

Pulneo: Prototipo móvil gamificado con detección acústica de actividad respiratoria para fisioterapia domiciliaria

Pulneo: Gamified Mobile Prototype with Acoustic Respiratory Activity Detection for Home-Based Physiotherapy

María Gabriela Viñansaca Cabrera¹ <https://orcid.org/0009-0007-9490-4180>, David Fernando Valladarez Muñoz² <https://orcid.org/0009-0002-3160-6243>

¹Universidad Politécnica Salesiana, Cuenca, Ecuador
mvinansacac@ups.edu.ec

²Corporación Ecuatoriana para el Desarrollo de la Investigación y la Academia (CEDIA), Cuenca, Ecuador
david.valladarez@cedia.org.ec



Esta obra está bajo una licencia internacional
Creative Commons Atribución-NoComercial 4.0

Enviado: 2026/07/12

Aceptado: 2026/08/27

Publicado: 2026/10/14

Resumen

La fisioterapia respiratoria domiciliaria requiere repetición y retroalimentación, pero muchas sesiones se realizan sin supervisión directa. Este artículo presenta Pulneo, un prototipo móvil gamificado que detecta localmente la actividad respiratoria mediante señales acústicas y la traduce en biorretroalimentación visual durante ejercicios guiados. El modelo se entrenó con 69 634 ventanas balanceadas: 34 817 de actividad respiratoria de 1061 sujetos sanos de Coswara y 34 817 negativas de 5 608 archivos de UrbanSound8K, con una partición 70/15/15 por grupos, sin solapamiento de sujetos ni de archivos. En 10 347 ventanas de prueba, el modelo TensorFlow Lite INT8 alcanzó 97,64 % de *accuracy*, 98,09 % de precisión, 97,18 % de *recall*, F1 de 97,63 % y ROC-AUC de 99,65 %. La cuantización redujo el modelo de 88 488 a 31 600 bytes (64,29 %) y mantuvo 99,55 % de concordancia con TFLite FP32. La aplicación interpreta la salida como P(respiración) y aplica un umbral calibrado en validación ($P \geq 0,29$). Pulneo no distingue inspiración de exhalación ni ha sido validado clínicamente; los resultados respaldan su funcionamiento técnico, no su eficacia terapéutica. Un autor fisioterapeuta con maestría en terapia respiratoria realizó una revisión conceptual interna para delimitar el uso previsto, la interpretación de la biorretroalimentación y los requisitos de seguridad para futuras evaluaciones.

Sumario: Introducción, Materiales y Métodos, Resultados, Discusión, Alcance y limitaciones, Conclusiones.

Como citar: Viñansaca, M. & Valladarez, D. (2026). Pulneo: Prototipo móvil gamificado con detección acústica de actividad respiratoria para fisioterapia domiciliaria. *Revista Tecnológica - Espol*, 38(E1), 136-152.
<https://doi.org/10.37815/rte.v38nE1.1602>

Palabras clave: aprendizaje automático, biorretroalimentación, computación en el borde, cuantización, inferencia local, redes neuronales convolucionales.

Abstract

Home-based respiratory physiotherapy requires repetition and feedback, yet many sessions are performed without direct supervision. This paper presents Pulneo, a gamified mobile prototype that detects respiratory activity on-device from acoustic signals and maps it to visual biofeedback during guided exercises. The model was trained on 69,634 balanced windows: 34,817 respiratory-activity windows from 1,061 healthy Coswara subjects and 34,817 negative windows from 5,608 UrbanSound8K files, using a group-wise 70/15/15 split with no subject or source-file overlap. On 10,347 test windows, the TensorFlow Lite INT8 model achieved 97.64% accuracy, 98.09% precision, 97.18% recall, 97.63% F1-score, and 99.65% ROC-AUC. Quantization reduced the model from 88,488 to 31,600 bytes (64.29%) while preserving 99.55% decision agreement with TFLite FP32. The application interprets the output as P(breathing) and applies a threshold calibrated on validation data ($P \geq 0.29$). Pulneo does not distinguish inhalation from exhalation and has not been clinically validated; the findings support technical performance, not therapeutic efficacy. An author who is a physiotherapist with a master's degree in respiratory therapy conducted an internal conceptual review to delimit intended use, biofeedback interpretation, and safety requirements for future evaluations.

Keywords: biofeedback, convolutional neural networks, edge computing, local inference, machine learning, quantization.

Introducción

La rehabilitación pulmonar combina evaluación, entrenamiento, educación y cambio de comportamiento para mejorar la condición física y psicológica de personas con enfermedad respiratoria crónica (Spruit *et al.*, 2013). Sin embargo, la práctica domiciliaria sigue planteando problemas de adherencia, continuidad y ejecución técnica cuando el profesional no está presente. Los programas domiciliarios y de telerehabilitación pueden reducir las barreras de acceso, pero requieren mecanismos comprensibles de seguimiento y de retroalimentación que no sustituyan el juicio clínico (Cox *et al.*, 2021; Holland *et al.*, 2017; McCarthy *et al.*, 2015; Watson *et al.*, 2023).

