A Multimodal Dataset for Automatic Edge-AI Cough Detection

Lara Orlandic,[1] Jérôme Thevenot,[1] Tomas Teijeiro[2], and David Atienza[1]

Abstract— **Counting the number of times a patient coughs per day is an essential biomarker in determining treatment efficacy for novel antitussive therapies and personalizing patient care. Automatic cough counting tools must provide accurate information, while running on a lightweight, portable device that protects the patient's privacy. Several devices and algorithms have been developed for cough counting, but many use only error-prone audio signals, rely on offline processing that compromises data privacy, or utilize processing and memory-intensive neural networks that require more hardware resources than can fit on a wearable device. Therefore, there is a need for wearable devices that employ multimodal sensors to perform accurate, privacy-preserving, automatic cough counting algorithms directly on the device in an edge Artificial Intelligence (edge-AI) fashion. To advance this research field, we contribute the first publicly accessible cough counting dataset of multimodal biosignals. The database contains nearly 4 hours of biosignal data, with both acoustic and kinematic modalities, covering 4,300 annotated cough events from 15 subjects. Furthermore, a variety of non-cough sounds and motion scenarios mimicking daily life activities are also present, which the research community can use to accelerate machine learning (ML) algorithm development. A technical validation of the dataset reveals that it represents a wide variety of signal-to-noise ratios, which can be expected in a real-life use case, as well as consistency across experimental trials. Finally, to demonstrate the usability of the dataset, we train a simple cough vs non-cough signal classifier that obtains a 91% sensitivity, 92% specificity, and 80% precision on unseen test subject data. Such edge-friendly AI algorithms have the potential to provide continuous ambulatory monitoring of the numerous chronic cough patients.**

I. INTRODUCTION

Chronic cough is a common condition that globally affects 7.9% of the general adult population [1]. It can significantly detriment individuals' quality of life by impairing sleep quality, contributing to anxiety, and increasing medical expenditures [2]. The number of times people cough per day is an indicator of the severity of conditions with symptomatic cough, and it is the primary effectiveness metric for many antitussive therapy trials [3]. However, objective cough counts remain difficult to measure due to the significant time and effort required to manually detect and annotate coughs [2], and subjective cough questionnaires show only a moderate correlation to actual cough counts [4]. Given the global burden of the condition, there is a need for objective, widespread cough quantification tools to monitor patients' symptoms and assess treatment efficacy [3].

One of the most widely adopted automatic cough counting tools in clinical studies is the Leicester Cough Monitor [3], [5]. The device consists of a microphone-enhanced necklace that records audio data for up to 24 hours, saves it to a central hardware unit, and then the data is downloaded and processed by an offline computer-based algorithm [5]. This state-of-the-art tool has many issues, including cough precision, data privacy, and portability. This system, as well as newer cough detection systems based solely on audio signals [6], [7], may exhibit false positives due to bystanters' coughs. Furthermore, processing biosignal data offline poses significant privacy risks, especially when communicating sensitive audio data to computers or servers. Finally, hardware capable of recording 24 hours of high-bandwidth audio data requires significant computational resources, thus resulting in a bulky and conspicuous design.

State-of-the-art cough detection models improve precision by leveraging sensor fusion, and many of these use tri-axial accelerometers [8], [9], [10]. In particular, Fan et al. observed that adding accelerometer signals to their audio-based cough detection model reduced the false positive rate by a factor of eight [8]. Furthermore, Drugman et al. analyzed the efficacy of different sensors for cough counting and observed that while audio signals produced the highest accuracy, accelerometers helped distinguish cough from speech events [10]. Otoshi et al. combined a chest-mounted accelerometer with a neck-worn strain sensor to obtain a 92% sensitivity and 96% specificity in cough detection [9].

To address the issues of patient data privacy and portability, edge Artificial Intelligence (edge-AI) is a novel technological paradigm that has the potential to satisfy all of these design constraints. Instead of streaming raw physiological data through power-hungry wireless links, edge-AI systems perform heavy data processing directly on the wearable device [11]. This can conserve the battery life and communication latency [12], while also preserving privacy by keeping most data localized to the device [11].

