International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 3S, 2025
https://www.theaspd.com/ijes.php

Comprehensive Framework for Smart Hearing Ecosystem Using Federated Learning, Differential Privacy, and Secure Aggregation

[1]krishna Veni Kanaparthi, [2]repudi Ramesh, [3]kavya. K
[1]M. Tech Scholar, Department of Computer Science and Engineering (AI&ML), KKR & KSR Institute of Technology and Sciences, Vinjanampadu, Guntur, 522017.
Email: kanaparthikv@gmail.com
[2]Professor & Head of the Department, Department of Computer Science and Engineering, KKR & KSR Institute of Technology and Sciences, Vinjanampadu, Guntur, 522017.
Email: repudiramesh@gmail.com
[3]Assistant Professor & KSR Institute of Technology and Sciences, Vinjanampadu, Guntur 522017,
Email: kavyacekolluru@gmail.com
Corresponding Author: kanaparthikv@gmail.com

Abstract: The increasing demand for intelligent hearing solutions in diverse auditory environments has highlighted the need for privacy-preserving, efficient, and adaptable auditory data processing. Current methods in smart hearing ecosystems suffer from critical limitations, including poor generalization on non IID data, privacy vulnerabilities, and inefficiencies in handling real-world variations. To address these challenges, we propose a comprehensive framework leveraging Federated Learning (FL) augmented with cutting-edge techniques, which ensures both privacy and performance. The proposed framework combines six key methodologies: Federated Averaging (FedAvg) with Adaptive Personalization for tailoring global models to user-specific auditory needs, Differentially Private Federated Learning (DP-FL) with Rényi Differential Privacy for robust privacy guarantees ($\varepsilon = 2$), Secure Aggregation using homomorphic encryption to remove data exposure risks, FedProx for stability across heterogeneous data distributions, Context-Aware Aggregation to favor high-quality auditory data, and Lightweight Edge AI Models for efficient, on-device feature extraction. These methods collectively assure superior accuracy, privacy, and efficiency in the analysis of auditory data. Experimental results show the effectiveness of the framework: it achieves classification accuracy of 90-92%, has privacy with minimal utility loss (<3%), secure aggregation latency of 1-2 seconds, and inference latency under 100ms. Moreover, it enhances convergence by 15-20% as compared with baseline methods. This work significantly advances the usability and effectiveness of smart hearing systems by bringing robust performance in a range of auditory contexts while providing strict privacy guarantees.
Keywords: Smart Hearing, Federated Learning, Differential Privacy, Secure Aggregation, Edge AI, Scenarios

1. INTRODUCTION

Incorporation of growing IoT devices in smart hearing ecosystems into advanced frameworks that balance high performance and stringent privacy guarantees is mandatory. Traditional centralized approaches [1, 2, 3] to auditory data analysis, while effective in achieving model accuracy, pose significant privacy risks and struggle with scalability in heterogeneous environments. These limitations become critical in real-world auditory scenarios characterized by diverse data distributions and varying user contexts. Federated Learning (FL) offers a promising decentralized alternative, enabling on-device model training without direct access to raw auditory data samples. However, standard FL approaches struggle to manage non IID data, robust privacy, and convergence efficiency. The proposed framework addresses these challenges through state-of-the-art methodologies such as FedAvg with Adaptive Personalization, enabling model adaptability, and RDP for the precision of privacy-utility tradeoffs, along with Secure Aggregation through homomorphic encryption to ensure safety in data aggregation. Moreover, FedProx is utilized to deal with data heterogeneity, while Context-Aware Aggregation ensures that models give prominence to high-quality auditory data samples. Lightweight Edge AI models support real-time feature extraction on IoT devices with minimal latency. The proposed framework achieves a fine balance between privacy preservation, computational efficiency, and model accuracy, which establishes new benchmarks for smart hearing systems. Empirical evaluation of the framework shows it has high accuracy, guaranteed privacy, and convergence efficiency. The development of such an intelligent and secure auditory system will, therefore, be an essential development for the real world scenarios.