La gamificación puede aportar metas visibles y refuerzo inmediato en tareas repetitivas y su uso en intervenciones dirigidas a personas con enfermedades respiratorias crónicas se ha estudiado como estrategia para apoyar el compromiso y el manejo de la actividad (Chen *et al.*, 2025). Los teléfonos inteligentes añaden micrófono, capacidad de cómputo local y una interfaz familiar, lo que ha motivado trabajos de análisis respiratorio acústico y biorretroalimentación móvil (Brown *et al.*, 2020; Larson *et al.*, 2012; Shih *et al.*, 2019). Breeze constituye una referencia relevante de interacción porque detecta fases respiratorias en tiempo real; Pulneo no reproduce ese objetivo. El presente trabajo utiliza un detector binario de actividad respiratoria y, por tanto, Breeze se considera un antecedente de biorretroalimentación acústica, mas no un clasificador directamente equivalente.

El procesamiento local resulta pertinente para señales acústicas potencialmente sensibles porque reduce la necesidad de transmitir audio crudo y elimina la dependencia de conectividad para la inferencia. Modelos compactos basados en MFCC y redes convolucionales pueden ejecutarse en dispositivos con recursos limitados, aunque su utilidad depende de conservar el mismo preprocesamiento entre entrenamiento y despliegue (Abadade *et al.*, 2024; Rivas-Navarrete *et al.*, 2025; Warden & Situnayake, 2019).

Pulneo integra un *pipeline* de entrenamiento, un contrato de preprocesamiento versionado y una aplicación Flutter que captura audio, extrae MFCC, ejecuta un modelo TFLite INT8 y actualiza una interfaz gamificada. El objetivo es estudiar la viabilidad técnica del uso de la actividad respiratoria acústica como indicador indirecto de continuidad durante ejercicios guiados; no diagnosticar, estimar función pulmonar ni demostrar la eficacia terapéutica.

Objetivo general

Diseñar e implementar un prototipo móvil gamificado para apoyar ejercicios respiratorios domiciliarios mediante detección acústica de actividad respiratoria ejecutada localmente con TensorFlow Lite, integrando biorretroalimentación visual, conteo de repeticiones y registro de sesión.

Objetivos específicos

- Construir un *pipeline* auditable de entrenamiento y evaluación para detección de actividad respiratoria usando respiración profunda como clase positiva y sonidos ambientales como clase negativa.
- Cuantificar el desempeño del clasificador y comparar TFLite FP32 frente a TFLite INT8 en las mismas muestras de prueba.
- Definir un contrato único de preprocesamiento, normalización, semántica de clases y umbral para el despliegue móvil.
- Implementar en Flutter captura de audio, extracción MFCC, inferencia en isolate y una máquina de estados de biorretroalimentación.
- Incorporar una revisión fisioterapéutica conceptual del protocolo para analizar la relación entre la señal acústica, la tarea respiratoria guiada, la dosificación, la seguridad y la interpretación de los indicadores de proceso.
- Delimitar explícitamente los alcances técnicos y clínicos antes de una evaluación con participantes.

Materiales y métodos

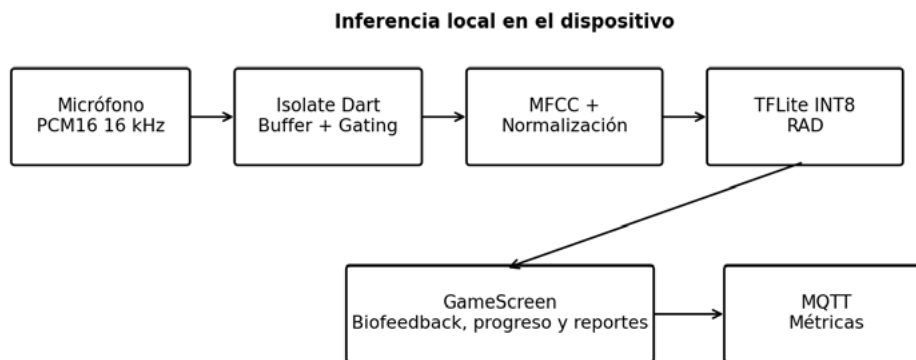
Diseño del sistema

Este trabajo se planteó como el desarrollo y la evaluación técnica de un prototipo, por lo que no corresponde a un ensayo clínico. La solución se organizó en dos componentes principales. El primero es un *pipeline* desarrollado en Python, utilizado para preparar los datos, entrenar y calibrar el modelo, evaluar su desempeño y generar los archivos necesarios para el despliegue. El segundo corresponde a una aplicación móvil desarrollada en Flutter, encargada de capturar el audio, extraer las características acústicas y ejecutar la inferencia directamente en el dispositivo.

Para mantener la consistencia entre el entrenamiento y la aplicación móvil, la configuración de despliegue se definió en el archivo `normalizer_mfcc.json`. Este archivo reúne los parámetros necesarios para reproducir el procesamiento de entrada: dimensiones esperadas, configuración de los MFCC, estadísticas utilizadas para la normalización, correspondencia entre clases, interpretación de la salida y umbral de decisión del modelo cuantizado en INT8.

Se utilizó la convención 0 = ruido/silencio y 1 = actividad respiratoria. La salida sigmoide del modelo se interpreta como la probabilidad estimada de actividad respiratoria, es decir, $P(\text{respiración})$. En la aplicación móvil, una ventana se clasifica como respiración cuando esta probabilidad es mayor o igual a 0,29. La arquitectura general del sistema se presenta en la Figura 1.

Figura 1
Arquitectura general de Pulneo



La Figura 1 resume el flujo de ejecución. El audio PCM16 mono, con una frecuencia de muestreo de 16 kHz, se procesa en un isolate de Dart, donde se aplica el contrato MFCC y se ejecuta el TFLite INT8. La salida alimenta la pantalla de juego (*gamescreen*) mediante códigos de estado semánticos. La telemetría MQTT es experimental y no transmite audio crudo.

