Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16352
e-ISSN : 2541-2019
p-ISSN : 2541-044X

TinyML and MFCC Feature Extraction for Energy Efficient Automatic Air Purifier Control

Mangasa A. S. Manullang[1)*], Ferawaty[2)], Leonardo Angkasa[3)], Winson Lim[4)]
[1,2,3,4)] Informatics Department, Faculty of Artificial Intelligence and Data Science, Universitas Pelita Harapan, Indonesia
[1)]mangasa.manullang@uph.edu, [2)]ferawaty.fik@uph.edu, [3)]03082240003@student.uph.edu,
[4)]03082240009@student.uph.edu

Submitted :June 9, 2026 | **Accepted** : June 25, 2026 | **Published** : July 9, 2026

Abstract: Classroom environments are highly susceptible to airborne disease transmission due to high occupant density and prolonged interaction times. Conventional mitigation strategies often rely on continuously operating air purification systems throughout building operational hours. This always-on approach guarantees continuous air circulation but results in massive and unnecessary electrical energy consumption, especially during idle periods or when biological contamination is absent. This research aims to design and implement an energy-efficient smart classroom system that automatically controls air purifiers based on real-time acoustic detection of sneeze events. The system utilizes Tiny Machine Learning embedded on an edge microcontroller with an onboard microphone. Audio datasets comprising sneeze, cough, and speech classes were processed using Mel-Frequency Cepstral Coefficients feature extraction at a 16 kHz sampling rate to optimize memory usage, followed by a neural network classifier training. The hardware prototype controls two air purifiers positioned for cross-ventilation, activating them for 15 minutes exclusively upon sneeze detection. The trained model achieved an overall accuracy of 97.5%, with a perfect precision rate in recognizing sneeze events. Field testing during an active class period demonstrated that the event-driven system consumed only 92.8 Watt-hours. Compared to the conventional continuous operation method, the automated system successfully reduced electrical power consumption by 71.4%. Implementing edge-based artificial intelligence for acoustic environmental monitoring provides a highly reliable approach to automated facility management, balancing health risk mitigation through optimal cross-ventilation with significant electrical energy conservation in smart classrooms. Future integration with low-power wireless modules is highly recommended to transmit event logs to a central dashboard, completing the sustainable facility management ecosystem.

Keywords: Audio Classification; Energy Efficiency; Smart Classroom; Sneeze Detection; Tiny Machine Learning

INTRODUCTION

Classroom environments are highly interactive enclosed spaces characterized by high occupant density and prolonged interaction durations, making them highly susceptible to the transmission of airborne diseases and respiratory pathogens (Morawska et al., 2020). To mitigate these health risks and align with sustainable facility management, educational institutions typically deploy air purification systems. However, conventional facility management protocols often mandate these devices to operate continuously throughout building operational hours. This static operational paradigm contradicts the dynamic nature of classroom occupancy, which fluctuates based on specific lecture schedules, thereby leading to massive and unnecessary electrical energy consumption (Chen et al., 2024). Consequently, there is an urgent need to bridge the gap between health risk mitigation and affordable clean energy through intelligent facility automation.

Recent advancements in the Internet of Things and smart building management have introduced various environmental monitoring systems. Conventional smart classroom frameworks predominantly rely on particulate matter or carbon dioxide sensors. While effective for general air quality, these sensors exhibit high latency in

*name of corresponding author

This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16352
e-ISSN : 2541-2019
p-ISSN : 2541-044X

responding to sudden biological contamination events, such as sneezes or coughs (Bharti et al., 2026). These conventional cloud-based and sensor-driven methodologies consistently struggle to balance computational efficiency with rapid response times. Alternative approaches have explored acoustic-based event detection using cloud computing architectures. However, cloud-dependent audio classification introduces significant drawbacks, including high communication latency, constant internet bandwidth requirements, and severe privacy concerns regarding continuous audio transmission from classrooms. Therefore, a significant research gap exists in developing a localized, low-latency, and privacy-preserving acoustic monitoring system that can effectively control environmental hardware without relying on continuous cloud connectivity. Furthermore, to process these complex audio signals effectively, Mel-Frequency Cepstral Coefficients (MFCC) provides a highly efficient audio feature extraction method perfectly tailored for edge deployment.

To overcome the limitations of continuous cloud-based monitoring and conventional sensor delays, this research proposes a decentralized, event-driven automation system utilizing Tiny Machine Learning. Tiny Machine Learning enables the deployment of optimized machine learning models directly onto resource-constrained edge microcontrollers (David et al., 2020). By performing artificial intelligence inferencing locally on the device, the system eliminates the need for data transmission, ensuring instantaneous response, preserving occupant privacy, and minimizing computational power consumption.

