WJEMSD

WASD
WORLD ASSOCIATION FOR SUSTAINABLE DEVELOPMENT

World Journal of
ENTREPRENEURSHIP, MANAGEMENT
AND SUSTAINABLE DEVELOPMENT

ISSN: 2042-5961 (Print) | 2042-597X (Online)

WJEMSD V22 N5 2026

DOI: 10.47556/J.WJEMSD.22.5.2026.1

RESEARCH

EchoBand and Pulses: Developing AI Solutions for Sustainable Sound Accessibility in Low-Income Settings

Cyrille Awatchede Ange Noutchogbe
Kobe Institute of Computing (KIC), Kobe, Japan
Email: angenoutchogbe@gmail.com
ORCID: 0009-0002-4683-031X

ABSTRACT

PURPOSE: To demonstrate that affordable, offline sound-awareness technology for deaf and hard-of-hearing (DHH) users is achievable in low-income settings using edge AI.

DESIGN/METHODOLOGY/APPROACH: Design Science Research and Tankyu Practice methodologies were applied in Benin. Four context-specific TinyML models were deployed on a USD 30 microcontroller-based device and evaluated with 26 DHH participants over 40 days.

FINDINGS: All four models achieved 90.7- 97.9% laboratory accuracy. On-device offline inference delivered a latency of 267 ms, within the 500 ms alert threshold. The mean SUS score was 72.4.

ORIGINALITY/VALUE: This is the first evaluation of a DHH sound-awareness wearable with participants in sub-Saharan Africa, with a community contribution model that inverts the recording burden of prior personalisation systems.

RESEARCH LIMITATIONS: The research is based on a single-city sample. Field accuracy was not quantitatively measured. Extension to other LMIC contexts is needed.

PRACTICAL IMPLICATIONS: A USD 30 offline device is technically feasible as a research prototype. Systematic data collection and false positive mitigation are required before production deployment.

CITATION: Noutchogbe, C. A. A. (2026): EchoBand and Pulses: Developing AI Solutions for Sustainable Sound Accessibility in Low-Income Settings. *World Journal of Entrepreneurship, Management and Sustainable Development (WJEMSD)*, Vol. 22, No. 5, pp. 403-418.

RECEIVED: 19 June 2026 / **REVISED:** 21 June 2026 / **ACCEPTED:** 3 July 2026 / **PUBLISHED:** 10 July 2026

KEYWORDS: *Assistive Technology; Deaf and Hard-of-Hearing; Tinyml; Sustainable Development; Edge AI; Low-Income Contexts; Community-Based Innovation; Wearable Devices; West Africa; Benin.*

INTRODUCTION

Sound is infrastructure. Before any technical development, the research team engaged directly with deaf communities in Benin, where a consistent finding emerged: the devices that exist do not work for us. A deaf worker in the market cannot hear a motorcycle approaching from behind. A deaf parent cannot hear their infant crying. A deaf employee cannot hear a fire alarm. These are not rare situations. For the 466 million people with disabling hearing loss worldwide, they happen every day (World Health Organisation, 2021; Chadha *et al.*, 2021).

A review of the literature reveals a striking gap: every peer-reviewed study on wearable sound-awareness systems for DHH users has been conducted in North America, Europe, or East Asia (Jain *et al.*, 2020; Do *et al.*, 2023; Goodman *et al.*, 2025). Not one has evaluated such a system with participants in sub-Saharan Africa. The gap is not merely geographic. Acoustic environments, sound priorities, and economic realities in West African cities are simply absent from the literature.

This paper presents EchoBand and Pulses: a wrist-worn device and Android application developed specifically for low-income settings (See Figure 1). EchoBand classifies environmental sounds entirely on-device using TinyML (machine learning on microcontrollers), requires no internet connection, and costs approximately USD 30.

The system makes two contributions that no prior system has made: the first field evaluation of DHH sound-awareness technology with participants in sub-Saharan Africa, and a community-based data contribution model that inverts the individual-recording burden found across all prior personalisation systems. The study pursues three research objectives: (RO1) to design and evaluate an affordable, offline sound-awareness wearable for DHH users in a low-income context; (RO2) to determine whether context-specific TinyML models can achieve acceptable accuracy and latency on commodity hardware costing approximately USD 30; and (RO3) to assess usability and acceptance among DHH participants in sub-Saharan Africa.

Against this background, the study makes three contributions that, to the authors' knowledge, have not appeared together in any prior work. First, it presents the first peer-reviewed evaluation of a DHH sound-awareness wearable with participants in sub-Saharan Africa, addressing a geographic gap in the published literature on this topic. Second, it demonstrates that watch-only TinyML inference, previously reported as unusable at 5.9 seconds (Jain *et al.*, 2022), is achievable in 267

milliseconds on comparable hardware using a four-model modular architecture at no additional cost. Third, it introduces a community-based data contribution model that inverts the individual recording burden found in all prior personalisation systems, and is grounded in collective action theory (Benkler, 2006; Olson, 1965; Ostrom, 1990).

