Decentralized Tiny Ensemble Learning for Privacy-Preserving Compliance Prediction in Operating Room Environments

Imane Abrya *, Safae Bouhaddou *, Meryeme Ayache * †, Nidal Nasser † and Amal Bouayad ‡
* STRS Lab, National Institute of Posts and Telecommunications INPT Rabat, Morocco
† College of Engineering & Advanced Computing, Alfaisal University, Riyadh 11533, Saudi Arabia.
‡ Hospital Center Provincial Lalla Aicha, Skhirat-Temara, Morocco.
abryaimane@gmail.com; Safaeb.ine@gmail.com; ayache@inpt.ac.ma;
nnasser@alfaisal.edu; abouayad@sante.gov.ma

***Abstract*—Healthcare environments, particularly operating rooms, face significant challenges in maintaining optimal conditions to prevent postoperative infections. These challenges include the absence of real-time monitoring systems, the limited availability of labeled medical data, and the difficulty of deploying intelligent solutions in resource-constrained settings. Additionally, there is a growing need to explore emerging technologies in IoT and TinyML to modernize healthcare monitoring and address existing gaps in system intelligence, scalability, and privacy.**

This paper presents a privacy-preserving intelligent system designed to predict the compliance of operating room environments with ISO standards by continuously monitoring key parameters such as temperature, humidity, air quality, and pressure. The system employs a decentralized Tiny ensemble learning architecture, utilizing ESP32 microcontrollers as edge learners and a Raspberry Pi 4 as the aggregator node. This approach improves prediction accuracy to 98.74% while preserving data privacy through localized processing. To address data scarcity, recent synthetic data generation techniques from the literature are employed to create comprehensive datasets for training and validation.

***Index Terms*—Operating Room Monitoring, Smart Healthcare Systems, TinyML, Ensemble Learning, IoMT, Edge AI, On-board Training.**

I. Introduction

Artificial Intelligence (AI) has become a transformative force in healthcare, allowing for better diagnostics, personalized treatment plans, and predictive analytics. AI began to be applied in this field during the late 20th century, notably with early expert systems such as MYCIN [1] developed in the 1970s, which assisted in diagnosing infections and suggesting appropriate antibiotics. These early systems laid the groundwork for today's advanced machine learning and deep learning methods, which now significantly improve patient safety, especially in the operating room.

With the rise of Industry 5.0, the emphasis has shifted towards creating human-centered, collaborative AI systems that encourage a partnership between human beings and machines [2]. Techniques such as ensemble learning, which are essential to this concept, enable various healthcare organizations or devices to collaboratively enhance model performance while safeguarding data privacy and security. These approaches foster distributed intelligence and support immediate decision-making in critical medical settings [3].

Simultaneously, the Internet of Things (IoT) has emerged as a key enabler in the digital transformation of healthcare, connecting a vast and growing network of devices that collect and transmit critical health and environmental data [4]. IoT-powered tools facilitate real-time monitoring and intervention, improving care delivery across hospitals, clinics, and even at home. This expansion is fueled by the rapid adoption of smart technologies while around 13.8 billion devices were connected globally in 2022, this number grew to approximately 17.7 billion by 2024 and is expected to exceed 40 billion by 2034 [5].

Despite the significant advancements in AI and IoT technologies, their integration into operating room (OR) environments is still lacking. Most existing systems are either centralized, lack real-time decision-making capabilities, or rely on only a few isolated features for prediction neglecting key environmental parameters essential for ensuring surgical safety. In particular, current approaches often overlook the combined effect of critical variables such as temperature, humidity, air quality, and pressure. This paper addresses these limitations by introducing a smart, privacy-preserving monitoring system that leverages TinyML and ensemble learning to ensure ISO-compliant operating room conditions in real-time, using decentralized edge computing.