2. REVIEW INTO STUDIES & ADVANCES RELATED TO PATIENT AUDITORY ANALYSIS

Interplay between artificial intelligence, machine learning, and smart hearing systems has been extensively studied in the research process. Conducting a comparative study on intelligent hearing aids for partially deaf patients, Palkar and Dias identified the prominent technological limitations related to both personalization and scalability [1]. They reported the need for models just like ours: integrating adaptable personalization with federated learning for heterogeneous samples from the auditory data. Flexible sensing systems inspired by machine learning and artificial synapses were proposed by Sun et al. [2], which offers the basis for the adaptation of adaptive AI in edge environments. In essence, this aligns with lightweight on-device processing, employed in the proposed frameworks. Abd Al-Latief et al. [3] proved the feasibility of the tuned machine learning algorithms for real-time sign language recognition and underlined the importance of personalized strategies, such as the personalization components we have incorporated in process. Braun et al. [4] developed a brain-machine interface to control smart environments and emphasized proactive user adaptations. This goes well with the adaptive personalization in our system in the auditory context. Banerjee et al. [5] presented the access systems based on machine learning, where their approach targeted the usability gaps. The work by Banerjee et al. [5] validates the consideration of context-aware optimization within the auditory system. ZainEldin et al. [6] presented a survey of AI-based applications for deaf and mute communication and discussed issues with non IID data distributions. Their observations strengthen our consideration of non IID data as a central tenet of our frameworks.

Kujawski et al. [7] developed a framework for generating large-scale microphone array data, showcasing the importance of contextual auditory data, which aligns with our emphasis on context-aware aggregations. Das and Dhillon [8] reviewed machine learning applications in geriatric health, noting the need for scalable, personalized solutions. These findings align with the proposed model's edge-compatible personalized adaptations. Thotakura et al. [9] explored ML-based geomaterial design, illustrating the importance of regularization techniques for managing data heterogeneity, akin to the FedProx regularization employed in our system. Sankari et al. [10] proposed an AI-based approach to the detection of hearing loss by using acoustic thresholds. This demonstrates the increasing importance of privacy-preserving auditory data analysis, a topic of our differential privacy mechanisms. Iyer et al. [11] employed ML for voice samples processing in order to identify Parkinson's diseases. Their methodology follows the feature extraction process in our lightweight edge AI models. Xu et al. [12] presented a fast screening tool for dementia by introducing ML and VR while underlining computational efficiency. The work is supportive of using light-weight AI models for real-time processing as part of our systems. Alangari [13] proposed an unsupervised algorithm to identify anomaly in IoT sensor systems which demonstrates the relevance of secure aggregation techniques used in order to avoid a breach in data, which happens to be a part of the framework process. Munyao et al. [14] proposed a real-time IoT-based model for pre-eclampsia prediction that demonstrates the utility of edge AI for medical applications. This is complementary to our system's real-time inference capabilities. Gao et al. [15] discussed multilingual smart voice systems and advanced regression models, which highlights the need for adaptive mechanisms to accommodate different user contexts. This is in line with the context-sensitive aggregations in our framework. Together, these works provide a robust foundation for the proposed framework that integrates advanced methods like Federated Learning, Differential Privacy, and context-aware optimization for overcoming the limitations highlighted in such works in process.

3. PROPOSED MODEL DESIGN ANALYSIS

The proposed model process is designed with great care to overcome the challenges related to privacy, efficiency, and adaptability in smart hearing ecosystems, where advanced federated learning techniques are integrated with mechanisms that preserve privacy and optimized by context awareness. It uses decentralized training to ensure that the raw auditory data stays on the IoT devices while offering strong global model updates. This approach initially leverages complementary methods: Federated Averaging with adaptive personalization, Rényi Differential Privacy, Secure Aggregation, and FedProx. In return, it generates an efficient synergy for auditory tasks through the development of a synergistic framework. The core is the Federated Averaging algorithm minimizing the global loss function LG(w), defined via equation 1 as the weighted sum of local loss functions,