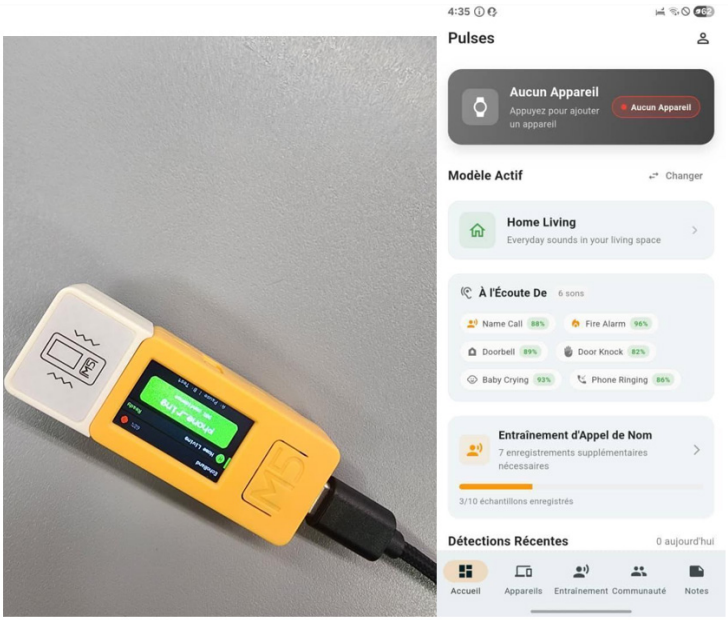

Figure 1: The EchoBand Wearable Device

Source: Author's own work

The left side of Figure 1 displays a vibration alert after classifying an environmental sound, and the Pulses Android companion application (right) shows the detected sound category and active context mode.

LITERATURE REVIEW AND THEORETICAL FRAMEWORK

The Lab-To-Field Gap in DHH Sound-Awareness Research

SoundWatch (Jain *et al.*, 2020) established the baseline for smartwatch-based sound classification with haptic and visual alerts. ProtoSound (Jain *et al.*, 2022) added few-shot personalisation, allowing users to train the system on their own acoustic environment. AdaptiveSound (Do *et al.*, 2023) went further, allowing users to provide corrective feedback during real-world use. SPECTRA (Goodman *et al.*, 2025) is the most recent work: an end-to-end interactive machine-learning pipeline in which DHH users build and refine their own classifiers.

Huang *et al.* (2023) deployed SoundWatch with 10 DHH participants for three weeks and found persistent failures: background noise, overlapping sounds, and acoustic confusion. Their title says it plainly: 'Not There Yet.' Training on cleaned, studio-recorded datasets does not prepare a model for a crowded market or an open office. A laboratory accuracy figure does not tell a DHH user whether the device will alert them in time. This study directly addresses this gap in the current study.

Every study in this line was conducted in the United States or Canada. Sub-Saharan Africa accounts for a disproportionate share of the global hearing loss burden (Mulwafu *et al.*, 2016), yet no wearable sound-awareness system has been evaluated in the region. No prior work has tested whether sound priorities, feedback preferences, or usability patterns differ in low-income urban settings. This study addresses this gap.

Compared to prior DHH sound-awareness systems, EchoBand presents a distinct profile. SoundWatch (Jain *et al.*, 2020), ProtoSound (Jain *et al.*, 2022), and SPECTRA (Goodman *et al.*, 2025) all depend on paired smartphones or cloud connectivity, were evaluated exclusively in high-income countries, and operate on devices costing upwards of USD 200. EchoBand runs entirely on a USD 30 wrist-worn device, requires no paired phone for core sound alerting, and was evaluated in sub-Saharan Africa. These differences are not incremental; they reflect a fundamentally different design target.

TinyML for Resource-Constrained Environments

TinyML refers to running machine learning on microcontrollers with 256 KB to 1 MB of random-access memory (RAM), no graphics processing unit (GPU), and no network dependency (Warden and Situnayake, 2019; Rusci and Tuytelaars, 2023; Heydari and Mahmoud, 2025). Recent work has further demonstrated on-device personalisation for keyword spotting systems with minimal user effort (Rusci *et al.*, 2024). For assistive technology in low-income contexts, these constraints represent the target environment, not obstacles to engineer around.