The remainder of this paper is structured as follows. Section II presents recent advances in AI-driven OR monitoring systems and highlights existing limitations. Section III provides the necessary background on key enabling technologies, including ensemble learning, IoT, MQTT, and TinyML. Section IV details the proposed architecture, edge deployment, and model training methodology. Section V discusses the experimental results and evaluates the system's performance. Finally, Section VI summarizes the main contributions and outlines directions for future work.

979-8-3503-9281-4/25/$31.00 ©2025 IEEE

II. Related Work

Operating rooms (ORs) are critical environments within healthcare systems that demand strict adherence to environmental and procedural standards to ensure patient safety and surgical success. However, maintaining optimal conditions in ORs remains a persistent challenge due to the dynamic nature of surgical procedures and limitations in real-time environmental monitoring [6]. Several studies have addressed the importance of maintaining controlled environments in operating rooms to ensure patient safety and reduce the risk of surgical site infections (SSIs). For instance, Gola et al. [7] emphasized the impact of thermal and air quality conditions on surgical outcomes, highlighting the need for continuous monitoring systems. Similarly, Baqer et al. [8] proposed an IoT-based environmental sensing framework that tracks temperature, humidity, and airborne contaminants in real-time to support compliance with healthcare regulations. Additionally, Ghazal et al. [9] explored the integration of smart sensors and AI algorithms in operating theaters to detect deviations from environmental norms and provide early warnings. These studies highlight the increasing demand for intelligent, standards-compliant monitoring systems that can operate effectively in complex and resource-constrained clinical environments.

In response to these challenges, recent efforts have shifted toward predictive modeling and intelligent control systems that leverage artificial intelligence (AI) and machine learning (ML) to enhance proactive decision-making and environmental stability in ORs. Jamali et al. developed a neuro-fuzzy decision support system that integrates fuzzy inference systems (FIS) and adaptive neuro-fuzzy inference systems (ANFIS), achieving validation accuracies of 97.3% and 95% for humidity and particle prediction, respectively [10]. Their hybrid system, based on Mamdani and Takagi-Sugeno-Kang models, demonstrated effective control of airflow to reduce contamination risks in ORs. Similarly, Colella et al. designed a FIS that utilizes parameters such as temperature, humidity, particle count, and staff movement to evaluate air quality and infection risk, thereby contributing to optimized HVAC system regulation and enhanced control of the surgical environment [11].

Beyond fuzzy systems, ensemble-based machine learning models have shown promise in healthcare operations and predictive analytics. Chu et al. applied models such as XGBoost, Random Forest, and artificial neural networks to predict OR usage time, illustrating the superior performance of ensemble methods in managing heterogeneous datasets and improving scheduling efficiency [12]. Arad et al. used a Random Forest classifier to identify risk factors for surgical "Never Events," highlighting the robustness of bagging techniques in handling imbalanced datasets [13]. These findings support the applicability of ensemble learning in clinical decision support, especially where model reliability and generalization are crucial.

Despite these advancements, gaps remain in leveraging ensemble methods for comprehensive prediction of dynamic OR environments. Existing studies often focus on isolated parameters such as particle concentration or room usage without integrating all key environmental variables. Furthermore, real-world datasets exhibit high variance across parameters like temperature, humidity, and air quality, which can compromise the reliability of single-model approaches. Ensemble bagging methods offer the advantage of reducing variance and combining multiple weak learners to enhance prediction accuracy [14]. Additionally, the potential for holistic modeling of the OR environment has not yet been fully explored.

This paper addresses those gaps by applying ensemble learning to jointly predict and monitor a set of interdependent environmental features, supporting more stable and intelligent control of surgical environments.

III. Background and Overview

This section offers a foundational overview of the key technologies central to this paper, including ensemble learning, IoT, the MQTT protocol, and TinyML, with particular emphasis on tools like TensorFlow Lite. These technologies collectively enable the creation of distributed, secure, and intelligent systems that can efficiently process sensitive data while meeting the performance and security demands of IoT environments.