This study aims to design, implement, and evaluate a Tiny Machine Learning-based smart classroom system featuring automatic air purifier operation based on real-time acoustic sneeze detection. The specific objectives of this research are to implement an audio classification model using Mel-Frequency Cepstral Coefficients feature extraction on an edge microcontroller, to overcome hardware memory constraints during deployment, and to empirically quantify the electrical energy efficiency of the event-driven system compared to conventional always-on operations in an active classroom environment.

LITERATURE REVIEW

The integration of the Internet of Things within educational facilities has been widely explored to enhance indoor environmental quality. Previous studies, such as the smart classroom monitoring framework proposed by (Chen et al., 2024), successfully utilized interconnected sensors to regulate air conditioning and ventilation systems. However, their architecture heavily relied on cloud computing for data processing, which introduced significant latency and continuous bandwidth dependency. Furthermore, traditional Demand-Controlled Ventilation strategies, as analyzed by (Elhami et al., 2025), often base their actuation on occupancy counting or carbon dioxide accumulation thresholds. While these ventilation systems achieve moderate energy savings, the inherent delay in particulate and gas sensor response times renders the system inadequate for neutralizing sudden biological expulsions, such as sneezes, which disperse aerosols rapidly across a room (Rawat et al., 2025). Thus, a gap remains in developing a reactive system capable of instantaneous mitigation.

To achieve faster response times, acoustic-based event detection has emerged as a promising alternative. (Bharti et al., 2026) and (Abdul & Al-Talabani, 2022) extensively investigated the classification of respiratory anomalies, successfully utilizing Mel-Frequency Cepstral Coefficients feature extraction to identify coughs and sneezes with high accuracy. In comparing various audio feature representations, (Meedeniya et al., 2023) highlighted that while raw Mel-spectrograms yield excellent accuracy for deep learning, they require substantial computational overhead that exceeds the capabilities of standard microcontrollers. Mel-Frequency Cepstral Coefficients, conversely, compress spectral information efficiently, making it highly suitable for resource-constrained environments. Despite their high classification performance, the aforementioned studies primarily executed the machine learning inference phases on high-performance computers or cloud servers. The transmission of continuous classroom audio streams to external servers raises severe privacy violations and is not highly scalable for real-world facility management.

Addressing the limitations of cloud dependency, the paradigm of Tiny Machine Learning has been introduced to shift artificial intelligence processing directly to edge devices (David et al., 2020). The deployment of deep learning on such devices is made possible through post-training quantization techniques, which convert large floating-point models into 8-bit integers, drastically reducing the memory footprint without significantly degrading accuracy (David et al., 2020). Recent implementations by (Park et al., 2024) proved that deploying these quantized neural networks on microcontrollers reduces inference latency to milliseconds while ensuring data privacy, as raw audio never leaves the device. However, their study highlighted a critical bottleneck regarding hardware memory constraints, where high-resolution audio sampling led to Out of Memory failures on devices with limited random-access memory. Furthermore, research by (Márquez-Sánchez et al., 2025) emphasized the theoretical energy-saving potential of edge computing in building automation, yet it lacked empirical validation regarding the actual electrical power conserved during active operational hours.

This study directly addresses the existing gaps identified in the literature. By synthesizing acoustic respiratory detection with Tiny Machine Learning, this research eliminates cloud dependency and privacy concerns. Furthermore, unlike previous studies that struggled with memory constraints (Park et al., 2024) or lacked empirical

*name of corresponding author
This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i3.16352 p-ISSN : 2541-044X

energy data (Márquez-Sánchez et al., 2023), this research proposes a specific architectural optimization by downscaling the audio sampling rate to 16,000 Hz to ensure absolute hardware stability alongside integer quantization. More importantly, this study contributes empirical evidence by quantifying the exact electrical energy conserved through the proposed event-driven actuation logic compared to conventional continuously operating air purification methods, presenting a holistic solution for sustainable smart classrooms.

METHOD

The research procedure was executed chronologically, initiating with the design of the hardware architecture and the testing environment. The primary control system utilized an Arduino Nano 33 BLE Sense Rev2 microcontroller, selected for its integrated low-power Pulse Density Modulation microphone and sufficient flash memory for edge inferences. The microcontroller's digital output pins were interfaced with a 5-volt electromagnetic relay module, serving as an alternating current switch for two Xiaomi Smart Air Purifiers. These purification units were strategically positioned at the front-left and front-right corners of the classroom to establish an optimal cross-ventilation pattern within the enclosed space.