There is a specific reason why prior DHH sound-awareness systems could not run on-device. Jain *et al.* (2022) tested watch-only inference and found an average latency of 5.9 seconds, which they termed "currently unusable", and recommended a paired smartphone as the minimum viable architecture. This study took a different architectural path. Rather than deploying one large universal classifier, four small, context-specific models were trained, each fitting within the ESP32's memory. Each model handles four locally relevant sounds. The watch-only latency achieved was 267 milliseconds, compared with the paired phone architecture average of 2.2 seconds.

Assistive Technology access in Low-Income Settings

The Global Report on Assistive Technology (United Nations, UNICEF and World Health Organisation, 2022) estimated that 2.5 billion people need at least one assistive device, yet fewer than 1 in 10 have access. The fieldwork in Benin confirmed this reality directly.

Prior research has consistently found that DHH users prefer sound-awareness systems that are discreet, context-aware, and easy to personalise (Findlater *et al.*, 2019). EchoBand was designed from the ground up to meet those preferences within a low-income context, not by adapting a North American prototype. That choice shaped every technical decision.

Community-Based Data Collection

Every personalisation system in the DHH literature asks the DHH user to record their own training data. ProtoSound, AdaptiveSound, and SPECTRA all work this way. The assumption is that users can identify the sounds they need, record clean examples, and label them accurately. That assumption is already strained in low-bandwidth environments. It breaks entirely for sounds a DHH user has never encountered before, like a specific factory alarm or a market signal they have never consciously registered.

EchoBand inverts this model. Community members, both DHH and hearing, can request that sounds they want detected be added and contribute recordings through the mobile app. The model is trained centrally and delivered over Bluetooth Low Energy (BLE). This approach draws on well-established ideas about collective action (Olson, 1965; Ostrom, 2009) and knowledge commons (Benkler, 2006). No prior work has applied these frameworks to DHH assistive technology.

Theoretical Framework: Tankyu Practice

The Tankyu Practice methodology (Sumitani, Kobe Institute of Computing) served as the research framework. Tankyu means contemplation. The method starts with genuine listening before any technical work begins. Community interviews with DHH individuals and deaf associations in Benin revealed that locally relevant sounds, including motorcycle horns, honking vehicles, and workplace alarms, were not well represented in the datasets used by prior systems. The sound categories were built directly from those conversations.

The system works because design decisions were made by the community, not by parameter tuning. Hardware choices, feedback patterns, and sound categories all emerged from fieldwork. TinyML was the tool used to implement what the community identified as needed.

RESEARCH METHODOLOGY

The Design Science Research (DSR) methodology (Hevner *et al.*, 2004) was employed, focusing on creating and evaluating artefacts to address real-world problems. Tankyu Practice provided an ethnographic grounding before any technical development began. In the first phase, structured interviews were conducted with DHH community members, CAEIS staff, and representatives of ASUNOES Benin to confirm the need for the solution, identify locally relevant sound categories, and understand participants lived acoustic environments. These conversations directly shaped the four context models and the decision to prioritise wrist-based haptic feedback over visual-only alerts.

Study Setting and Participants

Fieldwork was conducted at CAEIS (Centre d'Accueil d'Education et d'Integration des Sourds) in Porto-Novo, Benin, from December 2025 to January 2026, in partnership with ASUNOES Benin, the national association of deaf students. CAEIS has served deaf individuals since 1997. Working with CAEIS and ASUNOES staff, 26 DHH participants were recruited through purposive sampling (See Table 1). Ages ranged from 12 to 40, with hearing loss ranging from mild to profound. All participants gave informed consent. All data collection was conducted through a sign language interpreter, as participants communicate primarily through sign, and the researcher does not have proficiency in the local sign language.

Table 1: Participant's Demographic Profile

Characteristic	*Category*	*N*	*Percentage*
Sex	Male	14	53.8%
	Female	12	46.2%
Age group	12-17 years	10	38.5%
	18-30 years	12	46.2%
	31-40 years	4	15.4%
Degree of hearing loss	Mild to moderate	8	30.8%
	Severe to profound	18	69.2%
Total		26	100%

Source: Analysed by author

(N = 26), CAEIS Porto-Novo, Benin (December 2025–January 2026)

Data Collection

Three instruments were used to collect data. First, a five-question pre-deployment questionnaire on daily sound awareness challenges and prior technology use; second, the system interaction logs recorded model outputs and battery consumption; third, a post-deployment semi-structured interviews on usability, unexpected benefits, adoption barriers, and remaining challenges. Interviews were recorded with permission, transcribed, and analysed through thematic coding.

Laboratory validation accuracy was measured on held-out test sets using accuracy, precision, recall, and F1-score. Acoustic classification accuracy was assessed by the researcher against labelled ground-truth recordings; because the researcher has normal hearing, correct and incorrect model outputs could be identified by direct listening. DHH participants evaluated a different set of questions: vibration-pattern intelligibility, overall system usability via the SUS questionnaire, and the relevance of the targeted sound categories to their daily lives. Field performance was additionally assessed through qualitative observation and participant-reported experiences over the 40-day deployment. This distinction is stated explicitly because conflating laboratory and field accuracy is precisely the problem Huang *et al.* (2023) identified in prior work.