Fundamento fisioterapéutico de la biorretroalimentación

Pulneo utiliza la presencia sostenida de actividad respiratoria acústica como señal de proceso durante un ejercicio guiado de exhalación. No mide flujo, volumen pulmonar, saturación de oxígeno, esfuerzo muscular ni fase respiratoria; por ello, el avance de la barra indica continuidad acústica detectada dentro de una instrucción externa, no una espiración terapéuticamente correcta. La calidad del patrón ventilatorio, el esfuerzo, la coordinación y la tolerancia requieren valoración profesional. La biorretroalimentación estructura la práctica y aporta una referencia temporal, sin sustituir el razonamiento clínico (Frownfelter & Dean, 2012; Pryor & Prasad, 2008).

La aplicación también distingue *clipping* o saturación del micrófono. Este evento se reporta como saturación acústica y no como “mala técnica”, porque puede depender de ganancia, distancia al micrófono, hardware o intensidad del sonido. La interpretación clínica permanece a cargo del profesional.

Conceptualización clínica y revisión conceptual fisioterapéutica del protocolo

La conceptualización clínica distinguió el objetivo terapéutico, la variable registrada por el teléfono y la respuesta de la interfaz. El ejercicio es una tarea respiratoria guiada; el teléfono solo detecta actividad acústica compatible con respiración, que la aplicación representa como continuidad. Por ello, un estado *breathing* no permite determinar volumen, flujo espiratorio, patrón diafrágico, compensaciones ni tolerancia clínica.

Desde la perspectiva fisioterapéutica, se revisó si la señal podía apoyar de forma coherente la práctica domiciliaria. El diseño priorizó continuidad y dosificación, no intensidad: la retroalimentación no debe incentivar una señal más fuerte y la saturación del micrófono se trata como un problema de adquisición, no como de desempeño respiratorio. Así se evita convertir la gamificación en una competencia por respirar con mayor intensidad (Bott *et al.*, 2009; Frownfelter & Dean, 2012; Pryor & Prasad, 2008).

La revisión fue realizada por un autor fisioterapeuta con maestría en terapia respiratoria como evaluación experta interna del prototipo. Se examinó la relación entre tarea, la señal acústica, la biorretroalimentación y las conclusiones permitidas, además de la detección positiva, el tiempo objetivo, la dosificación, el conteo de repeticiones, la saturación, los reportes y el papel del profesional.

No fue una validación formal: no se aplicó un instrumento estandarizado, ni participaron evaluadores independientes, ni se realizaron rondas de consenso. Los resultados son criterios de una revisión experta interna. Se estableció que la duración fuera configurable, que una detección continua contara como una repetición, que el nuevo conteo requiriera transición a *noBreathing* y que las alertas de adquisición no se presentaran como valoración automática de la técnica.

La pertinencia terapéutica se definió por el uso previsto, sin asumir eficacia clínica. Pulneo complementa una tarea indicada y explicada por un profesional; no selecciona tratamientos, no ajusta dosis, no decide la finalización clínica ni sustituye la valoración de signos, síntomas o compensaciones. Duración, repeticiones, frecuencia y criterios de interrupción deben individualizarse según valoración y protocolo clínico (Spruit *et al.*, 2013).

Los datos de sesión (repeticiones, tiempo de práctica, interrupciones y clipping) describen la interacción con el prototipo, pero no la función pulmonar, la respuesta fisiológica, la calidad técnica ni el beneficio terapéutico. En telemonitoreo pueden aportar información al profesional, siempre que esa información se interprete junto con la evaluación clínica y el reporte del paciente.

La evaluación se planteó por etapas: validar técnicamente el detector en condiciones reales y distintos dispositivos; comprobar la concordancia con anotaciones profesionales y usabilidad supervisada; y, solo después, estudiar la adherencia o los resultados clínicos. Esta secuencia evita equiparar el desempeño en conjuntos de datos públicos con validación fisioterapéutica.

Conjuntos de datos y partición

La clase positiva se construyó con Coswara (Sharma *et al.*, 2020): en `combined_data.csv` se seleccionaron solo sujetos con `covid_status = healthy` y el archivo `breathing-deep.wav` de cada sujeto. Cada audio se cargó en mono a 16 kHz, se normalizó por su máximo absoluto y se segmentó en ventanas de 1 s con salto de 0,5 s; los audios menores a 1 s se completaron con ceros. Se omitieron archivos no legibles y ventanas con $RMS < 0,0001$, obteniéndose 34 817 ventanas de 1061 sujetos. Para la clase negativa, los WAV de UrbanSound8K (Salamon *et al.*, 2014) se barajaron con semilla 42 y recibieron el mismo procesamiento y filtro RMS. Se procesaron archivos hasta acumular 34 817 ventanas de 5608 archivos, igualando la clase positiva; no se usaron negativos sintéticos.

La partición fue aleatoria, estratificada por clase y realizada a nivel de grupo con semilla 42. Las proporciones objetivo fueron 70 % entrenamiento, 15 % validación y 15 % prueba. Todas las ventanas de un mismo sujeto Coswara o archivo negativo permanecieron en un único subconjunto. La auditoría de la partición verificó cero solapamientos de grupos entre entrenamiento, validación y prueba. Esta decisión reduce el riesgo de que ventanas altamente relacionadas del mismo sujeto o archivo aparezcan en más de un subconjunto y permite una estimación más conservadora que una partición aleatoria a nivel de ventana.