To our knowledge, no cough detection models have been optimized specifically for edge-AI implementation, taking into account the energy and memory constraints of edge devices. While multimodal cough counting devices and algorithms have shown promising preliminary results, most rely on artificial neural networks [8], [9], which impose a significant memory and processing overhead. For example, to perform cough detection from audio signals, a state-of-the-art convolutional neural network (CNN) includes more than 17,000 parameters and 10 million operations per classification [13]. This network exceeds the memory capacity of most processors running on wearable devices, which usually

[1]L. Orlandic, J. Thevenot, and D. Atienza are with the Embedded Systems Laboratory (ESL) of EPFL, Lausanne, Switzerland. Corresponding author: lara.orlandic@epfl.ch

[2]T. Teijeiro is with BCAM - Basque Center for Applied Mathematics, Bilbao, Spain

TABLE I: Recorded biosignals

Sensing Modality	Recorded Signals	Sampling Rate
Acoustic	Body-facing microphone, Outward-facing microphone	16 kHz
Kinematic	Accelerometer X, Y, Z, Gyroscope Yaw, Pitch, Roll	0.1 kHz

Fig. 1: Placement and orientation of each sensor embedded in the wearable device.

contain hundreds of KiB of Flash and RAM [14]. In order for cough counting wearable devices to achieve clinical-grade accuracy, preserve patient privacy, and ensure wearability, the constraints of embedded sensors and edge computing need to be integrated in the definition of suitable wearable architectures and corresponding algorithms.

The first step of developing multimodal edge-AI cough counting algorithms is access to an extensive, finely labeled dataset. Unfortunately, none of the datasets collected in previous cough detection studies are publicly available [5], [8], [9], [10], [13]. In this work, we contribute such a dataset to the research community, complete with an open-sourced code repository to aid in data pre-processing and data segmentation for rapid model prototyping [15], [16]. We designed an experiment in which subjects performed cough and non-cough events in everyday, noisy environments while wearing a device incorporating audio and kinematic signals. Finally, we demonstrate how this dataset can be used to develop multimodal edge-AI cough classifiers.

II. MATERIAL AND METHODS

A. Acquisition system

A novel lightweight, battery-powered, chest-mounted wearable device was designed to collect the biosignals included in this dataset. The sensor data was acquired and synchronized using a STM32G4 series MCU (32-bit ARM Cortex-M4, 170 MHz), and subsequently saved to an on-board Flash memory (GigaDevices GD5F1GQ4-family SPI NAND Flash). This hardware was placed inside a custom frame made of biocompatible material (polylactic acid/thermoplastic polyurethane) and coated in certified skin-safe silicone (Dragon Skin™).

An overview of the biosignal information recorded by the device is presented in Table I. In addition to the accelerometers used in previous studies [8], [9], [10], we employ a 9-axis inertial measurement unit (IMU) that contains an accelerometer, gyroscope, and magnetometer. The audio and IMU signals were selected due to their proven effectiveness in cough classification compared to other sensors [10]. The acoustic information of the cough is recorded by two digital microphones (SPH0645LM4H-B, Knowles, Itasca, USA), each sampled at 16 kHz with 24-bit PCM encoding. One microphone faced toward the subject's body, and the other one faced away from the body, as shown in Fig. 1. The goal of this setup was to evaluate the effects of microphone orientation on signal quality; the body-facing microphone captures less background noise than the outward-facing microphone, but may contain more noise from the device rubbing against the body. This hypothesis is investigated in Section III-B.
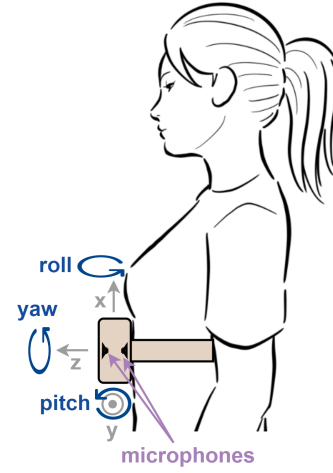
To monitor the chest vibrations due to coughing and other activities, one 9-axis IMU (BNO080, Hillcrest Labs, Rockville, USA) is used. The on-board tri-axial magnetometer provides the absolute orientation of the module, which is fused with the other IMU sensors to assess the chest acceleration in the x, y, and z directions, as well as rotation along these axes as the gyroscope yaw, pitch, and roll. The orientation of these signals when the device was placed on the subject's chest is illustrated in Fig. 1. The IMU signals were each sampled at 100 Hz and encoded with 16 bits. In all experiments, the device was placed at the midline of the chest, just below the subject's pectorals. Three different chest locations were investigated in a pilot study, and this one produced the audio signals with the highest SNR.

All of the signals recorded throughout the experiment are shown in Fig. 2, which depicts multiple cough episodes. For visualization purposes, each signal is normalized to its maximum value, and the accelerometer z signal is negated such that a chest acceleration into the body (i.e. in the negative z direction) appears as a peak in the signal. No additional preprocessing was performed on the signals. We can observe strong accelerations in the z-direction during coughing, as well as roll and yaw rotations potentially due to the subject's reflex of rotating their body to cough into their elbow.