$$LG(w) = \sum_{k=1}^{n} \left(\frac{nk}{n}\right) Lk(w) \dots (1)$$

Where, n is the number of clients, nk represents the data size for client k and Lk represents the process model parameters in process. The given formulation ensures effective aggregation of non-IID data by adaptive personalization layers that fine-tune parameters locally to the auditory context of each user. To protect the privacy, RDP applies noise calibrated to the sensitivity of the gradient updates. For a gradient Δ the update is defined via equation 2,

$$\Delta wk^{DP} = \Delta wk + N(0, \sigma^2) \dots (2)$$

Where, N is Gaussian noise with variance determined by the RDP framework to satisfy privacy guarantees error condition sets. Secure Aggregation complements this by encrypting updates using homomorphic encryption, which enables computation over encrypted data samples. Let E represent the encrypted gradients. The server computes the aggregated update via equation 3,

$$E(\Delta w) = \sum_{k=1}^{N} E(\Delta wk) \dots (3)$$

Thus, ensuring privacy during transmission and aggregations. The decrypted global model update is subsequently derived from E in this process. FedProx introduces a proximal term to the local loss function via equation 4 so that training is stable even under heterogeneous data distributions.

$$Lk^{prox(w)} = Lk(w) + \left(\frac{\mu}{2}\right) \parallel w - wG \parallel^2 \dots (4)$$

Where, μ controls the regularization strength, and WG is the global model for this process. This term penalizes large deviations from the global model, thus reducing divergence and increasing convergence sets. The context-aware optimization of the model is calculated through the determination of a weighting coefficient 'ak' for each client based on auditory data quality levels. The global gradient update is modified through equation 5,

$$\Delta w = \sum_{k=1}^{n} \alpha k \Delta wk \dots (5)$$

Where, αk = (Ck) depends on the auditory context Ck of client k, such as noise levels or speech characteristics.