La estrategia de partición controla el solapamiento directo entre sujetos y archivos, pero el diseño del conjunto mantiene una consideración adicional de validez externa: existe una asociación entre la clase y la procedencia, ya que las ventanas positivas provienen de Coswara y las negativas de UrbanSound8K. En consecuencia, no puede descartarse que el clasificador aproveche, además de rasgos propios de la actividad respiratoria, características del dominio de adquisición como dispositivo, ganancia, protocolo de grabación o ruido de fondo. Este aspecto no invalida el experimento ni su utilidad como evaluación técnica inicial, pero delimita su interpretación: el rendimiento reportado cuantifica discriminación dentro de los dominios públicos utilizados y constituye una referencia de ingeniería para el prototipo, no una estimación directa de generalización a grabaciones respiratorias domiciliarias. La siguiente fase de validación contempla datos recogidos con la propia aplicación y negativos del entorno objetivo para medir esa generalización de forma específica.

Preprocesamiento acústico

Los audios se convirtieron a mono y se remuestrearon a 16 kHz. Se utilizaron ventanas de 1 s (16 000 muestras) con salto de 0,5 s (8000 muestras). De cada ventana se extrajeron 40 MFCC con 128 bandas Mel, FFT de 512 puntos, ventana Hann de 400 muestras, salto STFT de 160 muestras y banda de 50 a 8000 Hz. Los MFCC constituyen una representación clásica de la envolvente espectral utilizada ampliamente en reconocimiento y análisis acústico (Davis & Mermelstein, 1980; Rabiner & Schafer, 2011), mientras que la parametrización temporal y espectral sigue principios estándar de procesamiento digital de señales (Oppenheim & Schafer, 2010).

El contrato de características especifica `center=false`, escala Mel Slaney, normalización Slaney y conversión `power_to_db` con `top_db=80`. La implementación móvil utiliza estos mismos parámetros y no aplica preénfasis. La forma de entrada resultante es $97 \times 40 \times 1$. El empleo de representaciones cepstrales y características espectrales en análisis respiratorio cuenta con antecedentes tanto en clasificación de sonidos respiratorios como en detección de fases respiratorias (Rocha *et al.*, 2019; Zhantleuova *et al.*, 2025). Los parámetros del contrato de inferencia se resumen en la Tabla 1.

Tabla 1
Parámetros de preprocesamiento acústico del contrato de inferencia.

PARÁMETRO	VALOR	USO
Frecuencia de muestreo	16000 Hz, mono	Entrada común de entrenamiento y app
Ventana / salto de audio	1 s / 0,5 s	Inferencia solapada al 50 %
MFCC / bandas Mel	40 / 128	Representación acústica compacta
FFT / ventana / salto STFT	512 / 400 / 160	Contrato de espectrograma
Mel	Slaney; norm=Slaney	Alineación Python-Dart
<code>power_to_db</code>	<code>top_db=80</code>	Escala logarítmica definida
Banda	50–8000 Hz	Rango procesado
Forma	$97 \times 40 \times 1$	Entrada fija de la CNN

Normalización y prevención de fuga de datos

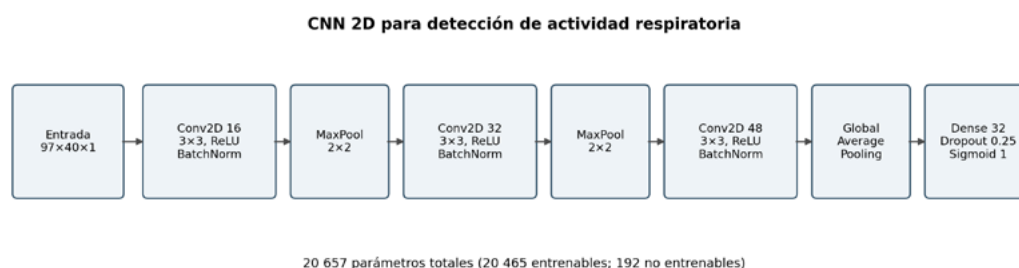
La media y la desviación estándar de cada coeficiente MFCC se estimaron exclusivamente con el conjunto de entrenamiento y luego se reutilizaron en validación, prueba y aplicación móvil. La operación fue $z = (x - \mu_{train}) / (\sigma_{train} + \epsilon)$. El contrato JSON conserva los 40 valores de media y desviación estándar, de modo que la inferencia móvil no recalcula estadísticas con datos de usuario.

Modelo de detección de actividad respiratoria

Se implementó una CNN 2D con tres convoluciones 3×3 de 16, 32 y 48 filtros, normalización por lotes, dos capas MaxPooling 2×2 , Global Average Pooling, una capa densa de 32 unidades con ReLU, Dropout de 0,25 y una salida sigmoide. El modelo contiene 20 657 parámetros totales, de los cuales 20 465 son entrenables y 192 no entrenables. Se optimizó con Adam, *binary crossentropy*, *batch size* 32, *learning rate inicial* 0,001, *early stopping* y reducción adaptativa de la tasa de aprendizaje. La selección de una arquitectura convolucional compacta se apoya en componentes estándar de aprendizaje profundo y en su aplicación previa en informática de la salud (Goodfellow *et al.*, 2016; Ravi *et al.*, 2017).