B. Experimental setup

In order to assess the accuracy of automatic cough quantification algorithms in everyday environments, an experiment was designed to elicit typical sounds in various noisy conditions. Similarly to the experiment performed in [10], we aimed to produce cough and non-cough sounds in the presence of both audio and kinematic noise conditions to develop robust multimodal cough detection algorithms. Recordings were collected from 20 healthy subjects (10 male, 10 female; age 26.5 ± 6.5 years; body mass index (BMI) 22.6 ± 4.5 kilograms per square meter). Institutional review board approval was obtained (HREC No.: 085-2022)
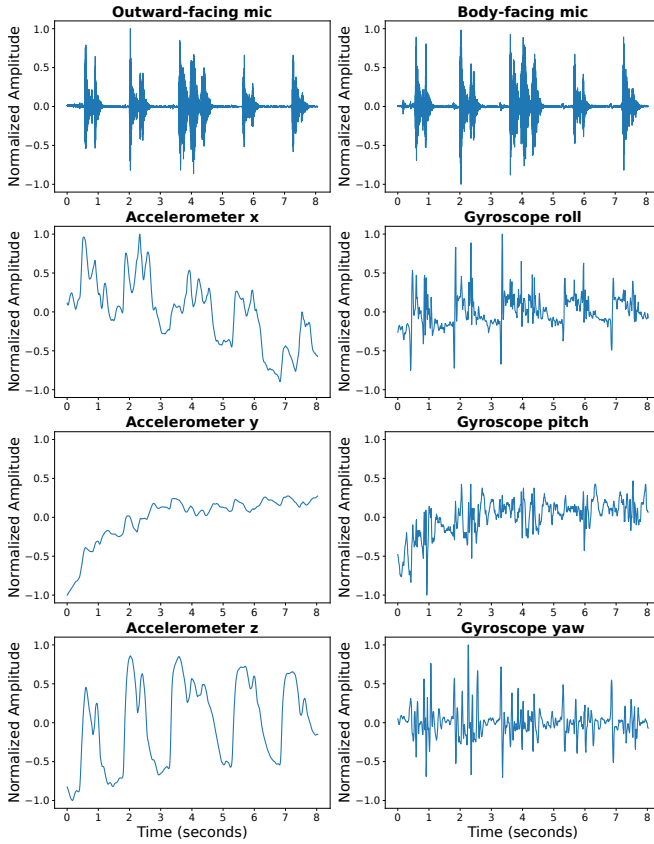
Fig. 2: All of the signals measured by the device during 11 cough episodes over 8 seconds.

TABLE II: Experimental conditions

Sound	Acoustic noise	Kinematic noise
Cough, Laugh, Throat clearing, Deep breathing	None, Traffic, Music, Bystander cough	Sitting, Walking

and all participants signed an informed consent prior to data acquisition.

The experiment consisted of every possible combination of the sounds and noise scenarios shown in Table II, thus producing 32 recordings per session. Each subject performed either two or three trials on separate days to assess repeatability. The subjects wore their everyday clothing throughout the trials, and they were given the choice whether to wear shoes or not. This induced real-life variability due to different individuals' clothing choices. The device was placed below the shirt and tightened to each subject's comfort level. The recordings each lasted approximately 10 seconds, during which the subject was asked to produce either a cough sound or one of three sounds that could produce similar audio or chest motion artefacts as a cough: laughing, throat clearing, and deep breathing.

Kinematic noise was introduced by the subject performing tasks while walking, which produces motion artefacts in the IMU signals, as well as audio noise of footsteps in some cases. Acoustic noise was introduced by the experimenter

playing loud music or traffic noises during the recording. The music consisted of one of four songs, and the traffic was a minutes-long recording of city traffic sounds. These recordings were started at random times across the experiments to induce variability in the noise amplitude and content. A useful final noise scenario is that of bystanders' coughing, in which the subject performed a sound in the presence of one other subject's recorded coughs. By having the bystander be another subject in the training dataset, we prevent the model from learning to exclude only one person's specific cough sounds. Such a condition can test the ability of cough detection algorithms to count only the coughs produced by the subject. All of the experiments were performed in the same room over the course of two months.

C. Data annotation

In this work, we performed an extensive, fine-grained annotation of the cough recordings that marks the start and endpoints of each individual cough sound. This was done in a semi-automated manner by a trained observer. The first step of the annotation was extracting relevant fiducial points of the outward-facing microphone and accelerometer z-direction signals, such as peaks and valleys, and then manually selecting the beginning and end of each cough. The fiducial points extracted from the audio signal were the starts, ends, and peaks of each cough burst, which were computed using a modified version of the COUGHVID segmentation algorithm [17]. This algorithm performs hysteresis thresholding on the signal power to delineate areas of rapid signal amplitude increase and decrease.