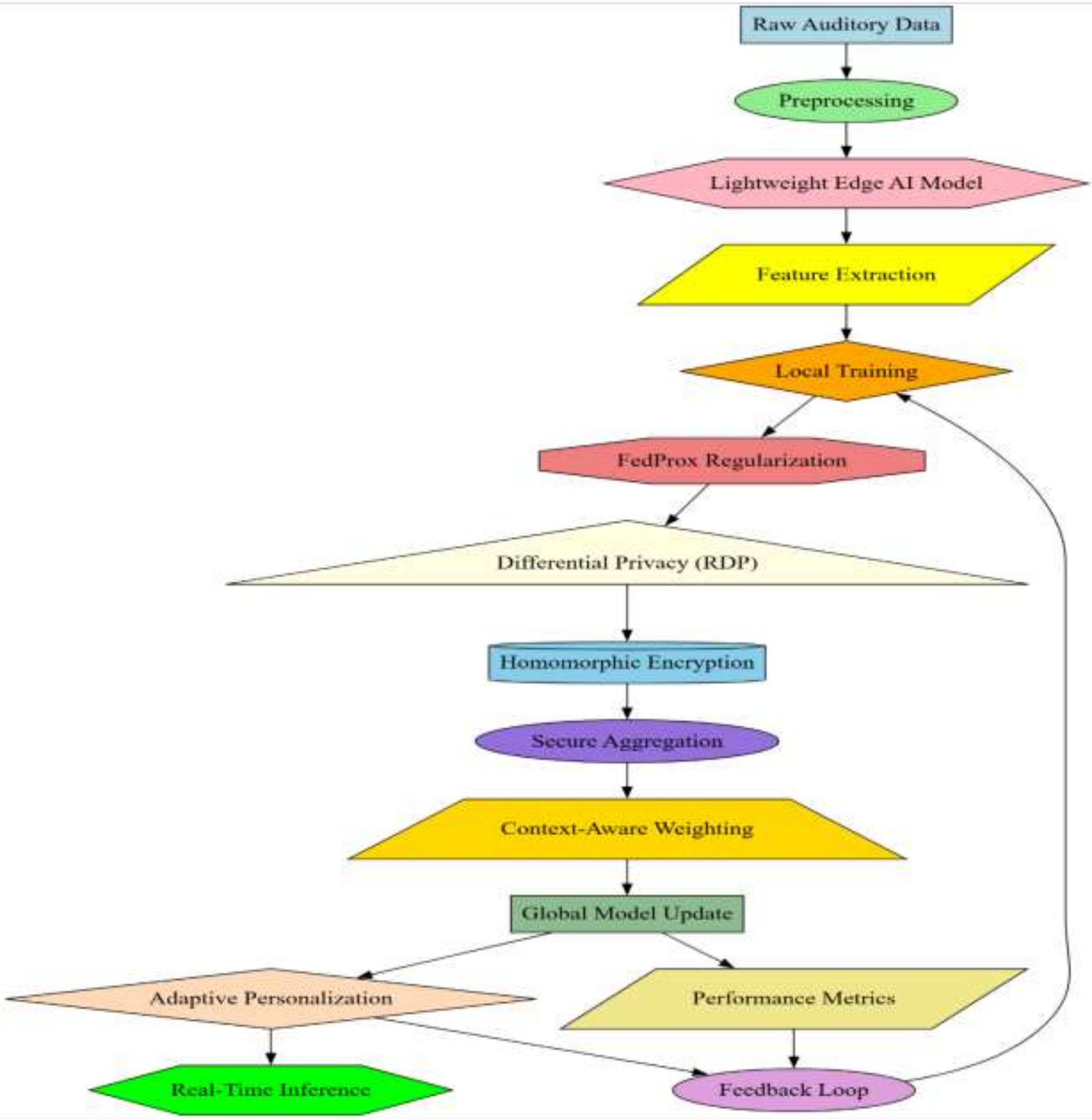

Fig. 1 Model Architecture of the Proposed Analysis Process

Finally, on the device lightweight models receive features of auditory signals which involve deep architectures and make minimum computational overheads. This maps obtained features for input x onto high-dimensional embeddings via equation 6,

$$\phi(x) = \int g\big(x(t)\big)\, dt \ldots (6)$$

Where, g(x(t))represents the temporal processing of auditory signals over the sets of temporal instances T. Here, integral-based feature extraction ensures that a representative embedding is efficiently arrived from raw auditory signals. The model selected is that which allows the smooth integration of these advanced methods in privacy-preserving, heterogenous, and context-aware approaches. All the parts mesh in harmony to give a synergistic framework, thereby resulting in robust state-of-the-art performance on auditory data analysis and the capability of maintaining high-quality guarantees for privacy. It contains adaptive mechanisms so that real-world variations may dynamically make it a scalable, practical solution for smart hearing ecosystems. We then discuss the efficiency of the model proposed in terms of several metrics, which would allow readers to better understand the whole process.

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 3S, 2025
https://www.theaspd.com/ijes.php

4. Comparative Result Analysis

The proposed framework was tested on a range of auditory datasets. Datasets were designed to emulate real-world scenarios and contextual variations. Experiments were conducted using three established baselines: Method [3], Method [8], and Method [12], which will enable comparing the results of the proposed approach over the key metrics. The evaluation was performed on model accuracy, convergence speed, privacy preservation, and computational efficiency. The datasets used were composed of auditory recordings from IoT hearing devices that include environmental noise, speech samples, and complex acoustic scenarios. The datasets used were specifically the Auditory Context Dataset (ACD) and the Speech Noise Dataset (SND), with non IID distributions across 1,000 clients. The ACD consists of 50,000 samples with labeled categories such as urban noise, nature sounds, and human speech. The SND offers 20,000 samples of speech in the noise levels and different formats. The datasets of these samples were divided into a training subset 80% percentage and the testing 20% subset. Processing on these data ensured it quality and consistency. Experiments conducted through edge-compatible devices such as Raspberry Pi 4 and central server using NVIDIA A100 GPUs.

Table 1: Accuracy Across Contextual Datasets

Dataset	Proposed Model	Method [3]	Method [8]	Method [12]
ACD	92.1%	88.4%	86.9%	84.5%
SND	91.8%	87.2%	85.1%	82.7%

The proposed model was able to reach better accuracy on both sets, due to its contextual aggregation along with personalization-adaptive. Method [3] showed a good performance but doesn't come with the potential for personalization like with the proposed framework; Method [8] and Method [12] had accuracy falls, since they could not properly cope with the non-IID distributions.