La clase positiva es actividad respiratoria (1) y la negativa ruido/silencio (0). El umbral de cada artefacto se calibró de manera independiente y exclusivamente sobre el conjunto de validación, seleccionando el valor que maximizó F1; una vez seleccionado, se mantuvo fijo durante la evaluación en test. Keras FP32 y TFLite FP32 seleccionaron 0,53, mientras que TFLite INT8 seleccionó 0,29. El umbral de despliegue de la aplicación es, por tanto, $P(\text{respiración}) \geq 0,29$. La arquitectura del detector se presenta en la Figura 2.

Figura 2
Arquitectura de la CNN 2D utilizada para el detector RAD



Cuantización y protocolo de evaluación

A partir del mismo modelo Keras se exportaron dos artefactos TFLite: TFLite FP32, utilizado como modelo de referencia para el despliegue, y TFLite INT8, utilizado como modelo cuantizado final. Keras FP32 se conserva como referencia del entrenamiento; la comparación destinada a aislar el efecto de cuantización se realiza entre los dos artefactos TFLite para mantener constante el entorno de ejecución. La cuantización INT8 utilizó 100 muestras representativas del conjunto de entrenamiento. TFLite FP32 y TFLite INT8 se ejecutaron con `tf.lite.Interpreter` sobre exactamente las mismas 10 347 ventanas de test, después de 20 muestras de calentamiento.

Se calcularon *accuracy*, precisión, *recall*/sensibilidad, especificidad, F1, ROC-AUC, PR-AUC, *balanced accuracy*, MCC y matriz de confusión. ROC-AUC y PR-AUC se calcularon sobre probabilidades continuas, antes del umbral. También se midió la concordancia entre las probabilidades y entre las decisiones, tamaño de archivo y latencia media por muestra. Las cifras de latencia corresponden al entorno de evaluación TFLite del *pipeline* y no constituyen un *benchmark* de un teléfono específico.

Entorno de software y reproducibilidad

La ejecución se realizó con Python 3.12.10 y semilla aleatoria 42. Los artefactos generados identifican TensorFlow 2.21.0 y Keras 3.15.1. El proyecto Flutter fija Dart SDK $\geq 3.9.2$ y Flutter $\geq 3.35.6$; las dependencias de inferencia y audio incluyen `tflite_flutter` 0.12.1, `record` 6.2.0 y `fft` 0.5.1. Esta configuración documenta el entorno principal utilizado para el entrenamiento, la evaluación y la integración móvil.

Implementación móvil

La aplicación Flutter captura PCM16 mono a 16 kHz y envía los buffers a un isolate. El isolate completa la inicialización del intérprete antes de abrir el flujo de audio, aplica *gating* y detección de *clipping*, calcula MFCC según el contrato, normaliza las características y ejecuta el modelo TFLite INT8. La salida cuantizada se descuantiza y se interpreta como P(respiración).

El protocolo de inferencia utiliza el tipo `AudioInferenceStatus` con cuatro estados: *fatalError* (−1), *noBreathing* (0), *breathing* (1) y *clipping* (2). Estos valores representan estados de la aplicación, mientras que el modelo subyacente efectúa una clasificación de dos clases. La pantalla de juego recibe estados semánticos tipados para actualizar la biorretroalimentación y las alertas.

La duración de la sesión se mide con un reloj de ejecución. El conteo de repeticiones requiere una transición a *noBreathing* para rearmar una nueva repetición, de modo que una detección respiratoria continua solo puede completar una repetición. La telemetría MQTT se mantiene como función experimental y configurable; no transmite audio crudo. El contrato de inferencia y sus códigos de estado se muestran en la Figura 3. Los componentes principales del sistema y su alcance actual se resumen en la Tabla 2.

Figura 3
Contrato de inferencia y códigos de estado utilizados por la aplicación

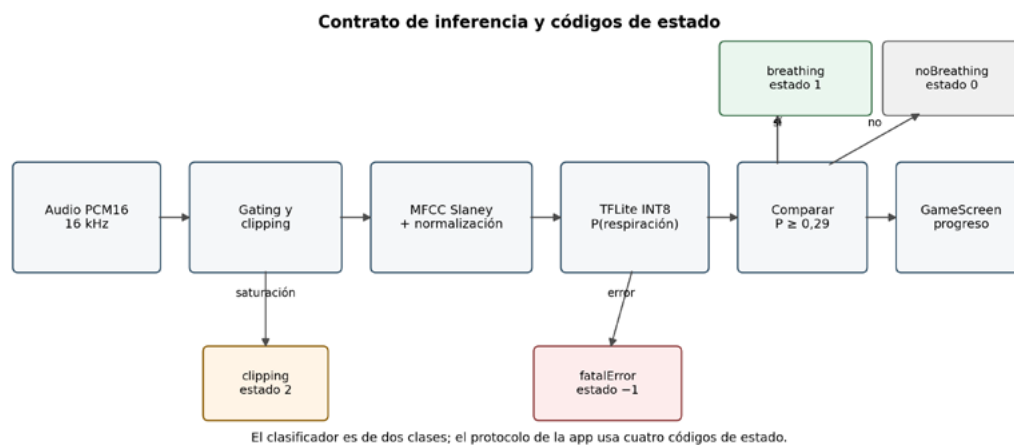


Tabla 2
Componentes principales del sistema y alcance actual

COMPONENTE	FUNCIÓN	ALCANCE ACTUAL
Pipeline Python	Datos, MFCC, entrenamiento, auditoría y exportación FP32/INT8	Evaluación técnica; no validación clínica
Modelo RAD	Dos clases: ruido/silencio frente a actividad respiratoria	No distingue inspiración/exhalación
Contrato JSON v2	Preprocesamiento, normalización, clases y <i>threshold</i> INT8	$P(\text{respiración}) \geq 0,29$
Isolate Dart	MFCC e inferencia fuera del hilo UI	Evita bloquear la interfaz
Pantalla de juego	Estado, progreso, repeticiones y <i>clipping</i>	Biorretroalimentación de proceso, no medición fisiológica
MQTT	Telemetría resumida	Experimental; sin audio crudo

Heurísticas previas a la inferencia

Las heurísticas de amplitud, *clipping* y factor de cresta pertenecen solo a la aplicación móvil: no se aplicaron a las 10 347 ventanas de test ni intervienen en las métricas reportadas.