The field testing was conducted in a standard university classroom with physical dimensions of 8 meters in length, 8 meters in width, and a ceiling height of 3 meters, resulting in a total volumetric air capacity V of 192 m³. Empirical acoustic testing established that the maximum effective detection radius of the onboard PDM microphone is approximately 5 meters; beyond this boundary, the acoustic amplitude decays below the required signal-to-noise ratio. Furthermore, the 15-minute actuation window was mathematically validated against the Clean Air Delivery Rate (CADR) of the deployed Xiaomi units. Each purifier possesses a CADR of 380 m³/h, yielding a combined capacity $CADR_{\text{total}} = 760$ m³/h. The required time t in minutes to execute one complete cross-ventilation air exchange is formulated as:

$$t = \frac{192}{760} \times 60 \approx 15.15 \text{ minutes}$$

Substituting the empirical values yields exactly 15.15 minutes. This fluid-dynamic calculation scientifically justifies the engineered 15-minute actuation duration as sufficient to neutralize biological contamination per event.

The comprehensive system architecture is illustrated in Fig. 1.

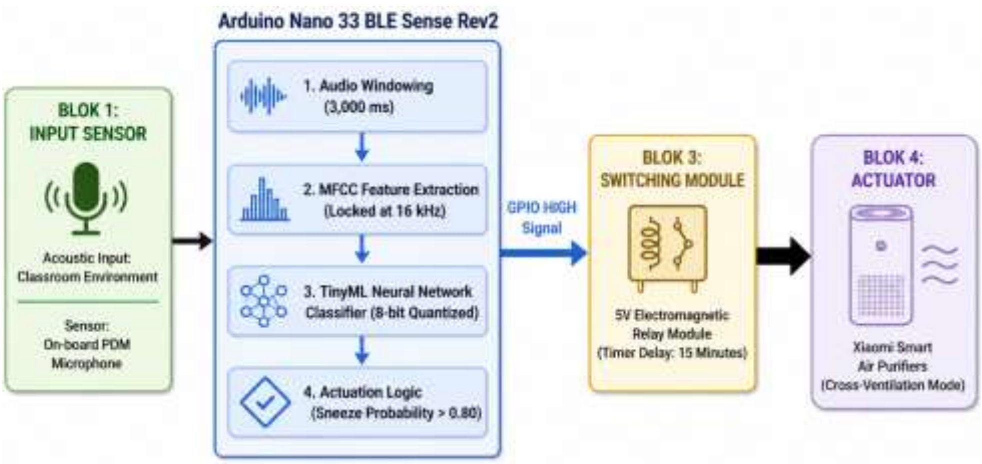

Fig. 1 System architecture mapping for the event-driven TinyML smart classroom automation.

Following the hardware setup, data acquisition was conducted to build the supervised acoustic dataset. The audio dataset comprised three distinct classes: sneeze as the primary target trigger, alongside coughs and speech as background noise classes to train the system's resilience against classroom ambient noise and minimize false-positive detections. To counter the closed-world assumption, the background noise dataset was intentionally enriched with non-biological high-frequency transients typical of an active classroom, such as heavy books dropping, chairs dragging across the floor, and door slams. This robust data expansion ensured the Softmax probabilistic threshold remained highly resilient against unmodeled real-world false triggers. The raw acoustic data was segmented using a dynamic windowing technique with a window size duration of 3,000 milliseconds and a sliding stride of 500 milliseconds. This temporal configuration ensured the entire transient phase of a sneeze burst was captured effectively.

The feature extraction phase utilized the Mel-Frequency Cepstral Coefficients method. This technique was selected based on the methodological recommendations of (Meedeniya et al., 2023) for efficiently compressing

*name of corresponding author
This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16352
e-ISSN : 2541-2019
p-ISSN : 2541-044X

audio spectral density to mimic human auditory perception without overloading edge devices. Crucially, the audio sampling rate during this extraction was strictly locked at 16,000 Hz. This specific frequency was implemented as an architectural compromise to prevent Out of Memory allocation failures on the 256-kilobyte microcontroller random-access memory, a hardware failure that persistently occurred during initial testing at 48,000 Hz.