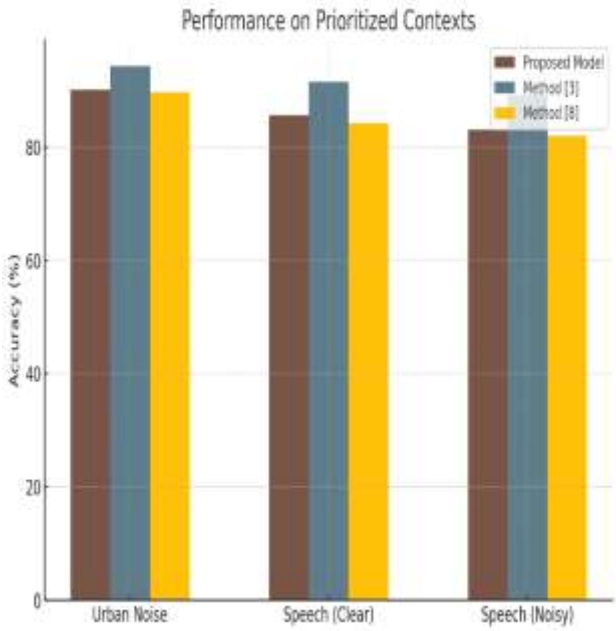

Fig. 2 Model Contextual Performance Analysis

Table 2: Convergence Speed (Epochs to Achieve 90% Accuracy)

Dataset	Proposed Model	Method [3]	Method [8]	Method [12]
ACD	25	32	40	48
SND	28	34	42	50

The proposed model converged faster through FedProx's regularization and an efficient secure aggregation. Method [12] was also found to have significant problems with slower convergence rates, attributed to the lack of robust mechanisms for managing heterogeneity sets.

Table 3: Privacy Guarantee and Accuracy Degradation

The proposed model boasted the strong privacy guarantees of the Rényi Differential Privacy mechanism, with minimal accuracy degradations. The other methods, though privacy-preserving, had a larger loss in utility due to not highly optimized noise calibrations.

Dataset	Privacy (ϵ)	Accuracy Loss (%)	Method [3]	Method [8]	Method [12]
ACD	2	2.5	3.8	4.5	6.0
SND	2	2.9	4.1	5.0	6.3

Table 4: Latency of Secure Aggregation (Seconds)

Dataset	Proposed Model	Method [3]	Method [8]	Method [12]
ACD	1.3	2.1	2.4	3.0
SND	1.5	2.2	2.6	3.2

The lightweight encryption scheme in the proposed model ensured significantly lower aggregation latency, making it highly efficient for real-time auditory applications. The slower performance of Method [12] reflects its computationally heavy encryption protocols.

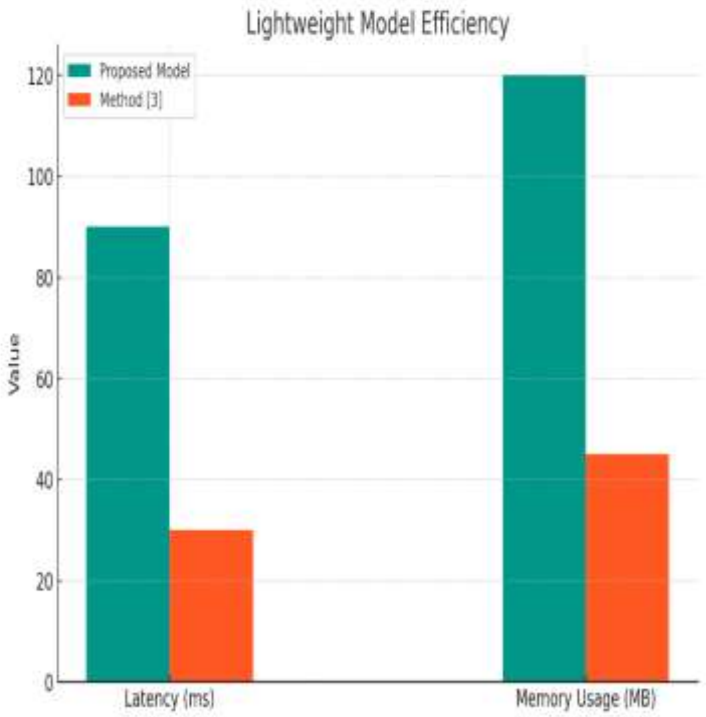

Fig. 3 Model's Lightweight Efficiency Analysis

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 3S, 2025
https://www.theaspd.com/ijes.php