The audio thresholding alone cannot accurately delineate coughs, as it may pick up high-amplitude noises as well as bystander coughs. This is why we also analyzed the peaks and valleys of the accelerometer z signal to determine whether each sound burst was accompanied by a chest acceleration. Peaks and valleys were defined as local minima and maxima of the second derivative of the signal, respectively. Once all of the fiducial points were extracted, the starts and stops of each cough were manually selected among the extracted points based on which points corresponded most closely to each observed cough sound. As an illustrative example, let us consider the recording in Fig. 3, which shows two bystander coughs followed by four of the subject's coughs, then one more bystander cough.

We can see from the fiducial points that the audio signal thresholding mistakenly identifies the bystander coughs, while the IMU signal exhibits a peak at every true cough but does not adequately mark the onset of the cough. In this example recording, the final annotation is selected as the regions between audio burst starts and IMU valleys containing an IMU peak in-between. However, the fiducial points selected to mark the start and end of each cough varied from one subject to the next and required significant manual effort to properly annotate. Therefore, a more sophisticated cough detection methodology employing ML is necessary to accurately determine the number of coughs in a given set of biosignals.
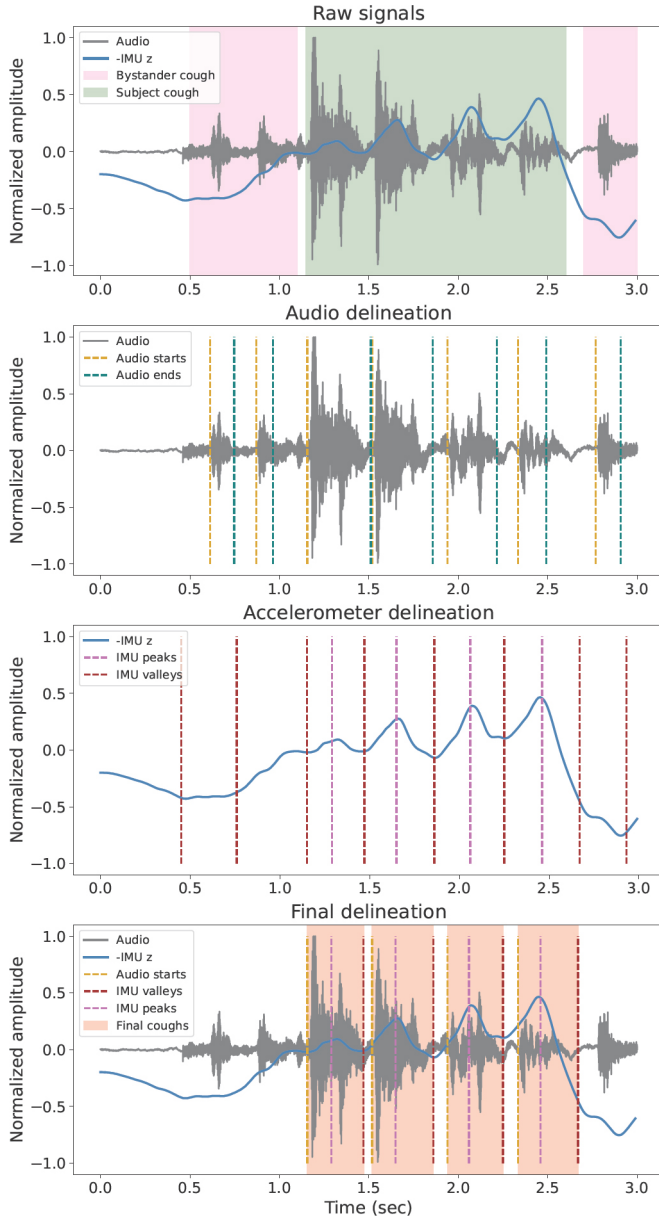
Fig. 3: Example cough recording annotation procedure.

Fig. 4: Directory structure of the data files for each subject.

All of the code used to extract the fiducial points has been made available at our public Git repository [16]. Overall, the public dataset contains 227 minutes of biosignals and nearly 4,300 annotated coughs.