Ethical Statement: The study adhered to institutional ethical guidelines for research involving vulnerable populations. All participants gave verbal informed consent, provided through the sign language interpreter. All data were anonymised for reporting.

Hardware Platform

EchoBand was built on the M5StickC Plus 2 microcontroller: an ESP32-PICO-V3-02 system-on-chip with 520 KB of static random-access memory (SRAM), 8 MB Pseudo-Static RAM (PSRAM), 8 MB flash, a 1.14-inch TFT display at 135 x 240 pixels, a built-in microphone, a vibration motor, and a 200 mAh battery. The module costs approximately USD 15. Assembled with enclosure and strap, the total cost is approximately USD 30 (See Figure 2).

All sound classification inference runs on the device itself; no internet connection is required at any point. The companion Android application, Pulses, handles context mode selection and model delivery over BLE. A 7-10 KB TensorFlow Lite (TFLite) model file (David *et al.*, 2021) transfers over BLE in under two seconds. The device retains the previous model as a fallback until the new one is confirmed operational.

Figure 2: Assembled EchoBand Device

Source: Author's own work

The assembled EchoBand device comprising the M5StickC Plus 2 microcontroller unit, vibration module, custom wristband enclosure with orange strap, and USB charging cable.

Machine Learning Models

Four context-specific models were trained using Edge Impulse: Home Living, Home Safety, Outdoor, and Workplace. Each cover four sounds identified through community fieldwork in Benin, rather than from a generic taxonomy. Training data collection began with original field recordings made in Porto-Novo, capturing sounds in markets, workplaces, and residential environments. Field recordings yielded between 30 and 60 samples per class, insufficient in isolation to train robust models. FreeSound and AudioSet were used to supplement field data, bringing the total number of samples per class to approximately 200.

Using community-sourced field recordings as the seed rather than purely generic data ensured that the acoustic signatures of locally relevant sounds were represented in the training distribution. Every audio clip, regardless of source, was treated with the same preprocessing pipeline before training: a high-pass filter to remove low-frequency rumble, peak normalisation, length trimming to a fixed window, silence padding where needed, and resampling to 16 kHz mono.

Each model uses a one-dimensional convolutional neural network (CNN): two convolutional blocks and a dense classification layer, operating on Mel-frequency cepstral coefficient (MFCC) features. INT8 post-training quantisation reduces each model to approximately 10 KB, requiring 20-40 KB of tensor arena during inference. A single model covering all 16 sound classes would require an estimated 80-150 KB tensor arena, risking memory allocation failures on the 520 KB SRAM budget. The four-model approach keeps the active tensor arena below 8% of available memory.

Context mode selection was designed to occur before leaving a location, not in real time. Automatic switching would require additional sensors and increase costs. Pre-departure selection aligns with how participants naturally plan their movements, and this design decision emerged from observing actual usage behaviour.

RESULTS

Model Performance

This section presents results in relation to the three research objectives stated in the introduction. Addressing RO1, the assembled EchoBand device was sourced at approximately USD 30 per unit (M5StickC Plus 2 at USD 15, an enclosure and a strap at approximately USD 15). All 40 deployment days were completed without any internet connection to the wristband; BLE model delivery required only a standard Android smartphone and was performed once per context switch. Both affordability and offline operation were therefore demonstrated in practice, not merely claimed. Addressing RO2, Table 2 presents laboratory validation accuracy for all four models.

Table 2: Laboratory Validation Accuracy by Context Model

Context model	*Overall accuracy*	*Lowest sound class*	*Highest sound class*
Home Living	93.4%	Phone Ring: 88.7%	Baby Cry / Doorbell: 100.0%
Home Safety	91.5%	Crash Impact: 80.0%	Door Forced: 95.1%
Outdoor	97.9%	Dog Bark: 95.1%	Bike Horn: 100.0%
Workplace	90.7%	Phone Ring: 84.8%	Alarm: 95.2%

Source: Analysed by author

(INT8 post-training quantisation, Edge Impulse held-out test sets, N ≈ 200 samples per class).

The Outdoor model performs best, with an accuracy of 97.9%, and three of the four sound achieving accuracies above 95% (See Figure 3). The Workplace model is the weakest at 90.7%, partly because the office phone ring overlaps acoustically with sounds in other categories. Crash Impact at 80.0% is the lowest individual class across all models, reflecting its acoustic similarity to other transient impact sounds.