En el entorno de ejecución, antes de MFCC/TFLite, descartan amplitud insuficiente, separan *clipping* y filtran impulsos. Controlan calidad de la entrada, pero no la técnica respiratoria ni las variables clínicas.

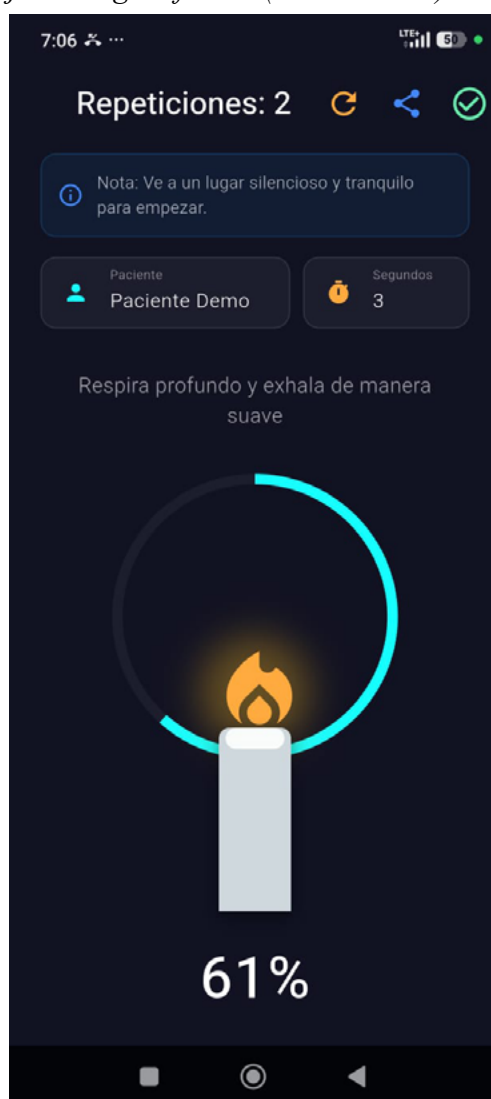
Gamificación y lógica de biorretroalimentación

La barra de progreso aumenta únicamente cuando el estado recibido es *breathing* y disminuye cuando la actividad se interrumpe. El *clipping* activa una alerta y no se contabiliza como respiración válida. Al completar el intervalo configurado, se registra una repetición y el sistema queda bloqueado hasta observar *noBreathing*, lo que impide múltiples repeticiones durante una detección continua.

Dado que el modelo no identifica la fase respiratoria, el estado *breathing* se usa como indicador acústico dentro de una instrucción guiada de exhalación. La interfaz se concibe como biorretroalimentación de proceso y no como sensor de flujo ni como clasificador específico de fase respiratoria. Esta elección mantiene sencilla la tarea de inferencia en la etapa actual y delimita con claridad qué debe comprobarse en la validación externa con datos del ejercicio objetivo. La interfaz implementada durante el ejercicio se muestra en la Figura 4.

Figura 4

Pantalla principal de ejercicio gamificado (GameScreen) durante una sesión guiada



En esta vista, el porcentaje y el anillo circular representan el progreso temporal del ejercicio configurado; el conteo superior registra repeticiones completadas y el tiempo objetivo se muestra en segundos. Estos elementos constituyen biorretroalimentación de proceso y no una medición de flujo, volumen pulmonar u otra variable fisiológica.

Plan de evaluación y criterios de seguridad

No se reclutaron pacientes ni se realizó intervención clínica. La evaluación futura se plantea en tres etapas: primero, comprobar el detector con audios del escenario objetivo anotados por profesionales, incluidos exhalaciones, inhalaciones, habla, tos, pausas, ruido doméstico y distintas distancias; segundo, con aprobación ética y consentimiento informado, evaluar comprensión, aceptabilidad, carga de uso y concordancia con la observación del fisioterapeuta; y tercero, estudiar adherencia y resultados terapéuticos. Cada etapa deberá definir los criterios de inclusión y exclusión, las reglas de detención, la supervisión y protección de datos, sin delegar decisiones clínicas a la aplicación.

La caracterización computacional de esta etapa se concentró en variables reproducibles del *pipeline*: tamaño de modelo y latencia del intérprete TFLite bajo el mismo entorno experimental. Las mediciones de consumo energético, memoria pico, batería, temperatura y latencia en modelos concretos de teléfonos inteligentes se reservan para la fase de *benchmark* móvil, donde podrán compararse bajo hardware y condiciones de uso definidos.