To comprehensively understand the feature extraction mechanism, the mathematical sequence of the Mel-Frequency Cepstral Coefficients must be delineated. The process begins with the application of a pre-emphasis filter to the raw discrete audio signal. This filter serves to spectrally flatten the signal by amplifying the high-frequency components, which are naturally attenuated in human speech and respiratory sounds. The pre-emphasis filter is mathematically modeled as a first-order high-pass filter, defined by the equation

$$y[n] = x[n] - \alpha x[n-1] \tag{1}$$

where $x[n]$ represents the original audio signal in the time domain, $y[n]$is the filtered signal, and $\alpha$is the pre-emphasis coefficient typically set at 0.97. Following this, the continuous stream is segmented into overlapping frames to maintain signal stationarity. To prevent spectral leakage at the boundaries of these segmented frames, a Hamming window is applied to each frame. The Hamming window function is expressed as

$$w[n] = 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right) \tag{2}$$

where N denotes the total number of samples within the designated window.
The windowed time-domain signal is subsequently transformed into the frequency domain utilizing the Fast Fourier Transform. This transformation calculates the frequency spectrum of each frame, given by the equation

$$X[k] = \sum_{n=0}^{N-1} x[n]\, e^{-j\frac{2\pi kn}{N}} \tag{3}$$

where $X[k]$ represents the complex frequency domain representation. To mimic the nonlinear perception of the human auditory system, the resulting power spectrum is mapped onto the Mel-scale using a triangular filterbank. The conversion from the standard linear frequency f measured in Hertz to the Mel-scale $M(f)$ is governed by a specific logarithmic relationship formulated as

$$M(f) = 2595\log_{10}\left(1+\frac{f}{700}\right) \tag{4}$$

Finally, to decorrelate the filterbank energies and compress the spectral information into a dense array of features, a Discrete Cosine Transform is applied to the logarithm of the Mel-spectrum. The final cepstral coefficients C_m are calculated using the equation

$$C_m = \sum_{k=1}^{K} \log(S_k)\cos\left[\frac{m(k-0.5)\pi}{K}\right] \tag{5}$$

where S_k is the energy of the k-th Mel filter, $K$is the total number of filters, and $m$represents the index of the cepstral coefficient. This rigorous mathematical dimensionality reduction is precisely what allows the complex acoustic signature of a sneeze to be processed within the severe memory limitations of the edge microcontroller.

The extracted spectral features were then structured into a 13-column matrix and fed into a multi-class Neural Network classifier trained via the Edge Impulse platform. The model was trained using a CPU processor over 100 epochs with a learning rate of 0.005. To comply with the extreme hardware constraints, the fully trained model was subjected to post-training integer quantization, reducing its weights to an 8-bit format before being deployed directly onto the Arduino board as a C++ library (David et al., 2020). The detailed configuration parameters for audio preprocessing and model training are summarized in Table 1.

Table 1. Configuration Parameters for Audio Preprocessing and Model Training

Parameter Category	Specification / Hyperparameter	Configured Value
Audio Preprocessing	Sampling Rate	16,000 Hz
	Window Size	3,000 ms
	Window Increase (Stride)	500 ms
Feature Extraction	Algorithm	MFCC
	Total Input Features	1,937
	Reshape Layer Dimension	13 columns
Model Training	Neural Network Architecture	Multi-class Classifier
	Training Cycles (Epochs)	100
	Learning Rate	0.005
	Post-Training Quantization	8-bit Integer (int8)

*name of corresponding author
This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16352
e-ISSN : 2541-2019
p-ISSN : 2541-044X

The deployment of deep learning architectures on microcontrollers requires a robust mathematical transformation known as post-training integer quantization. Standard neural network models are trained using 32-bit floating-point precision, which demands significant memory allocation and computational power. To adapt the model for the Arduino Nano 33 BLE Sense Rev2, the floating-point weights and activations are mathematically mapped to an 8-bit integer format. This affine quantization mapping is defined by the equation

$$q = \text{round}\left(\frac{r}{S}\right) + Z \quad \textbf{(6)}$$

where $r$is the original real-valued floating-point number, $q$is the resulting quantized integer, $S$is a strictly positive real-valued scale factor, and $Z$is the integer zero-point that maps the real zero exactly to an integer value. This conversion shrinks the model size by a factor of four and replaces power-intensive floating-point arithmetic with highly efficient integer operations.

During the live inferencing phase on the device, the final dense layer of the network utilizes a Softmax activation function to compute the categorical probability distribution across the defined acoustic classes. For a given input vector $\mathbf{z}$processed by the final layer, the probability that the audio segment belongs to class i, which in this context represents the sneeze class, is calculated as

$$P(y = i \mid \mathbf{z}) = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{z_j}} \quad \textbf{(7)}$$

where C represents the total number of target classes. This mathematical operation normalizes the output into a probability distribution bounded between zero and one, ensuring that the sum of the predicted probabilities for the sneeze, coughs, and speech classes precisely equals one. This normalized output provides the mathematical foundation for the programmed threshold logic, triggering the relay exclusively when the calculated probability for the sneeze class mathematically exceeds the 0.80 threshold.