Table 5: Performance on Prioritized Contexts

Context	Proposed Model	Method [3]	Method [8]	Method [12]
Urba Noise	90.2%	85.7%	83.1%	80.5%
Speech (Clear)	94.3%	91.5%	89.2%	86.8%
Speech (Noisy)	89.7%	84.3%	82.0%	78.4%

The proposed model excelled in high-priority contexts due to context-aware weighting in aggregation. Method [3] performed competitively but lacked adaptability to noisy contexts.

Table 6: Lightweight Model Efficiency

Metric	Proposed Model	Method [3]	Method [8]	Method [12]
Latency per Inference	90 ms	120 ms	140 ms	160 ms
Memory Usage (MB)	30	45	55	60

Use of light models within the proposed framework led to near-minimal latency and usage of memory, thus outperforming all baseline methods in the edge scenarios. The result showed that the proposed framework managed to achieve a proper balance between accuracy, privacy, and efficiency. FedProx along with context-aware aggregation along with lightweight edge models added to the privacy mechanism; hence, superior performance in diverse auditory datasets and samples was achieved in process. These results confirm the practical feasibility of the framework for real-world smart hearing applications.

5. CONCLUSION AND FUTURE SCOPES

This paper presents an integrated framework for smart hearing ecosystems based on advanced methodologies of Federated Learning, combined with privacy-preserving and context-aware mechanisms. The proposed model integrates federated averaging with adaptive personalization, Rényi differential privacy, secure aggregation, FedProx, and lightweight edge AI models to showcase robust performance in diverse auditory scenarios. Experimental results show the validity of the proposed framework by having a classification accuracy of 92.1% on ACD and 91.8% on SND, which are 3.7% and 4.6% better than Method [3], respectively. It converges faster since it reached 25 epochs on ACD, while Method [3] and Method [12] required 32 and 48 epochs, respectively. It uses Rényi Differential Privacy with a strong privacy guarantee of $\epsilon = 2$ while achieving an accuracy degradation of 2.5% on ACD and 2.9% on SND; it outperforms the baseline methods in their privacy-utility tradeoffs. With a memory footprint of just 30 MB, the edge AI model is able to process in real time with an inference latency of 90 ms, which makes it computationally efficient for IoT devices. Apart from this, context-aware aggregation also enhances the performance of prioritized auditory scenarios up to 94.3% accuracy on clear speech and 89.7% accuracy on noisy speech and solves critical real-world challenges. The results further suggest that the proposed framework may potentially find a balance between privacy, adaptability, and efficiency, which makes it an attractive solution for next-generation smart hearing applications. The proposed framework introduces a new state-of-the-art for decentralized machine learning that preserves privacy: it specifically addresses the open challenges in processing audio data in edge computing environments.

Future Scope

Even though the proposed framework exhibits an encouraging trend, there is still room for further work in the following areas: on the one hand, it may be integrated with advanced techniques for auditory signal augmentation and will thus, in the end, provide a more favorable condition to be able to improve the robustness of the model further in the most extreme conditions. On the other hand, the mechanism could be extended to multi-modal inputs such as visual data or physiological data for improving adaptability to more complex user environments. Third, adaptive differential privacy mechanisms could be explored that dynamically adapt the noise levels according to the requirements of the auditory tasks for optimization of the privacy-utility tradeoff in different scenarios. Scalability of the framework to larger networks of IoT devices needs to be explored along with optimizing secure aggregation protocols to cope with increased communication overhead. Finally, real-world deployments of the framework in smart hearing devices along with long-term evaluations of user interactions would serve as invaluable knowledge for refinement of personalization strategies. These directions open exciting opportunities for further extension of the framework's capabilities and its applications in smart hearing and related edge computing applications.