Resultados

Composición del conjunto de datos y partición

El conjunto de datos final incluyó 69 634 ventanas perfectamente balanceadas: 34 817 por clase. Los 1061 sujetos Coswara y 5608 archivos negativos se distribuyeron por grupos, sin solapamiento entre subconjuntos. La fracción real de ventanas fue 70,10 % en entrenamiento, 15,05 % en validación y 14,86 % en prueba, próximas a las proporciones objetivo de 70/15/15. La composición de cada subconjunto se presenta en la Tabla 3.

Tabla 3
Tamaño y composición real de train, validation y test

SPLIT	VENT.	RESP.	RUIDO	SUJETOS	NEG.
Train	48 810	24 413	24 397	743	3926
Val.	10 477	5229	5248	159	841
Test	10 347	5175	5172	159	841
Total	69 634	34 817	34 817	1061	5608

Rendimiento del clasificador

En test ($n = 10\,347$), Keras FP32 y TFLite FP32 coincidieron en las métricas principales y en la matriz de confusión: 97,59 % de *accuracy*, 98,28 % de precisión, 96,89 % de *recall*, 98,30 % de especificidad, F1 de 97,58 % y ROC-AUC de 99,71 %, con umbral 0,53 calibrado en validación. Para aislar el efecto de la cuantización, la comparación reportada en la Tabla 4 utiliza TFLite FP32 como modelo de referencia de despliegue y TFLite INT8 como modelo cuantizado, ambos ejecutados con `tf.lite.Interpreter` sobre las mismas 10 347 ventanas de prueba. TFLite INT8 alcanzó 97,64 % de *accuracy*, 98,09 % de precisión, 97,18 % de *recall*, 98,11 % de especificidad, F1 de 97,63 % y ROC-AUC de 99,65 %. Su umbral de 0,29 se fijó previamente en validación (F1 = 96,79 %) y no se reajustó en test.

Tabla 4
Rendimiento en test de TFLite FP32 y TFLite INT8

TIPO	TH.	ACC.	PREC.	REC.	F1	AUC
TFLite FP32	0,53	97,59	98,28	96,89	97,58	99,71
TFLite INT8	0,29	97,64	98,09	97,18	97,63	99,65

En la Tabla 4, las métricas se expresan en porcentaje, excepto el umbral. La fila TFLite FP32 corresponde al artefacto float exportado a TensorFlow Lite, no al modelo Keras; ambos coincidieron en las métricas principales de test, pero se utiliza TFLite FP32 para mantener el mismo entorno de ejecución que INT8. PR-AUC fue 99,71 % para FP32 y 99,67 % para INT8; balanced *accuracy*, 97,59 % y 97,64 %; y MCC, 0,9520 y 0,9529, respectivamente. La matriz de confusión del TFLite INT8 se presenta en la Tabla 5.

Tabla 5
Matriz de confusión del TFLite INT8 en test

CLASE REAL	PRED. RUIDO/SILENCIO	PRED. RESPIRACIÓN
Ruido/silencio	5074 (TN)	98 (FP)
Respiración	146 (FN)	5029 (TP)

Comparación FP32 frente a INT8 y huella del modelo

La comparación pareada entre TFLite FP32 e INT8 se realizó sobre las mismas 10 347 muestras y con el mismo entorno de ejecución. El F1 aumentó en 0,00054 en INT8, mientras que ROC-AUC disminuyó en 0,00058. La correlación entre probabilidades fue 0,99699, el MAE fue 0,01164 y el acuerdo de decisiones usando los umbrales de despliegue fue 99,55 %. En conjunto, estas cifras muestran que la capacidad discriminativa y el orden relativo de las puntuaciones se conservaron en gran medida tras la cuantización.

Interpretación de los umbrales de decisión. TFLite FP32 e INT8 mostraron puntuaciones altamente correlacionadas, aunque sus umbrales óptimos de validación fueron 0,53 y 0,29. Esto es compatible con salidas globalmente similares, porque el umbral que maximiza F1 depende de la distribución cercana a la frontera de decisión y de la calibración de cada artefacto. Ambos umbrales se optimizaron por separado en *validation* y se aplicaron sin reajuste en test; por tanto, la diferencia se interpreta como calibración del punto de decisión y no como evidencia de pérdida discriminativa ni se atribuye causalmente a la cuantización.

El tamaño del modelo TFLite disminuyó de 88 488 a 31 600 bytes, una reducción de 64,29 % (2,80×). El archivo Keras ocupa 310 023 bytes. La cuantización se utilizó principalmente para reducir la huella de almacenamiento; en el entorno experimental, la latencia media fue de 0,401 ms para INT8 y 0,215 ms para FP32 después del calentamiento. Por ello, el beneficio observado en esta ejecución fue la compacidad del artefacto, mientras que la eventual ventaja de velocidad, energía o memoria debe determinarse mediante *benchmarks* específicos por dispositivo. El efecto de la cuantización sobre tamaño, métricas y latencia se resume en la Tabla 6.

Tabla 6
Efecto de cuantización TFLite FP32 frente a TFLite INT8

INDICADOR	TFLITE FP32	TFLITE INT8	CAMBIO
Tamaño	88 488 bytes	31 600 bytes	-64,29 %
F1	97,577 %	97,632 %	+0,054 pp
ROC-AUC	99,711 %	99,653 %	-0,058 pp
Latencia media*	0,215 ms	0,401 ms	+87,00 %
Acuerdo de decisiones	—	99,55 % vs. FP32	0,45 % desacuerdo

* Medición con tf.lite.Interpreter, una muestra por invocación, después de 20 muestras de calentamiento; no corresponde a un benchmark de teléfonos inteligentes.