III. DATA RECORDS

A. Dataset usage

The data of 15 of the 20 subjects has been published to our Zenodo repository [15]. The remaining five subjects are reserved for a private testing set, as described in Section III-C. The files are arranged in a hierarchical structure as shown in Fig. 4. For each experimental condition in Table II, there are three to four corresponding files: .wav audio files for the body-facing and outward-facing microphones, a .csv file for the IMU data, and in the case of cough recordings, a .json file
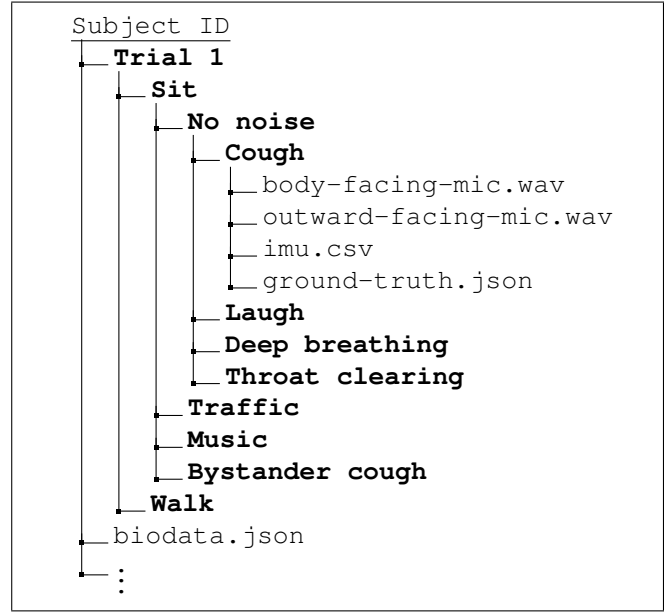
containing the cough location annotations. The raw audio is saved in 32-bit signed integer format, which corresponds to the sampled 24-bit PCM format shifted 8 bits to the left. The annotations file should be loaded as a dictionary object, with the cough start times being under the key start_times and cough end times under end_times. The gender and BMI of each subject are recorded in the biodata.json file.

In order to facilitate Machine Learning (ML) model development, several functions have been provided on our public Git repository [16] to transform the raw data into segmented model inputs. The functions read the biosignal files in each experimental segment and generate short cough and non-cough signal segments with which to train models. This segmentation procedure enables edge-AI model development, as such algorithms must perform inference on short data buffers to economize the limited memory resources of the device.

The positive samples are generated by locating each ground-truth cough sound in the cough recording and selecting a window of data around it. The window length is a hyperparameter of the model training and can be selected by the user of the dataset. As a data augmentation procedure, the window is randomly shifted around the cough to account for coughs at different locations within each window. In the case that the window length is less than the cough length, then the window is randomly shifted within the cough duration. For the non-cough recordings, a user-defined number of randomly-selected segments, equal in length to the cough segments, is extracted. A function is provided to automatically generate the data segments for each subject, taking an equal number of segments of each produced sound.

B. Technical validation

To ensure that this dataset provides biosignals in a variety of everyday environments, we analyze the signal-to-noise ratio (SNR) of the cough audio recordings of different phases

of the experiment. In doing so, we analyze the amount of audio noise present across the experimental segments to ensure that they represent a wide range of real-life scenarios. Moreover, we assess the differences in audio quality between the body-facing and outward-facing microphones.
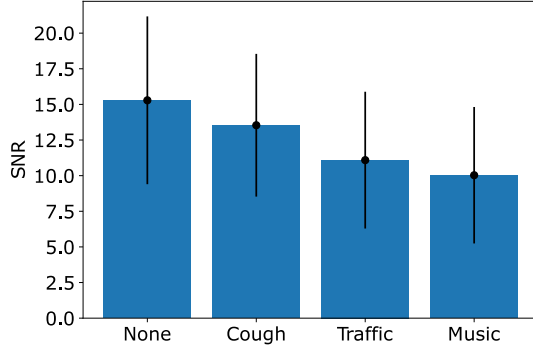

Fig. 5: A comparison of the outer-facing microphone SNR across different audio noise conditions of the experiment.

Fig. 5 depicts the SNR across different phases of the experiment for all recordings in the public dataset. The SNR was computed by comparing the signal power, in units of dB, of the cough indices of a given recording to that of the rest of the signal. We can see from the figure that the various noise scenarios successfully induce audio noise by significantly reducing the SNR, especially in the case of the traffic and music conditions. Furthermore, the large standard deviation in each condition indicates that a variety of different noise levels are present within the database, thus enabling models to perform generalization.

Next, the SNR of all cough recordings in each trial was computed to assess the repeatability of the experiment and standardization of the experimental conditions across trials. This analysis revealed no significant difference in SNR across trials. An unpaired, two-tailed t-test of the SNR values in each trial produced $p = 0.2$ between trials 1 and 2, $p = 0.65$ between trials 1 and 3, and $p = 0.41$ between trials 2 and 3. Finally, we analyzed the SNR difference between all recordings of the body-facing and outward-facing microphones. A paired, two-tailed t-test revealed that the SNR of the outward-facing microphone was significantly higher than that of the body-facing microphone ($p = 3 \times 10^{-16}$), with the difference being more pronounced during walking ($p = 2.6 \times 10^{-10}$) than sitting ($p = 1.9 \times 10^{-7}$). This confirms the hypothesis that the body-facing microphone contains more noise than the outward-facing one due to friction against the skin.