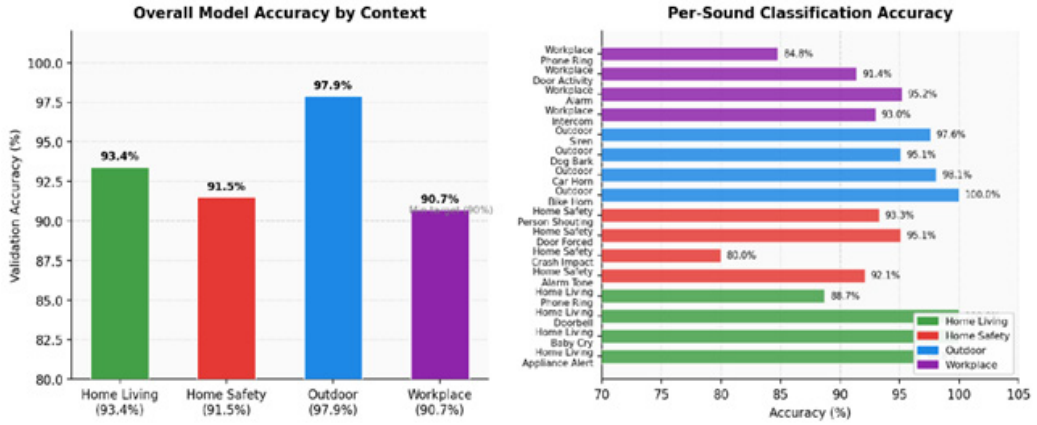

Figure 3: Laboratory Validation Accuracy Results

Source: Constructed by author

Overall accuracy per context model (INT8 post-training quantisation). Right panel shows per-sound classification accuracy across all 16 sound categories. All values were derived from held-out test sets in Edge Impulse Studio.

System Performance

Addressing RO2, end-to-end detection latency averaged 267 milliseconds, well within the 500 ms thresholds for assistive technology alerts (World Health Organisation, 2021). For comparison, cloud-based sound recognition systems typically add 2 to 5 seconds of latency due to network round-trip time. The vibration alert fires on the wrist in approximately 187 ms from audio capture: 100 ms capture, 85 ms TFLite inference, and 2 ms threshold check (Figure 4). BLE notification to

the phone adds approximately 80 ms when the app is in use. Battery life under continuous inference reaches approximately 3 hours, achieved through three optimisation steps: a voice activity gate that reduces inferences from 40 to 12 per minute, display sleep after 5 seconds of idleness, and an inference cadence of 1.5 seconds.

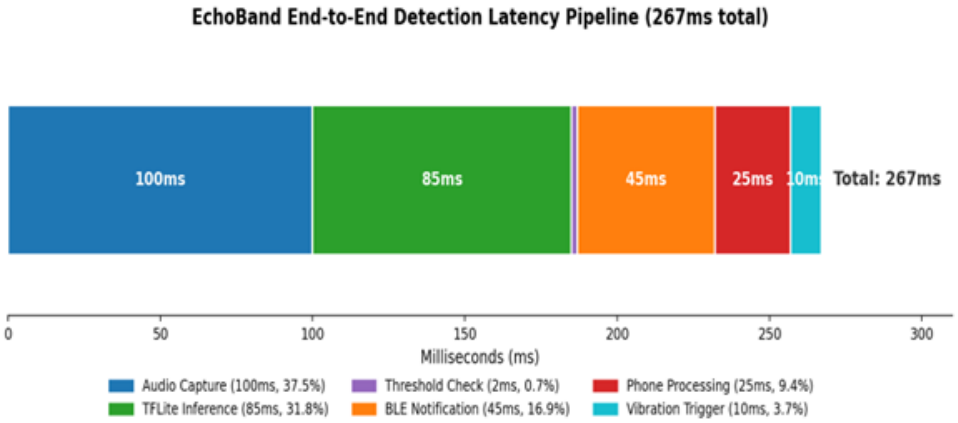

Figure 4: End-To-End Detection Latency Pipeline

Source: Constructed by author

Audio capture (100 ms) + TFLite inference (85 ms) + threshold check (2 ms) = 187 ms on-device, plus BLE notification (+80 ms) = 267 ms total end-to-end latency

User Evaluation

Addressing RO3, eighteen participants completed a full usability session. The mean System Usability Scale (SUS) score was 72.4, with individual scores ranging from 62.5 to 87.5. A score of 72.4 falls within the 'Good' band according to established benchmarks. Fifteen participants tested six distinct vibration patterns; all 15 correctly differentiated the six patterns after approximately five minutes of familiarisation. This outcome aligns with systematic evidence showing that haptic feedback is technically feasible and usable for DHH populations (Flores Ramones and Del-Rio-Guerra, 2023; Orzech *et al.*, 2024).