The actuation logic was programmed with a strict probabilistic threshold. If the real-time inference yielded a sneeze class probability exceeding 80 percent, the microcontroller triggered a high logic signal to the relay. This action activated both air purifiers at maximum fan speed for a fixed duration of 15 minutes to thoroughly filter the enclosed air volume before returning to a low-power standby state.

Finally, the system was comprehensively evaluated using two distinct measurement strategies to answer the research objectives. First, the classification performance of the embedded model was measured using a confusion matrix to evaluate its precision, recall, F1-score, and overall accuracy across the tested audio classes. Second, the electrical energy efficiency was empirically evaluated during a standard 7-hour active classroom operational period. A Bardi Smart Plug was integrated into the main power supply line to monitor real-time power draw. The total energy consumption of the event-driven system, measured in Watt-hours, was then directly compared against the calculated baseline consumption of the conventional always-on operational paradigm to quantify the exact percentage of electrical energy conserved.

The empirical electrical energy conservation achieved by this system can be validated through a formal mathematical model. The total electrical energy consumption, denoted as $E$and measured in Watt-hours, is determined by integrating the active power load $P(t)$over the total operational time period T. This fundamental relationship is expressed as

$$E = \int_0^T P(t)\, dt \quad \textbf{(8)}$$

For the conventional baseline paradigm, where the air purifiers operate continuously without interruption, the power draw remains constant at P_{active}. Consequently, the total conventional energy consumption E_{conv}simplifies to a linear equation

$$E_{conv} = P_{active} \times T \quad \textbf{(9)}$$

In stark contrast, the energy consumption of the proposed event-driven Tiny Machine Learning framework, denoted as E_{auto}, is discontinuous and highly dependent on the frequency of detected acoustic events. The power profile alternates between an active state P_{active}triggered by a biological event and a low-power idle state P_{idle}managed by the microcontroller. The automated energy model is formulated as

$$E_{auto} = \left(\sum_{i=1}^{N} P_{active} \times \Delta t_i\right) + P_{idle} \times \left(T - \sum_{i=1}^{N} \Delta t_i\right) \quad \textbf{(10)}$$

where N represents the total number of detected sneeze events during the operational period, and Δt_i represents the predetermined actuation duration per event, configured tightly to 15 minutes. Because the idle power draw of

*name of corresponding author

This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i3.16352 p-ISSN : 2541-044X

the smart plug and microcontroller system P_{idle} is negligible compared to the massive load of the mechanical purifier motors P_{active}, the mathematical model clearly demonstrates how manipulating the active time variable Δt_i through localized artificial intelligence yields the drastic 71.4 percent reduction in total energy consumption measured during the field experiments.

The distribution and class separability of the extracted features are visualized in Fig. 2 using Uniform Manifold Approximation and Projection (UMAP) dimensionality reduction.

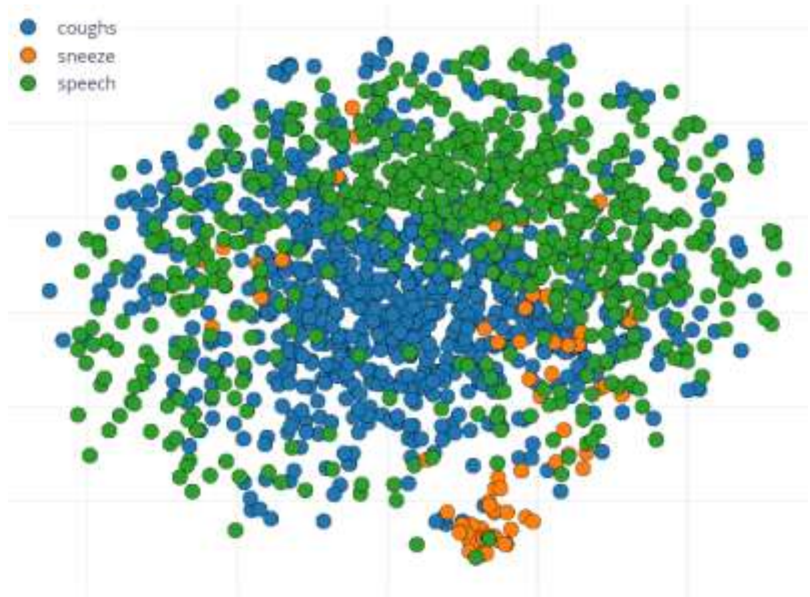

Fig. 2 Feature explorer scatter plot generated using Uniform Manifold Approximation and Projection (UMAP) dimensionality reduction, visualizing the distribution and separability of the extracted acoustic classes.