A. TinyML

The rapid advancement of Internet of Things (IoT) technologies has revolutionized healthcare by enabling continuous monitoring and real-time data collection through interconnected devices [15]. In particular, the Internet of Medical Things (IoMT) has emerged as a specialized subset of IoT, focusing on medical applications that improve patient care and operational efficiency in clinical environments. However, deploying machine learning models directly on these edge devices is often constrained by their limited computational power and energy resources [16]. To address this, Tiny Machine Learning (TinyML) [17] has gained significant attention as it enables the execution of lightweight, efficient AI models on resource-constrained hardware such as microcontrollers and embedded systems, allowing for low-latency, energy-efficient, and bandwidth-conscious inference directly at the edge [18].

TensorFlow Lite [19] is used to convert large models into compact and efficient formats optimized for deployment and training on low-power devices, making it particularly well-suited for TinyML applications at the edge. Furthermore, to transfer TinyML models between edges, MQTT (Message Queuing Telemetry Transport) [20] protocol is often used. In fact, MQTT is a lightweight, open-source messaging protocol designed for efficient and reliable machine-to-machine communication in resource-constrained IoT systems, using a broker to manage message publishing and subscribing with adjustable Quality of Service levels.

B. Ensemble Learning

To further enhance the accuracy and robustness of edge-based inference in resource-constrained environments, ensemble learning techniques are often employed alongside TinyML, combining multiple models to improve overall predictive performance [21]. In fact, ensemble learning methods, particularly bagging techniques, have been widely adopted to enhance predictive performance and robustness by aggregating multiple weak learners into a strong composite model. This approach not only improves accuracy but also increases the resilience of the system against noisy or incomplete data, a common challenge in IoMT environments. Together, these technologies provide a promising foundation for developing intelligent, privacy-preserving healthcare monitoring systems capable of operating reliably at the network edge.

IV. Proposed System

This section presents the design of our smart environmental monitoring system for operating rooms (ORs), which combines ensemble learning, synthetic data generation, edge AI deployment, and lightweight communication protocols.

A. *System Architecture and Component Interactions*

The proposed system implements a bagging-based ensemble learning architecture, consisting of a Raspberry Pi 4 (RPi4) as the central coordination and aggregation node, and three ESP32 microcontrollers serving as independent inference units, as shown in Fig. 1. The RPi4 loads a preprocessed dataset, splits it into 80% training and 20% testing using Scikit-learn, and saves the test set for consistent evaluation.

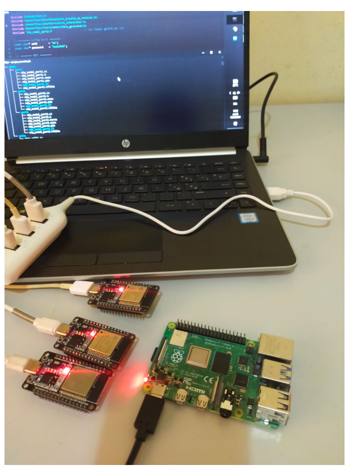

Fig. 1. Experimental Prototype of The Proposed System

Following the bagging paradigm (Fig. 2), the RPi4 generates three bootstrapped training subsets and trains a separate MLP model on each. Each model is validated using an internal 80/20 split and 5-fold cross-validation to ensure generalization. The trained models are converted from PyTorch to quantized TFLite format and subsequently transformed into C header files (.h) for deployment on ESP32 devices.

Each ESP32 receives a distinct model and the same test set via MQTT. They perform local inference and return accuracy scores to the RPi4, which aggregates the results to compute final ensemble performance. This design achieves lightweight communication, real-time edge inference, and enhanced robustness through model diversity and parallelism.