Qualitative feedback revealed two consistent clusters. Participants valued the wrist form factor, the haptic-plus-visual combination, the context-specific models, and offline operation. One participant said: 'Je peux sentir quand mon ami est derriere moi maintenant' ('I can feel when my friend is behind me now'), describing how the vibration improved her awareness in the market. A second participant noted: 'The system would let me know when my child cries, even when I am cooking and cannot see her.' These responses suggest the system met real, everyday needs beyond the test scenarios.

On the other hand, the USD 30 cost was still cited as a barrier by several participants. Waterproofing was requested. Aesthetics were raised as concerns; participants wanted the device to look more like a conventional watch.

DISCUSSION

The results demonstrate that context-specific TinyML models trained on local data effectively address the specific acoustic needs of DHH users in Benin, outperforming generic datasets. By employing a modular four-model architecture rather than a single universal classifier, this study resolved the latency issues identified in prior research, achieving a 267ms end-to-end response time—a 22-fold improvement over earlier designs. While these results confirm the technical feasibility of low-cost, offline sound awareness, field testing exposed a persistent "false-positive" challenge, primarily due to limited training data volume and environmental noise. Taken together, these findings position AI not merely as a technical feature but as a foundational input shaping the venture's sustainable business model (Mohamad *et al.*, 2026).

This gap between validation accuracy and real-world performance highlights the need for larger, more diverse local datasets. Finally, the community-based data contribution model offers a sustainable, scalable pathway for assistive technology deployment in resource-constrained settings. This model reflects principles of social entrepreneurship, in which a venture generates simultaneous economic, social and environmental value (Alqaoud and Alhaimer, 2024), and addresses the wider argument that AI's capacity to support socio-economic development in African contexts depends on deployment at a scale and cost appropriate to local conditions (Uwizeyimana, 2024). Longitudinal validation of the contribution mechanism remains a priority.

PRACTICAL IMPLICATIONS

For researchers: this study fills a geographic gap that is larger than it appears. The DHH wearable literature is almost entirely North American. A 26-participant study in West Africa surfaces design priorities, specifically acoustic localisation and community data collection, that do not appear in the literature. More field studies in low- and middle-income countries (LMICs) are needed.

For practitioners: a USD 30 offline device is technically feasible as a research prototype and proof of concept. Significant engineering work remains before the system is suitable for independent use. The primary outstanding challenge is reducing false positives, which requires a substantially larger and more acoustically representative training dataset than the approximately 200 samples per class used in this study. Organisations seeking to adopt this architecture should plan for a dedicated local data collection phase before any deployment.

For Policymakers: WHO and UNICEF (2022) called for the expansion of assistive technology (AT) access in LMICs. This work offers one concrete pathway. Edge AI on commodity microcontrollers, trained on local data, does not require cloud infrastructure or high-income supply chains. It can be replicated by any organisation willing to conduct the fieldwork.

LIMITATIONS AND FUTURE DIRECTIONS

EchoBand was evaluated with 26 participants in one West African city. The sample is small. Findings from Porto-Novo do not automatically apply to rural Benin, Anglophone West Africa, or other LMIC contexts. No broader claims are made beyond what the data supports.

Field accuracy was not measured quantitatively. Laboratory validation figures were reported alongside qualitative field observations. These represent different types of evidence. Future work should deploy in-field logging that records model outputs and participant-confirmed correctness in real environments.

Field deployment additionally revealed a persistent false positive problem that is not captured by laboratory validation figures. The system triggered vibration alerts for sounds that did not correspond to a target event, as reported across multiple participants in acoustically busy environments. Post-deployment analysis identified two contributing causes. First, training data for each sound class comprised approximately 200 samples, of which only 30 to 60 were authentic field recordings from Benin; the remainder were drawn from FreeSound and AudioSet, introducing an acoustic domain mismatch between training and deployment conditions. Second, the four-class-per-model architecture concentrates classification boundaries across a narrow sound space, making the model vulnerable to out-of-distribution acoustic events that resemble a target class. Resolving the false-positive problem requires a dedicated second data-collection phase in Benin, targeting substantially larger per-class samples from authentic environments.

The study also encountered practical challenges in conducting the research. Data collection required international travel and extended fieldwork in Porto-Novo. All participant interactions were conducted through a sign language interpreter, as the researcher is not proficient in the local sign language, and participants communicate primarily through sign. Environmental sound recording in Benin yielded between 30 and 60 samples per class, insufficient in isolation for robust model training and requiring supplementation with FreeSound and AudioSet data. This introduces acoustic domain differences that may partially account for any gap between laboratory and field performance. Recruiting through a single centre enabled close collaboration but limited sample size and diversity. These challenges are characteristic of early-stage assistive technology research in low-resource settings and are reported here to support replication.

The community contribution mechanism needs a dedicated longitudinal study to assess contribution rates, data quality over time, and whether community-contributed data actually improves model accuracy in practice.

