Y., & Qin, J. (2024). Distributed and collaborative lightweight edge federated learning for IoT zombie devices detection. ACM Transactions on Sensor Networks.
33. Campolo, C., Genovese, G., Singh, G., & Molinaro, A. (2023). Scalable and interoperable edge-based federated learning in IoT contexts. Computer Networks, 223, 109576.
34. Savaglio, C., Mazzei, P., & Fortino, G. (2024). Edge intelligence for Industrial IoT: Opportunities and limitations. Procedia Computer Science, 232, 397–405.
35. Kolla, S. K., & Reddy, V. A. R. (2024). Evaluating Cloud-Native vs. Hybrid Architectures for Health Benefit Administration Systems. International Journal of Medical Toxicology and Legal Medicine, 27(5), 1042-1053.
36. Ahmad, S., Shakeel, I., Mehfuz, S., & Ahmad, J. (2023). Deep learning models for cloud, edge, fog, and IoT computing paradigms: Survey, recent advances, and future directions. Computer Science Review, 49, 100568.
37. Bourechak, A., Zedadra, O., Kouahla, M. N., Guerrieri, A., Seridi, H., & Fortino, G. (2023). At the confluence of artificial intelligence and edge computing in IoT-based applications: A review and new perspectives. Sensors, 23(3), 1639.
38. Jiang, Z., Guo, Y., & Wang, Z. (2021). Digital twin to improve the virtual-real integration of industrial IoT. Journal of Industrial Information Integration, 22, 100196.