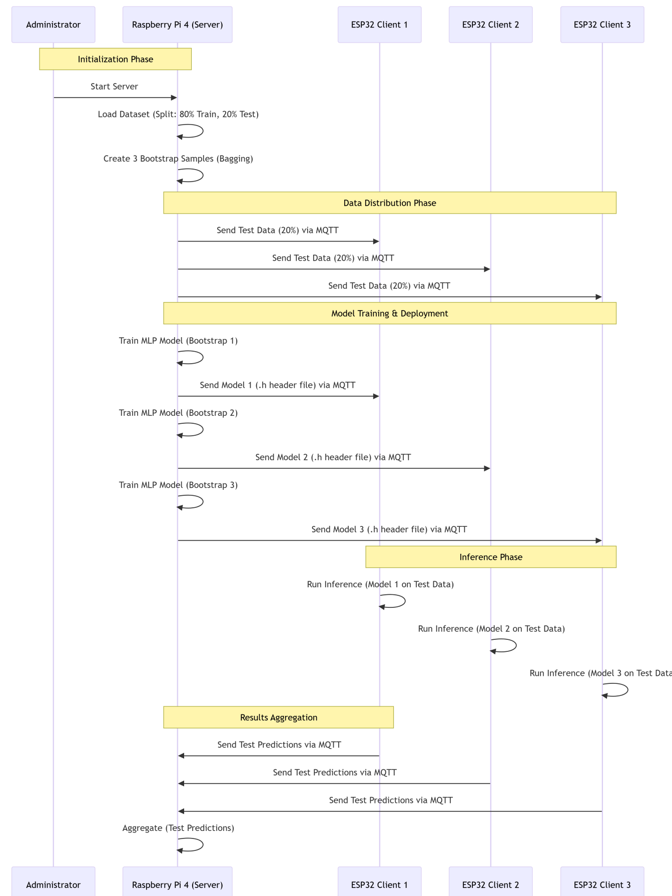

Fig. 2. Data Flow Sequence Diagram between Raspberry Pi and ESP32 Nodes

B. *Synthetic Dataset Generation*

Given the lack of large labeled datasets in the medical domain, especially concerning OR environmental conditions, we developed a synthetic dataset generation framework that produces statistically valid and medically relevant data samples.

Our final dataset consists of 60,000 samples generated using a combination of statistical modeling techniques:

- **Temperature and CO_2**: Modeled using Gaussian Mixture Models (GMMs) to capture multi-modal behavior around clinical operating room standards (e.g., 18–24°C).
- **Humidity, Pressure, and Ventilation**: Simulated using uniform distributions constrained by the hospital environmental ranges.
- **Air Quality Parameters (PM2.5, PM10, COV)**: Generated using Gaussian Multivariate Copulas to preserve natural dependencies among pollutants.

To label the generated data, we applied a hierarchical rule-based classification inspired by ISO 14644 standards [22]. Samples were categorized as Normal (all parameters within safe limits), Warning (values close to critical thresholds), or

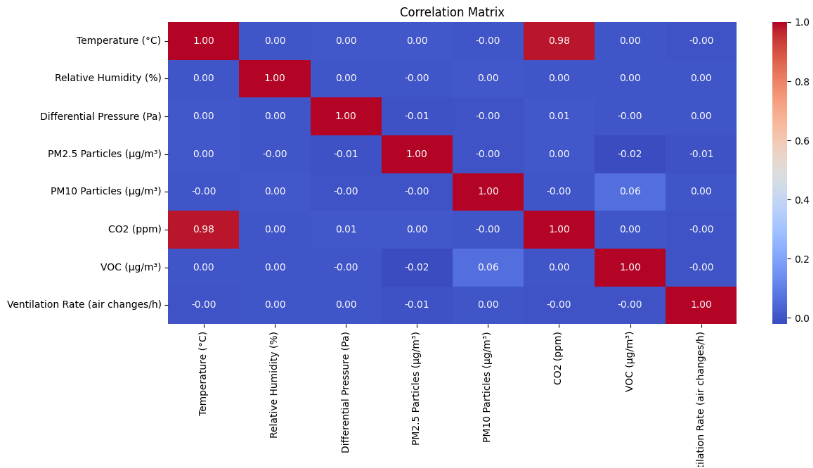

Fig. 3. Correlation Matrix of The Synthetic Data

Critical (one or more parameters exceeding safety thresholds). This labeling strategy ensures interpretability of the dataset and statistical realism.