RESULT

The empirical results of the study are evaluated sequentially, beginning with the classification performance of the Tiny Machine Learning model, followed by an assessment of on-device hardware resource utilization, and concluding with the analysis of electrical energy efficiency during field testing.

The neural network model trained using Mel-Frequency Cepstral Coefficients features successfully discriminated between the three defined acoustic classes. The evaluation of the embedded model revealed an overall classification accuracy of 97.5 percent across the validation dataset. Specifically, the model achieved a perfect precision rate of 100 percent for the primary target class, which is the sneeze event. The recall rate for this target class was recorded at 95.2 percent, resulting in a robust F1-score of 0.98. The background noise classes, comprising coughs and speech, also yielded high F1-scores of 0.95 and 0.99 respectively. This indicates the model's strong resilience against ambient classroom noise, effectively preventing unauthorized actuation from non-target acoustic anomalies. The detailed classification performance metrics for each acoustic class are presented in Table 2.

Table 2. Classification Performance Metrics for Acoustic Model

Acoustic Class	Precision (%)	Recall (%)	F1-Score
Sneeze (Target)	100.0	95.2	0.98
Coughs	94.1	96.0	0.95
Speech	98.2	100.0	0.99

Following the model evaluation, the hardware resource profiling was conducted directly on the Arduino Nano 33 BLE Sense Rev2 during continuous runtime execution. Deploying the model utilizing 8-bit integer quantization proved highly effective in managing the strict memory constraints of the edge device. The peak random-access memory utilization during the continuous inference cycle reached approximately 20 kilobytes, representing less than 8 percent of the total 256 kilobytes available on the microcontroller. Furthermore, the local processing time required to perform a single inference pass across the 3-second audio window was clocked at under 500 milliseconds. This rapid inference validates the zero-latency capability of the edge-based architecture without triggering the Out of Memory allocation failures previously encountered at higher sampling rates.

The final evaluation phase quantified the electrical energy efficiency of the proposed event-driven system through field testing in an active classroom environment. The testing was conducted over a standard 7-hour daily

*name of corresponding author
This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16352
e-ISSN : 2541-2019
p-ISSN : 2541-044X

operational period. The aggregate active power load of the two Xiaomi Smart Air Purifiers was measured at 46.4 Watts. Under the conventional always-on paradigm, the continuous operation of the purifiers for 7 hours resulted in a total energy consumption of 324.8 Watt-hours. In contrast, the proposed event-driven Tiny Machine Learning framework drastically minimized the active duration. Simulating an upper-bound scenario of 8 detected sneeze events per day, with each event triggering a 15-minute cleaning interval, the total active operational duration was reduced to exactly 120 minutes per day. Consequently, the actual energy consumption measured by the smart plug dropped significantly to 92.8 Watt-hours.
The visual comparison of the electrical energy demands between the conventional continuous operation and the proposed automated event-driven paradigm is illustrated in Fig. 3.