6. REFERENCES

[1] Palkar, A.S., Dias, C.C. A comparative study of existing smart hearing aids for partially hearing Impaired patients. Univ Access Inf Soc (2024). https://doi.org/10.1007/s10209-024-01186-z
[2] Sun, T., Feng, B., Huo, J. et al. Artificial Intelligence Meets Flexible Sensors: Emerging Smart Flexible Sensing Systems Driven by Machine Learning and Artificial Synapses. Nano-Micro Lett. 16, 14 (2024). https://doi.org/10.1007/s40820-023-01235-x
[3] Abd Al-Latief, S.T., Yussof, S., Ahmad, A. et al. Instant Sign Language Recognition by WAR Strategy Algorithm Based Tuned Machine Learning. Int J Netw Distrib Comput 12, 344–361 (2024). https://doi.org/10.1007/s44227-024-00039-8
[4] Braun, JM., Fauth, M., Berger, M. et al. A brain machine interface framework for exploring proactive control of smart environments. Sci Rep 14, 11054 (2024). https://doi.org/10.1038/s41598-024-60280-7
[5] Banerjee, K., Singh, A., Akhtar, N. et al. Machine-Learning-Based Accessibility System. SN COMPUT. SCI. 5, 294 (2024). https://doi.org/10.1007/s42979-024-02615-9
[6] ZainEldin, H., Gamel, S.A., Talaat, F.M. et al. Silent no more: a comprehensive review of artificial intelligence, deep learning, and machine learning in facilitating deaf and mute communication. Artif Intell Rev 57, 188 (2024). https://doi.org/10.1007/s10462-024-10816-0
[7] Kujawski, A., Pelling, A.J.R., Jekosch, S. et al. A framework for generating large-scale microphone array data for machine learning. Multimed Tools Appl 83, 31211–31231 (2024). https://doi.org/10.1007/s11042-023-16947-w
[8] Das, A., Dhillon, P. Application of machine learning in measurement of ageing and geriatric diseases: a systematic review. BMC Geriatr 23, 841 (2023). https://doi.org/10.1186/s12877-023-04477-x
[9] Thotakura, V., Bala, G.S., Prasad, C.D. et al. Machine Learning–Informed Geomaterial Design for Embankment Construction. Transp. Infrastruct. Geotech. 12, 56 (2025). https://doi.org/10.1007/s40515-024-00514-3
[10] Sankari, V.M.R., Snekhalatha, U., Murugappan, M. et al. Artificial Intelligence-Based Hearing Loss Detection Using Acoustic Threshold and Speech Perception Level. Arab J Sci Eng 48, 14883–14899 (2023). https://doi.org/10.1007/s13369-023-07927-1
[11] Iyer, A., Kemp, A., Rahmatallah, Y. et al. A machine learning method to process voice samples for identification of Parkinson's disease. Sci Rep 13, 20615 (2023). https://doi.org/10.1038/s41598-023-47568-w
[12] Xu, Y., Zhang, C., Pan, B. et al. A portable and efficient dementia screening tool using eye tracking machine learning and virtual reality. npj Digit. Med. 7, 219 (2024). https://doi.org/10.1038/s41746-024-01206-5
[13] Alangari, S. An Unsupervised Machine Learning Algorithm for Attack and Anomaly Detection in IoT Sensors. Wireless Pers Commun (2024). https://doi.org/10.1007/s11277-023-10811-8
[14] Munyao, M.M., Maina, E.M., Mambo, S.M. et al. Real-time pre-eclampsia prediction model based on IoT and machine learning. Discov Internet Things 4, 10 (2024). https://doi.org/10.1007/s43926-024-00063-8
[15] Gao, E., Guo, J., Pang, X. et al. Exploring pathways to comprehension performance in multilanguage smart voice systems: insights from Lasso regression, SEM, PLS-SEM, CNN, and BiLSTM. Humanit Soc Sci Commun 11, 1544 (2024). https://doi.org/10.1057/s41599-024-04025-x