C. Data Preprocessing and Model Preparation

To prepare the data for training, we normalized all input features using `MinMaxScaler` from Scikit-learn, which maps values to the [0,1] range. This step is essential for ensuring numerical stability during MLP training. The preprocessed dataset was saved to enable consistent use during model training and inference.

D. Model Training, Conversion, and Deployment

Each bootstrapped subset was used to train an MLP using PyTorch. The model architecture consists of:

- Input layer matching the number of features
- Two hidden layers with ReLU activations and 20% dropout
- Output layer with 3 units (corresponding to classification categories)

Training was performed using cross-entropy loss, the Adam optimizer (learning rate = 10^{-3}), weight decay for regularization, 50 epochs, and a batch size of 32. Post-training, each model was evaluated on the same test dataset to ensure a fair comparison.

Each model was then converted through the following steps to ensure compatibility with ESP32 devices:

1) PyTorch $\rightarrow$ ONNX $\rightarrow$ TensorFlow SavedModel
2) (Optionally) SavedModel $\rightarrow$ Keras v3 format
3) SavedModel $\rightarrow$ TFLite (.tflite)
4) TFLite $\rightarrow$ C Array (.cc) $\rightarrow$ Header File (.h)

Header files were embedded into ESP32 firmware, enabling real-time inference at the edge. The system design ensures model diversity, parallel inference, and low-latency communication via MQTT while maintaining high accuracy and robustness in environmental classification.

V. Performance Results

A. Evaluation of the Synthetic Dataset

To validate the statistical coherence and modeling quality of our 60,000-sample synthetic dataset, we performed visual and quantitative analyses. The class distribution exhibits a realistic imbalance aligned with real-world OR monitoring, approximately 24,500 Critical, 23,500 Warning, and 12,000 Normal samples. This distribution benefits machine learning by emphasizing risk-prone scenarios, which are more frequent in dynamic surgical environments.

The correlation matrix (Fig. 3) confirms the effectiveness of our modeling approach. A strong correlation ($r = 0.98$) exists between temperature and CO_2, reflecting their joint generation via GMM. Low or near-zero correlations between other variables (humidity, pressure, ventilation) affirm their independent simulation using uniform distributions. Dependencies among air pollutants were preserved without introducing spurious linearity($r = 0.06, r = 0.02$), as shown by moderate but realistic correlations among PM2.5, PM10, and COV.

Distribution visualizations further support these findings. Violin plots (Fig. 4) reveal expected patterns: bimodal distributions in temperature, trimodal in CO_2, uniform spreads in pressure and ventilation, and long-tailed behavior in particulate matter variables. These visual results confirm that the dataset captures meaningful clinical variability. Class-wise plots highlight distinct separation: Critical states appear at extremes, Warning states cluster near thresholds, and Normal states lie within safe operating ranges.

B. Bagging-Based Ensemble Learning Evaluation

The proposed bagging-based ensemble system, composed of three independently trained MLP models deployed on ESP32 nodes, achieved a final prediction accuracy of *98.74%*, demonstrating strong generalization and robustness. The diversity among models, introduced via bootstrapped training subsets, helped reduce overfitting and increased fault tolerance. Parallel inference on edge devices ensured low-latency predictions,