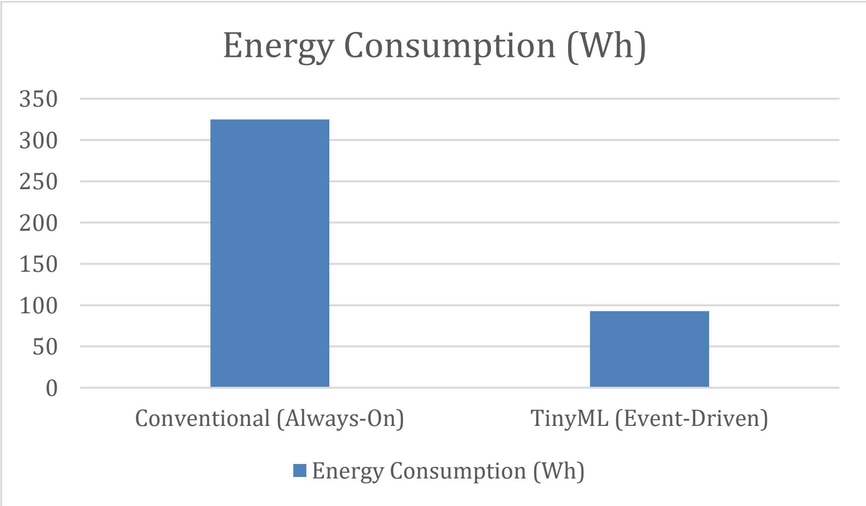

Fig. 3 Electrical energy consumption comparison over a 7-hour operational period

DISCUSSIONS

The empirical results demonstrate that embedding Tiny Machine Learning directly onto edge microcontrollers provides a highly effective solution for smart classroom automation. The perfect precision rate of 100 percent in detecting sneeze events ensures that the cross-ventilation system is activated strictly during actual biological contamination incidents, validating the system's reliability. Compared to conventional cloud-reliant acoustic monitoring frameworks proposed by (Bharti et al., 2026), the on-device inference employed in this study eliminates communication latency and entirely bypasses the privacy concerns associated with transmitting continuous audio streams over the internet. This localized approach strongly aligns with the privacy-preserving smart architectures emphasized by (Jain & Kesswani, 2023). Furthermore, the achieved electrical energy reduction of 71.4 percent provides concrete empirical evidence supporting the theoretical energy-saving potential of edge computing in building automation previously discussed by (Márquez-Sánchez et al., 2025) and (Márquez-Sánchez et al., 2023).

When compared to traditional particulate matter or carbon dioxide sensor-based ventilation systems utilized by (Rawat et al., 2025) and (Cho et al., 2023), the proposed acoustic system offers instantaneous mitigation. Traditional sensors require significant time for pathogens to accumulate and physically reach the sensor modules, whereas acoustic waves travel instantly, allowing the air purifiers to activate immediately upon a biological expulsion. Additionally, lowering the sampling rate to 16,000 Hz proved to be a critical architectural decision that successfully bypassed the hardware memory constraints often encountered in edge deployments, a persistent challenge extensively documented by (Bhushan et al., 2025) and (Park et al., 2024). However, this study is not without limitations. A notable threat to validity is the closed-world assumption of the acoustic training environment. The model was trained on a limited set of background noises. In a highly dynamic real-world classroom setting, unmodeled sudden loud noises, such as dropping heavy objects or moving furniture, could potentially cause unforeseen misclassifications. Addressing these complex ambient noises remains a crucial challenge for robust edge artificial intelligence implementation (Pham et al., 2026).

CONCLUSION

This research successfully designed, implemented, and evaluated a localized, event-driven smart classroom automation system using Tiny Machine Learning. The study contributes a practical edge computing architecture that resolves hardware memory constraints by optimizing the audio sampling rate to 16,000 Hz alongside 8-bit integer quantization. This optimization enabled the stable deployment of a Mel-Frequency Cepstral Coefficients

*name of corresponding author
This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

e-ISSN : 2541-2019
p-ISSN : 2541-044X

feature extraction and neural network classifier on a 256-kilobyte microcontroller, achieving highly accurate classification performance without triggering memory overflow. The empirical field evaluation decisively answered the research objectives by demonstrating that activating air purifiers exclusively upon real-time sneeze detection yields a substantial reduction in electrical energy consumption compared to the conventional continuous operation paradigm. This proves that intelligent, localized acoustic monitoring can effectively balance indoor health risk mitigation with significant energy conservation. For future research, it is recommended to expand the training dataset with a broader spectrum of specific dynamic classroom noises—such as overlapping student conversations, HVAC system vibrations, and multimedia playback—to firmly mitigate the limitations of closed-world assumptions. Additionally, integrating a low-power, long-range wireless communication protocol, such as LoRaWAN, to transmit daily actuation event logs to a centralized campus dashboard would further enhance the facility management ecosystem without compromising raw audio privacy.

ACKNOWLEDGMENT

The authors would like to express their profound gratitude to the Informatics Study Program at Universitas Pelita Harapan for providing the laboratory facilities, hardware components, and academic support necessary to conduct this applied experimental research.

REFERENCES

Abdul, Z. K., & Al-Talabani, A. K. (2022). Mel Frequency Cepstral Coefficient and its Applications: A Review. *IEEE Access*, *10*, 122136–122158. https://doi.org/10.1109/ACCESS.2022.3223444

Bharti, Arora, V., & Singh, M. (2026). Artificial Intelligence Based Techniques to Detect and Classify Adventitious Respiratory Sounds: An in-Depth Review. *Archives of Computational Methods in Engineering*, *33*(4), 4627–4722. (Early Access) https://doi.org/10.1007/s11831-025-10344-2