C. Testing procedure

In order to evaluate the success of ML algorithms for cough detection, 5 subjects in the dataset have been withheld from publication. The cough samples in these recordings have been annotated in the same way as those of the public

TABLE III: Cough classifier features

Signals	Feature Count	Feature Type	Feature List
Inner/ Outer Facing Mic	65	Freq. Domain	MFCC, Power Spectral Denisity, Dominant Frequency, Spectral Centroid / Rolloff / Spread / Skewness / Kurtosis / Decrease / Slope / Flatness / Standard Deviation
		Time Domain	Energy Envelope Peak Detection, Zero-Crossing Rate, Crest Factor, RMS Power
Accel x,y,z, Accel norm, Gyro y, p, r, Gyro norm	5	Time Domain	Line Length, Zero-Crossing Rate, Kurtosis, Crest Factor, RMS Power

training data. These 5 subjects were randomly selected from the dataset in a gender-stratified fashion, such that 3 males and 2 females were in the test set to keep the gender balance in both datasets as equal as possible. The evaluation of models on the private test set is available to the entire scientific community. Any researcher that demonstrates promising results on the public dataset using cross-validation may apply for an independent evaluation by contacting the research team of this work.

IV. EDGE-AI ALGORITHM DEVELOPMENT

To demonstrate the utility of our dataset in developing multimodal, edge-AI cough detection algorithms, we have trained several cough-vs-non-cough segment classification ML models based on state-of-the-art audio and IMU signal features. The dataset was prepared by segmenting the cough and non-cough signals into windows of 0.4 seconds using the code described in Section III-A. Such a short window length economizes the memory of resource-constrained wearable devices on which such an edge-AI algorithm would run. The dataset was augmented by randomly shifting the coughs within the windows twice. No filtering was performed on the signals, but the mean of each IMU signal segment was subtracted to center it at zero.

In order to extract useful information from the biosignals, the features listed in Table III were extracted from each segment. For each microphone signal, 65 audio features used in previous audio-based cough detection methods were computed [17]. For the IMU signals, five time-domain features were employed, most of which were previously used on cough detection algorithms [18], as well as the line length feature used in seizure classification [19]. These features were extracted on all individual IMU signals, as well as the L2 vector norms of the three accelerometer and gyroscope signals, resulting in 40 total IMU features. The majority of these features can be executed on wearable devices in a battery-preserving manner using a biomedical edge-AI hardware accelerator [20].

A model selection and optimization technique similar to that of [21] was applied, performing cross-validation (CV)

at every step of the procedure to assess the generalization capabilities of the model and investigate the variations in performance across folds. Once the features of each segment were computed, a 5-fold Leave-n-Subjects-Out CV procedure was performed to evaluate the performance of six classification algorithms: Logistic Regression, Gaussian Naive Bayes, Linear Discriminant Analysis, Decision Tree, Random Forest, and eXtreme Gradient Boosting. At each CV fold, standard normal scaling was performed on the features, and the SMOTE technique was applied to account for the class imbalance between cough and non-cough samples [22]. The ML model with the highest average Area Under the ROC Curve (AUC) across the cross-validation folds was selected. Next, Recursive Feature Elimination with CV (RFECV) was performed to remove features that did not contribute to the classification outcome. Finally, the selected model with the optimal feature list underwent hyperparameter optimization and its final CV AUC was reported.

To assess the efficacy of using multimodal cough detection signals, three different models are trained and evaluated in terms of CV AUC: One that uses only the outer microphone audio signal features, one using only IMU signal features, and one combining both microphone and IMU features. In each model, the eXtreme Gradient Boosting classifier performed best. The average and standard deviation AUC scores across the five CV folds are depicted in Fig. 6. We can see that the combined model performs the best, with a CV AUC of 0.96 ± 0.01. The model trained with only outer mic features exhibits a CV AUC of 0.92 ± 0.01, and the model trained with only IMU features has the lowest CV AUC of 0.90 ± 0.02.
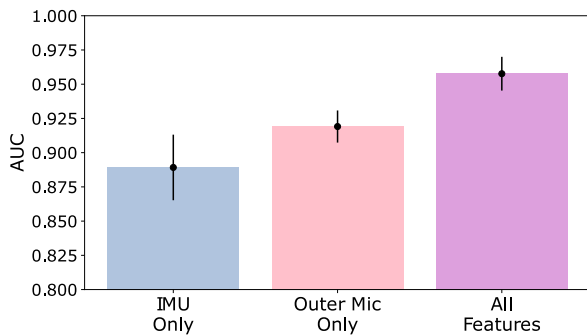

Fig. 6: Comparison of the cross-validation AUC values and standard deviations of the three trained models