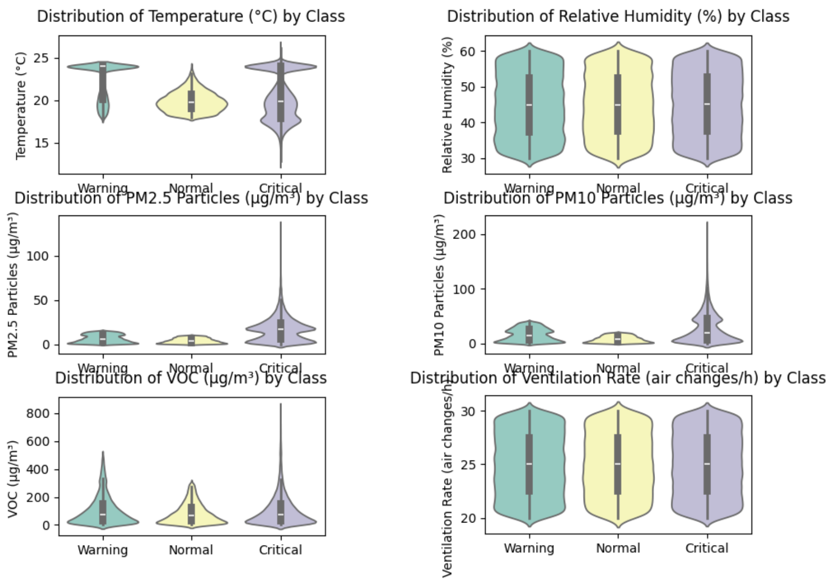

Fig. 4. Violin Plot of The Synthetic Data

while lightweight communication using MQTT maintained system efficiency, validating the system's applicability in real-time, privacy-preserving OR monitoring under resource constraints.

To further evaluate robustness, the ensemble's performance was tested under varying levels of Gaussian noise (0.025 to 0.2). As shown in Fig. 5, the ensemble consistently outperformed individual learners, maintaining an accuracy of *79.6%* at the highest noise level, compared to *78.4%* for a single ESP32 model, highlighting its capacity to mitigate variance in noisy conditions. Additionally, hyperparameter tuning of the MLP architecture revealed that the optimal configuration was achieved with 64 neurons and a dropout rate of 0.1, resulting in a test accuracy of *97.2%* as illustrated in Fig. 6. These results underscore the effectiveness of bagging in enhancing predictive stability, noise resilience, and overall model performance in complex operating room environments.

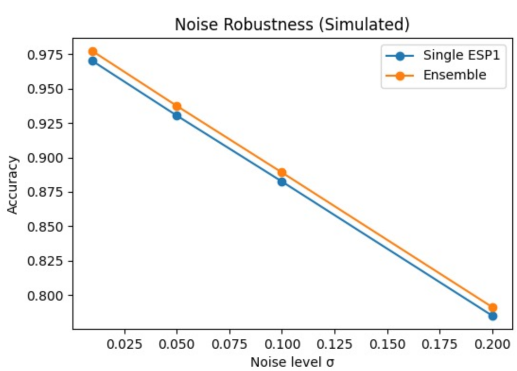

Fig. 5. Robustness Under Noisy Conditions

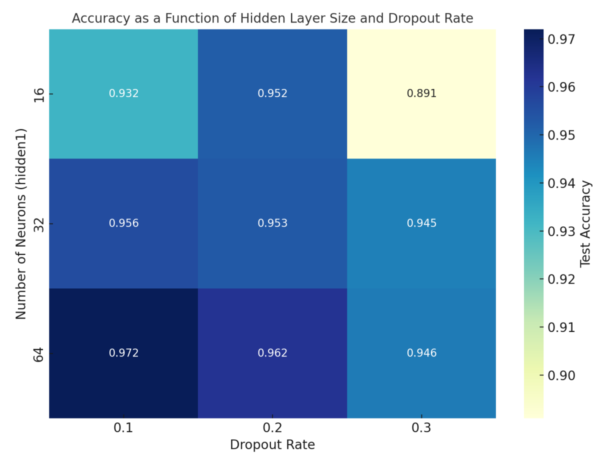

Fig. 6. Accuracy for Different Hidden Dropout parameters

C. Discussion

While the proposed TinyML ensemble system demonstrates promising results with 98.74% accuracy, several limitations warrant careful consideration for both the modeling approach and synthetic data generation strategy.