Immediate priorities are a comprehensive second round of data collection in Benin, targeting at least 500 samples per sound class from authentic field environments, as the current dataset of approximately 200 samples per class is insufficient to match laboratory accuracy under real-world conditions; and systematic in-field measurement of false-positive rates across varied noise environments. Hardware alternatives with better wearable form factors also warrant evaluation. The LILYGO T-Watch S3, based on the ESP32-S3 microcontroller (16 MB flash, 8 MB pseudo-static random-access memory (RAM), integrated microphone, BLE 5.0, DRV2605 haptic motor driver, watch form factor, approximately USD 43), directly addresses participant feedback on device aesthetics and offers an improved platform for the next EchoBand iteration. Longer-term directions include expanding field evaluation to other West African cities, open-sourcing the model pipeline, and testing cross-community data sharing across the contribution network.

CONCLUSIONS

EchoBand and Pulses demonstrate that affordable, standalone, offline assistive technology for DHH users in low-income settings is technically feasible as a research prototype. The system is not production-ready. Field deployment revealed a false-positive rate attributable to insufficient environment-specific training data and the narrow class boundaries of small models, each covering four sounds. These are correctable data and architecture challenges, not fundamental limitations of on-device inference. Prior work concluded that watch-only inference was unusable.

This study shows that the right architectural choice — deploying four small, context-specific models rather than one large general classifier — achieves 267 milliseconds latency, a 22-fold improvement over prior art. Matching that latency gain with equivalent reliability in real-world conditions requires a substantially larger and more representative training dataset than was available for this first iteration.

The technical contribution is real, but the more significant finding is geographic and methodological. This is the first peer-reviewed evaluation of DHH sound-awareness technology with participants in sub-Saharan Africa. The study produced a community-based personalisation model that no prior system has attempted and reversed a technical conclusion, specifically that watch-only inference is unusable, which the field had accepted for several years.The broader lesson is clear. Designing for the actual constraints of a context produces better outcomes than designing for ideal conditions and adapting afterwards. Every key decision, from USD 30 hardware and four-model modular architecture to pre-departure context selection and community data collection, emerged directly from field constraints. Those constraints were not obstacles; they were the design brief.

REFERENCES

Benkler, Y. (2006): The Wealth of Networks: How Social Production Transforms Markets and Freedom. New Haven: Yale University Press. Available at: https://www.benkler.org/Benkler_Wealth_Of_Networks.pdf

Chadha, S., Kamenov, K., and Cieza, A. (2021): The world report on hearing, 2021. Bulletin of the World Health Organisation, Vol. 99, pp. 242-242A. Available at: https://doi.org/10.2471/BLT.21.285643

David, R., Duke, J., Jain, A., Janapa Reddi, V., Jeffries, N., Li, J., Kreeger, N., Nappier, I., Natraj, M., Wang, T., Warden, P., and Wiesler, S. (2021): TensorFlow Lite Micro: Embedded machine learning for TinyML systems. Proceedings of Machine Learning and Systems (MLSys), Vol. 3, pp. 800-811. Available at: https://doi.org/10.48550/arXiv.2010.08678

Do, H., Dang, Q., Huang, J.Z., and Jain, D. (2023): AdaptiveSound: An interactive feedback-loop system to improve sound recognition for deaf and hard-of-hearing users. Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2023), New York: ACM, pp. 1-13. Available at: https://doi.org/10.1145/3597638.3608390

Findlater, L., Chinh, B., Jain, D., Froehlich, J., Kushalnagar, R., and Lin, A.C. (2019): Deaf and hard-of-hearing individuals' preferences for wearable and mobile sound awareness technologies. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, New York: ACM, pp. 1-13. Available at: https://doi.org/10.1145/3290605.3300276

Flores Ramones, A. and Del-Rio-Guerra, M.S. (2023): Recent developments in haptic devices designed for hearing-impaired people: A literature review. Sensors, Vol. 23, No. 6, Article 2968. Available at: https://doi.org/10.3390/s23062968

Goodman, S.M., McDonnell, E.J., Froehlich, J.E., and Findlater, L. (2025): SPECTRA: Personalizable sound recognition for deaf and hard-of-hearing users through interactive machine learning. Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems, New York: ACM, pp. 1-16. Available at: https://doi.org/10.1145/3706598.3713294

Hevner, A., March, S., Park, J., and Ram, S. (2004): Design science in information systems research. MIS Quarterly, Vol. 28, No. 1, pp. 75-105. Available at: https://doi.org/10.2307/25148625

Heydari, S. and Mahmoud, Q.H. (2025): Tiny machine learning and on-device inference: A survey of applications, challenges, and future directions. Sensors, Vol. 25, No. 10, Article 3191. Available at: https://doi.org/10.3390/s25103191