Finally, the most successful model was tested on the private test set and achieved a final AUC score of 0.97. The ROC curve of the classifier is shown in Fig. 7. The maximal average F-1 score of this ROC curve is 0.914, corresponding to a sensitivity of 91%, a specificity of 92%, and a precision of 80%. A SHAP analysis [23] of the classifier revealed that the top 5 features determining the model outcome were the gyroscope yaw line length, body-facing microphone MFCC standard deviation of the first component, gyroscope roll line length, gyroscope yaw RMS,

and body-facing microphone spectral kurtosis. This analysis shows that the model uses information from both types of sensors to enhance its classification outcome.
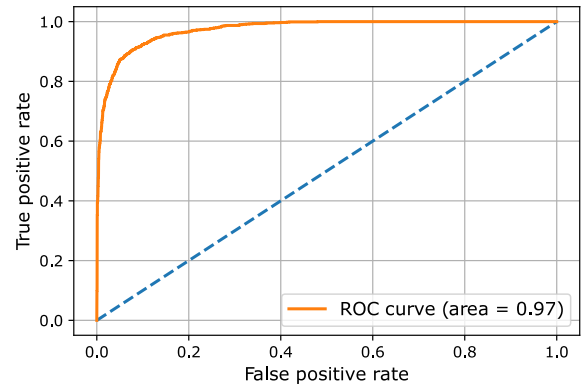

Fig. 7: ROC curve of the classifier using multimodal features.

V. DISCUSSION AND CONCLUSIONS

In this work, we have presented the first publicly available, multimodal cough counting dataset to assist the research community in developing, testing, and deploying cough detection ML models. We provide fine-grained annotations of the start and stop of each individual cough, which enable the development of both cough detection and segmentation algorithms. An experiment was conducted in line with state-of-the-art cough detection algorithm studies [10], providing training and testing data for coughs that occur in the presence of audio and kinematic noise scenarios. An analysis of the SNR of the audio signals confirms the presence of varying noise levels, thereby mimicking a real-world environment. Furthermore, our data segmentation code is open-sourced to streamline edge-AI cough detection model development, encourage seamless merging with existing datasets, or to perform semi-supervised learning on unlabeled audio and IMU data. The data can be downloaded at our Zenodo repository and processed using the code on our Git repository [15], [16].

While nearly 4 hours and 4,300 cough sounds are provided to the public, five testing subjects' biosignals have been withheld from publication and serve as a benchmark testing dataset. These samples may provide a fair comparison of different algorithms that count coughs based on audio and IMU biosignals. However, users of the dataset should note the limited age range of the subjects in this study and may choose to include more data from increasingly diverse participants.

Finally, we demonstrated the utility of this dataset by providing a sample ML model development pipeline. We trained three different classifiers using state-of-the-art features from combinations of sensors to justify the need for sensor fusion in noisy scenarios. The most successful model combined information from both the audio and IMU sensors to achieve final sensitivity and specificity values of 91% and

92%, respectively. Although the model trained with only IMU features performed the worst, the fact that its mean AUC was only 2% below that of the outer microphone model is striking given that the sampling rate of the IMU is 160x less than that of the microphone. Such information can allow edge-AI algorithm designers to perform energy-saving optimizations, such as only sampling the microphone signal when necessary.

Although the classifier performs well in distinguishing cough segments from non-cough segments, these results are only preliminary. Further testing must be done to ensure that in a continuous recording scenario, every cough is detected and can be counted separately from neighboring coughs. Furthermore, a more detailed analysis of the classifier window length must be performed to analyze the trade-offs between cough counting accuracy and memory consumption on a wearable edge-AI platform.

ACKNOWLEDGMENT

REFERENCES

[1] W.-J. Song, Y.-S. Chang, S. Faruqi, J.-Y. Kim, M.-G. Kang, S. Kim, E.-J. Jo, M.-H. Kim, J. Plevkova, H.-W. Park, S.-H. Cho, and A. H. Morice, "The global epidemiology of chronic cough in adults: a systematic review and meta-analysis," *European Respiratory Journal*, vol. 45, no. 5, pp. 1479–1481, May 2015.

[2] H.-K. Won and W.-J. Song, "Impact and disease burden of chronic cough," *Asia Pacific Allergy*, vol. 11, no. 2, 2021.

[3] J. I. Hall, M. Lozano, L. Estrada-Petrocelli, S. Birring, and R. Turner, "The present and future of cough counting tools," *Journal of Thoracic Disease*, vol. 12, no. 9, pp. 5207–5223, 2020.

[4] S. S. Birring, S. Matos, R. B. Patel, B. Prudon, D. H. Evans, and I. D. Pavord, "Cough frequency, cough sensitivity and health status in patients with chronic cough," *Respiratory Medicine*, vol. 100, no. 6, pp. 1105–1109, 2006.