The ensemble model gave only a slight accuracy boost (+0.2%) compared to the best single model because the three ESP32 models were too similar. They were trained the same way, using similar data, so their predictions were almost always the same. This limits the benefit of combining them. Also, using three models instead of one makes the system about three times slower, which could be a problem in real-time situations like emergencies in the operating room.

Although the synthetic dataset is statistically well-constructed, it hasn't been compared to real operating room

(OR) data. It is based on assumptions that may not fully capture the complex and changing conditions found in actual ORs, such as equipment heat, staff movement, or procedure types. Also, the labeling method, based on ISO standards, simplifies real regulations, which often involve more detailed and context-specific factors than simple threshold rules can cover.

VI. Conclusion

This paper presented a privacy-preserving intelligent monitoring system for operating rooms, leveraging TinyML and bagging-based ensemble learning on edge devices. By integrating ESP32 microcontrollers with a Raspberry Pi 4 aggregator, the system enables real-time inference, decentralized processing, and robust classification of environmental conditions in compliance with ISO standards.

To address the challenge of data scarcity, a synthetic dataset comprising 60,000 samples was generated using a combination of Gaussian Mixture Models, uniform distributions, and multivariate copulas. The resulting dataset exhibited realistic statistical properties and meaningful class separability, making it suitable for effective machine learning training.

Experimental results demonstrated that the ensemble model achieved a high classification accuracy of 98.74%, showing enhanced robustness under noisy conditions and improved generalization through model diversity. These findings validate the feasibility of deploying intelligent monitoring systems in resource-constrained medical environments.

Future work will focus on integrating real-time data from actual operating room sensors, extending the model to include anomaly detection, and exploring federated learning to enable cross-hospital collaboration without compromising patient data privacy.

Acknowledgment

The authors would like to thank Alfaisal University for funding and supporting this work. They also extend their gratitude to the National Institute of Posts and Telecommunications (INPT) and the Hospital Center Provincial Lalla Aicha for their collaboration and valuable support.

References

[1] E. H. Shortliffe, R. Davis, S. G. Axline, B. G. Buchanan, C. C. Green, and S. N. Cohen, "Computer-based consultations in clinical therapeutics: explanation and rule acquisition capabilities of the mycin system," *Computers and biomedical research*, vol. 8, no. 4, pp. 303–320, 1975.

[2] N. Taimoor and S. Rehman, "Reliable and resilient ai and iot-based personalised healthcare services: A survey," *IEEE Access*, vol. 10, pp. 535–563, 2021.

[3] N. Chaudhary, V. Gupta, K. Sandhir, R. Gupta, S. Chhabra, and A. K. Singh, "Privacy preserving ensemble learning classification model for mental healthcare," *2022 Seventh International Conference on Parallel, Distributed and Grid Computing (PDGC)*, pp. 513–518, 2022.

[4] F. Alshehri and G. Muhammad, "A comprehensive survey of the internet of things (iot) and ai-based smart healthcare," *IEEE access*, vol. 9, pp. 3660–3678, 2020.

[5] "IoT connections worldwide 2034 — statista — statista.com." Available at: https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/. [Accessed 17-06-2025].

[6] K. Mohammed, A. Zaidan, B. Zaidan, O. S. Albahri, M. Alsalem, A. S. Albahri, A. Hadi, and M. Hashim, "Real-time remote-health monitoring systems: a review on patients prioritisation for multiple-chronic diseases, taxonomy analysis, concerns and solution procedure," *Journal of medical systems*, vol. 43, pp. 1–21, 2019.

[7] M. Gola, G. Settimo, and S. Capolongo, "Indoor air quality in inpatient environments: a systematic review on factors that influence chemical pollution in inpatient wards," *Journal of Healthcare Engineering*, vol. 2019, no. 1, p. 8358306, 2019.

[8] N. S. Baqer, H. A. Mohammed, A. Albahri, A. Zaidan, Z. Al-Qaysi, and O. Albahri, "Development of the internet of things sensory technology for ensuring proper indoor air quality in hospital facilities: Taxonomy analysis, challenges, motivations, open issues and recommended solution," *Measurement*, vol. 192, p. 110920, 2022.