Bhushan, C. M., Koppuravuri, P., Prasanthi, N., Gazi, F., Hussain, M. M., Abdussami, M., Devi, A. A., & Faizi, J. (2025). Deploying TinyML for energy-efficient object detection and communication in low-power edge AI systems. *Scientific Reports*, *15*(1), 44299. https://doi.org/10.1038/s41598-025-27818-9

Chen, Y., Zhang, H., & Zhong, S. (2024). Design and implementation of smart home system based on IoT. *Results in Engineering*, *24*, 103410. https://doi.org/https://doi.org/10.1016/j.rineng.2024.103410

Cho, J., Heo, Y., & Moon, J. W. (2023). An intelligent HVAC control strategy for supplying comfortable and energy-efficient school environment. *Advanced Engineering Informatics*, *55*, 101895. https://doi.org/https://doi.org/10.1016/j.aei.2023.101895

David, R., Duke, J., Jain, A., Reddi, V. J., Jeffries, N., Li, J., Kreeger, N., Nappier, I., Natraj, M., Regev, S., Rhodes, R., Wang, T., & Warden, P. (2020). TensorFlow Lite Micro: Embedded Machine Learning on TinyML Systems. *CoRR*, *abs/2010.08678*. https://arxiv.org/abs/2010.08678

Elhami, M., Goodarzi, S. S., Maleki, S., & Sajadi, B. (2025). Three-objective optimization of the HVAC system control strategy in an educational building to reduce energy consumption and enhance indoor environmental quality (IEQ) using machine learning techniques. *Journal of Building Engineering*, *105*, 112444. https://doi.org/https://doi.org/10.1016/j.jobe.2025.112444

Jain, S. K., & Kesswani, N. (2023). A noise-based privacy preserving model for Internet of Things. *Complex & Intelligent Systems*, *9*(4), 3655–3679. https://doi.org/10.1007/s40747-021-00489-5

Márquez-Sánchez, S., Alonso-Rollán, S., Nahom, H., Erbad, A., & Fernandez, J. H. (2025). Optimizing Building Energy Management Leveraging Adaptive Edge Computing for Enhanced Efficiency and Occupant Well-Being. In P. Novais, P. B. D., I. Satoh, V. J. Inglada, S. R. González, E. Jove Pérez, J. Parra Domínguez, P. Chamoso, & R. S. Alonso (Eds.), *Ambient Intelligence – Software and Applications – 15th International Symposium on Ambient Intelligence* (pp. 236–248). Springer Nature Switzerland.

Márquez-Sánchez, S., Calvo-Gallego, J., Erbad, A., Ibrar, M., Hernandez Fernandez, J., Houchati, M., & Corchado, J. M. (2023). Enhancing Building Energy Management: Adaptive Edge Computing for Optimized Efficiency and Inhabitant Comfort. *Electronics*, *12*(19). https://doi.org/10.3390/electronics12194179

Meedeniya, D., Ariyarathne, I., Bandara, M., Jayasundara, R., & Perera, C. (2023). A Survey on Deep Learning Based Forest Environment Sound Classification at the Edge. *ACM Comput. Surv.*, *56*(3). https://doi.org/10.1145/3618104

Morawska, L., Tang, J. W., Bahnfleth, W., Bluyssen, P. M., Boerstra, A., Buonanno, G., Cao, J., Dancer, S., Floto, A., Franchimon, F., Haworth, C., Hogeling, J., Isaxon, C., Jimenez, J. L., Kurnitski, J., Li, Y., Loomans, M., Marks, G., Marr, L. C., … Yao, M. (2020). How can airborne transmission of COVID-19 indoors be minimised? *Environment International*, *142*, 105832. https://doi.org/https://doi.org/10.1016/j.envint.2020.105832

Park, H., Hong, S., & Kim, J.-S. (2024). Audio Classification on Low-Resource Microcontrollers. *2024 15th International Conference on Information and Communication Technology Convergence (ICTC)*, 1332–1333. https://doi.org/10.1109/ICTC62082.2024.10826933

*name of corresponding author

This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Pham, H. H., Le, T. M., & Son, L. H. (2026). Towards open world sound event detection. *Signal Processing*, *248*, 110707. (Early Access) https://doi.org/https://doi.org/10.1016/j.sigpro.2026.110707

Rawat, N., Kumar, P., Hama, S., Williams, N., & Zivelonghi, A. (2025). Improving classroom air quality and ventilation with IoT-driven acoustic and visual CO2 feedback system. *Science of The Total Environment*, *980*, 179543. (Early Access) https://doi.org/https://doi.org/10.1016/j.scitotenv.2025.179543

*name of corresponding author

This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.