[5] S. S. Birring, T. Fleming, S. Matos, A. A. Raj, D. H. Evans, and I. D. Pavord, "The Leicester Cough Monitor: preliminary validation of an automated cough detection system in chronic cough," *The European Respiratory Journal*, vol. 31, no. 5, pp. 1013–1018, 2008.

[6] S. Larson, G. Comina, R. H. Gilman, B. H. Tracey, M. Bravard, and J. W. López, "Validation of an Automated Cough Detection Algorithm for Tracking Recovery of Pulmonary Tuberculosis Patients," *PLOS ONE*, vol. 7, no. 10, 2012.

[7] S. Matos, S. S. Birring, I. D. Pavord, and D. H. Evans, "Detection of cough signals in continuous audio recordings using hidden Markov models," *IEEE transactions on bio-medical engineering*, vol. 53, no. 6, pp. 1078–1083, 2006.

[8] J. Fan, G. Comina, R. Gilman, J. López, and B. Tracey, "Cough monitoring for pulmonary tuberculosis using combined microphone/accelerometer measurements," *The Journal of the Acoustical Society of America*, vol. 135, p. 2268, Apr. 2014.

[9] T. Otoshi, T. Nagano, S. Izumi, D. Hazama, N. Katsurada, M. Yamamoto, M. Tachihara, K. Kobayashi, and Y. Nishimura, "A novel automatic cough frequency monitoring system combining a triaxial accelerometer and a stretchable strain sensor," *Scientific Reports*, vol. 11, 2021.

[10] T. Drugman, J. Urbain, N. Bauwens, R. Chessini, C. Valderrama, P. Lebecque, and T. Dutoit, "Objective Study of Sensor Relevance for Automatic Cough Detection," *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 3, pp. 699–707, May 2013.

[11] Y.-L. Lee, P.-K. Tsung, and M. Wu, "Technology trend of edge AI," in *2018 International Symposium on VLSI Design, Automation and Test (VLSI-DAT)*, 2018, pp. 1–2.

[12] R. Braojos, H. Mamaghanian, A. Dias, G. Ansaloni, D. Atienza, F. J. Rincón, and S. Murali, "Ultra-low power design of wearable cardiac monitoring systems," in *2014 51st ACM/EDAC/IEEE Design Automation Conference (DAC)*, Jun. 2014, pp. 1–6.

[13] F. Barata, K. Kipfer, M. Weber, P. Tinschert, E. Fleisch, and T. Kowatsch, "Towards Device-Agnostic Mobile Cough Detection with Convolutional Neural Networks," in *2019 IEEE International Conference on Healthcare Informatics (ICHI)*, 2019, pp. 1–11.

[14] "STM32G4 Series of mixed-signal MCUs with DSP and FPU instructions - STMicroelectronics." [Online]. Available: https://www.st.com/en/microcontrollers-microprocessors/stm32g4-series.html

[15] L. Orlandic, J. Thevenot, T. Teijeiro, and D. Atienza, "A Multimodal Dataset for Automatic Edge-AI Cough Detection," 2023, type: dataset. [Online]. Available: https://zenodo.org/record/7562332

[16] "Edge Artificial Intelligence (edge-AI) Cough Counting," 2023. [Online]. Available: https://github.com/esl-epfl/edge-ai-cough-count

[17] L. Orlandic, T. Teijeiro, and D. Atienza, "The COUGHVID crowd-sourcing dataset, a corpus for the study of large-scale cough analysis algorithms," *Scientific Data*, vol. 8, no. 1, p. 156, Jun. 2021, number: 1 Publisher: Nature Publishing Group.

[18] M. Pahar, I. Miranda, A. Diacon, and T. Niesler, "Deep Neural Network Based Cough Detection Using Bed-Mounted Accelerometer Measurements," in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Jun. 2021, pp. 8002–8006.

[19] N. Koolen, K. Jansen, J. Vervisch, V. Matic, M. De Vos, G. Naulaers, and S. Van Huffel, "Line length as a robust method to detect high-activity events: Automated burst detection in premature EEG recordings," *Clinical Neurophysiology*, vol. 125, no. 10, pp. 1985–1994, 2014.

[20] B. W. Denkinger, M. Peón-Quirós, M. Konijnenburg, D. Atienza, and F. Catthoor, "VWR2A: a very-wide-register reconfigurable-array architecture for low-power embedded devices," in *Proceedings of the 59th ACM/IEEE Design Automation Conference*, ser. DAC '22. New York, NY, USA: Association for Computing Machinery, 2022, pp. 895–900.

[21] L. Orlandic, A. A. Valdes, and D. Atienza, "Wearable and Continuous Prediction of Passage of Time Perception for Monitoring Mental Health," in *2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS)*, 2021, pp. 444–449.

[22] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, Jun. 2002.

[23] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," in *Advances in Neural Information Processing Systems*, vol. 30. Curran Associates, Inc., 2017.