Huang, J.Z., Wood, R., Chhabria, H. and Jain, D. (2023): 'Not there yet': Feasibility and challenges of mobile sound recognition to support deaf and hard-of-hearing people. Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2023), New York: ACM, pp. 1-15. Available at: https://doi.org/10.1145/3597638.3608431

Jain, D., Ngo, H., Patel, P., Goodman, S., Findlater, L., and Froehlich, J.E. (2020): SoundWatch: Exploring smartwatch-based deep learning approaches to support sound awareness for deaf and hard-of-hearing users. Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2020), New York: ACM, pp. 1-13. Available at: https://doi.org/10.1145/3373625.3416991

Jain, D., Nguyen, K.H.A., Goodman, S.M., Grossman-Kahn, R., Ngo, H., Kusupati, A., Du, R., Olwal, A., Findlater, L., and Froehlich, J.E. (2022): ProtoSound: A personalized and scalable sound recognition system for deaf and hard-of-hearing users. Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, New York: ACM, pp. 1-16. Available at: https://doi.org/10.1145/3491102.3502020

Mohamad, M., Balıkel, A.E., Chowdhury, D., Russell, C., and Mahmoud, S. (2026): Artificial Intelligence for Business Model Innovation in Digital Sustainability Companies: Multiple Cases from the Netherlands. World Journal of Entrepreneurship, Management and Sustainable Development, Vol. 22, Nos. 1-2, pp. 71–87. Available at: https://doi.org/10.47556/J.WJEMSD.22.1-2.2026.5

Mulwafu, W., Kuper, H., and Ensink, R.J.H. (2016): Prevalence and causes of hearing impairment in Africa. Tropical Medicine and International Health, Vol. 21, No. 2, pp. 158-165. Available at: https://doi.org/10.1111/tmi.12640

Olson, M. (1965): The Logic of Collective Action: Public Goods and the Theory of Groups. Cambridge: Harvard University Press. Available at: https://doi.org/10.4159/9780674041660

Orzech, G., Luo, Y., and Huang, G. (2024): Haptic technology for hearing loss: A systematic review of technical feasibility, usability, and user experience. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Vol. 68, No. 1. Available at: https://doi.org/10.1177/10711813241275928

Ostrom, E. (1990): Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge: Cambridge University Press. Available at: https://doi.org/10.1017/CBO9780511807763

Rusci, M. and Tuytelaars, T. (2023): Few-shot open-set learning for on-device customisation of keyword spotting systems. Proceedings of Interspeech 2023, ISCA, pp. 2768-2772. Available at: https://doi.org/10.21437/Interspeech.2023-1904

Rusci, M., Paci, F., Fariselli, M., Flamand, E., and Tuytelaars, T. (2024): Self-learning for personalized keyword spotting on ultra-low-power audio sensors. IEEE Internet of Things Journal, Vol. 12, pp. 10210-10221. Available at: https://doi.org/10.1109/JIOT.2024.3515143

Sumitani, T. (n.d.): Tankyu Practice: Inquiry-based methodology for real-world problem solving. Kobe Institute of Computing. Available at: https://www.kic.ac.jp/eng/tankyu.html (Accessed: April 2026).

United Nations (UN), United Nations International Children's Emergency Fund (UNICEF), and World Health Organisation (WHO). (2022): Global Report on Assistive Technology: Creating Opportunities and Removing Barriers for Persons with Disabilities. Geneva: World Health Organisation. Available at: https://www.who.int/publications/i/item/9789240049451

Uwizeyimana, D.E. (2024): How can artificial intelligence technologies enhance public policymaking? World Journal of Entrepreneurship, Management and Sustainable Development, Vol. 20, Nos. 3/4, pp. 299–312. Available at: https://doi.org/10.47556/J.WJEMSD.20.3-4.2024.7

Warden, P. and Situnayake, D. (2019): TinyML: Machine Learning with TensorFlow Lite on Arduino and Ultra-Low-Power Microcontrollers. Sebastopol: O'Reilly Media. Available at: https://tinymlbook.com/wp-content/uploads/2020/11/TinyML_preview.pdf

World Health Organisation (WHO). (2021): World Report on Hearing. Geneva: World Health Organisation. Available at: https://www.who.int/publications/i/item/world-report-on-hearing

BIOGRAPHY

Cyrille Awatchede Ange Noutchogbe is a graduate researcher in the Information and Communication Technology (ICT) Innovation programme at the Kobe Institute of Computing (KIC), Japan, supervised by Professor Takahara Toshiro. He is a recipient of the Japan International Cooperation Agency (JICA) African Business Education (ABE) Initiative scholarship. His research focuses on embedded machine learning, edge AI, and the design of affordable assistive technology for deaf and hard-of-hearing communities in low-income settings.