[9] T. M. Ghazal, "Retracted article: Internet of things with artificial intelligence for health care security," *Arabian Journal for Science and Engineering*, vol. 48, no. 4, pp. 5689–5689, 2023.

[10] N. Jamali, M. R. Gharib, and B. O. Koma, "Neuro-fuzzy decision support system for optimization of the indoor air quality in operation rooms.," *International Journal of Robotics & Control Systems*, vol. 3, no. 1, 2023.

[11] Y. Colella, A. S. Valente, L. Rossano, T. A. Trunfio, A. Fiorillo, and G. Improta, "A fuzzy inference system for the assessment of indoor air quality in an operating room to prevent surgical site infection," *International Journal of Environmental Research and Public Health*, vol. 19, no. 6, p. 3533, 2022.

[12] J. Chu, C.-H. Hsieh, Y.-N. Shih, C.-C. Wu, A. Singaravelan, L.-P. Hung, and J.-L. Hsu, "Operating room usage time estimation with machine learning models," in *Healthcare*, vol. 10, p. 1518, Multidisciplinary Digital Publishing Institute, 2022.

[13] D. Arad, A. Rosenfeld, and R. Magnezi, "Factors contributing to preventing operating room "never events": a machine learning analysis," *Patient Safety in Surgery*, vol. 17, no. 1, p. 6, 2023.

[14] G. Ngo, R. Beard, and R. Chandra, "Evolutionary bagging for ensemble learning," *Neurocomputing*, vol. 510, pp. 1–14, 2022.

[15] M. Habibur, M. H. Hasan, M. S. Sadik, M. Estehad, M. Rahatul, P. Kumar, A. Rahman, *et al.*, "Impact of internet of things (iot) on healthcare in transforming patient care and overcoming operational challenges," *Integrative Biomedical Research (Former Journal of Angiotherapy)*, vol. 8, no. 11, pp. 1–8, 2024.

[16] M. S. Murshed, C. Murphy, D. Hou, N. Khan, G. Ananthanarayanan, and F. Hussain, "Machine learning at the network edge: A survey," *ACM Computing Surveys (CSUR)*, vol. 54, no. 8, pp. 1–37, 2021.

[17] S. Salerno, "Tiny machine learning," in *Tiny Machine Learning Quickstart: Machine Learning for Arduino Microcontrollers*, pp. 1–40, Springer, 2025.

[18] L. Dutta and S. Bharali, "Tinyml meets iot: A comprehensive survey," *Internet of Things*, vol. 16, p. 100461, 2021.

[19] R. David, J. Duke, A. Jain, V. Janapa Reddi, N. Jeffries, J. Li, N. Kreeger, I. Nappier, M. Natraj, T. Wang, *et al.*, "Tensorflow lite micro: Embedded machine learning for tinyml systems," *Proceedings of Machine Learning and Systems*, vol. 3, pp. 800–811, 2021.

[20] H. Ap and K. K, "Secure-mqtt: an efficient fuzzy logic-based approach to detect dos attack in mqtt protocol for internet of things," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, p. 90, 2019.

[21] R. Jain, V. B. Semwal, and P. Kaushik, "Deep ensemble learning approach for lower extremity activities recognition using wearable sensors," *Expert Systems*, vol. 39, no. 6, p. e12743, 2022.

[22] Z.-R. He, T.-I. Lin, P.-J. Ko, S.-L. Tey, M.-L. Yeh, H.-Y. Wu, C.-Y. Wu, Y.-C. S. Yang, S.-N. Yang, and Y.-N. Yang, "The beneficial effect of air cleanliness with iso 14644-1 class 7 for surgical intervention in a neonatal intensive care unit: a 10-year experience," *Medicine*, vol. 97, no. 36, p. e12257, 2018.