Integración técnica del prototipo

La integración móvil utiliza la misma semántica definida por el entrenamiento: P(respiración) se evalúa directamente con el *threshold* INT8 de 0,29. El protocolo de estados distingue *noBreathing*, *breathing*, *clipping* y *fatalError*. La barra de progreso avanza durante el estado *breathing*, se interrumpe fuera de ese estado y aplica una alerta independiente ante *clipping*. El reporte de sesión utiliza la duración medida y el contador de repeticiones exige una transición a *noBreathing* entre repeticiones.

Estos resultados respaldan la coherencia interna del *pipeline* y del contrato de despliegue. La evaluación cuantitativa presentada corresponde al clasificador sobre los conjuntos de datos descritos; la correspondencia entre el estado acústico y una exhalación guiada en condiciones reales constituye el objetivo específico de la siguiente fase de validación con datos del escenario de uso.

Resultado de la revisión conceptual fisioterapéutica

La revisión conceptual fisioterapéutica definió la biorretroalimentación como indicador de continuidad acústica dentro de una tarea guiada y fijó cuatro decisiones: no asociar mayor amplitud con mejor ejecución; separar *clipping* del juicio de técnica; contar una sola repetición por episodio continuo; y mantener la dosificación temporal bajo control del protocolo o del profesional.

Duración, repeticiones, interrupciones y saturación se consideran resultados de proceso; no se presentan como medidas de ventilación, función pulmonar, disnea ni respuesta terapéutica.

La evaluación futura seguirá una secuencia: concordancia técnica en el dominio de uso, usabilidad y concordancia observacional bajo supervisión, y después estudios de adherencia o efecto terapéutico. La revisión actual es experta e interna; una validación formal de contenido requerirá varios evaluadores y un método definido, y la validación clínica, estudios con participantes.

Discusión

El INT8 alcanzó F1 de 97,63 % y ROC-AUC de 99,65 % en un test separado por grupos. La ausencia de solapamiento de sujetos y archivos reduce la fuga de datos y permite usar estas cifras para comparar versiones del *pipeline* y la cuantización. Sin embargo, describen el experimento Coswara/UrbanSound8K y no demuestran desempeño clínico ni generalización domiciliaria.

La cuantización redujo el tamaño del modelo en un 64,29 % y mantuvo 99,55 % de acuerdo con TFLite FP32. El cambio de umbral de 0,53 a 0,29, calibrado independientemente en validation, no implica por sí mismo pérdida discriminativa. En el entorno medido, INT8 fue más compacto pero más lento; la estabilidad del umbral, la latencia, la energía y la memoria deberán evaluarse en teléfonos inteligentes concretos.

La principal consideración para la validez externa es la diferencia entre el dominio de entrenamiento y el escenario final de uso. La clase positiva procede de breathing-deep.wav de sujetos sanos de Coswara, con ciclos de respiración profunda registrados bajo el protocolo de ese conjunto de datos, mientras que la aplicación guía una exhalación prolongada y controlada frente al micrófono de un teléfono. Esta elección permitió construir y depurar un clasificador experimental de actividad respiratoria con datos públicos y reproducibles antes de realizar recolección propia. El rendimiento obtenido caracteriza la discriminación dentro de esos dominios públicos; la generalización al ejercicio domiciliario se trata como una pregunta experimental independiente. Fase respiratoria, distancia, orientación, micrófono, ganancia, duración, ruido y características del usuario pueden modificar la distribución acústica, por lo que la evaluación futura debe medir explícitamente el desempeño en esas condiciones.

Esta delimitación también permite diferenciar Pulneo de Breeze debido a sus distintos objetivos. Breeze fue diseñado para detectar fases respiratorias, mientras que Pulneo clasifica actividad respiratoria frente a ruido y no separa inspiración de exhalación. En Pulneo, la actividad acústica se utiliza como una señal de proceso dentro de una instrucción guiada, una simplificación deliberada que facilita la inferencia local y la integración con la interfaz gamificada en esta etapa. La evolución prevista incorpora un conjunto específico del escenario objetivo, con exhalaciones etiquetadas, múltiples teléfonos, distancias y ruidos domésticos, además de negativos difíciles como voz, tos, manipulación del equipo e inhalaciones, con el fin de evaluar si el detector mantiene su utilidad fuera de los dominios públicos originales.

Desde el punto de vista clínico, Pulneo se diseñó como una herramienta de apoyo y no como un instrumento de medición fisiológica: el micrófono aporta información acústica, mientras que variables como flujo, volumen, esfuerzo y respuesta terapéutica requieren sensores o evaluaciones diferentes. Esta delimitación permite mantener una arquitectura móvil sencilla y no invasiva para explorar la biorretroalimentación y el registro del proceso sin atribuir a la señal acústica capacidades que no posee. En esta etapa tampoco se evaluaron adherencia, calidad de vida o concordancia con fisioterapeutas; esas variables pertenecen a las fases posteriores de evaluación clínica y de usabilidad. La contribución actual es de ingeniería y diseño clínico-conceptual: un flujo auditable de datos, un clasificador compacto, un contrato explícito entre Python y Dart y una aplicación que realiza inferencia local sin transmitir audio crudo.

La conceptualización clínica aporta además un criterio para interpretar la gamificación. En este contexto, la barra de progreso no representa una magnitud fisiológica continua; representa el cumplimiento temporal de una condición acústica durante una tarea previamente explicada. Esa distinción cambia el papel de la biorretroalimentación: su finalidad inmediata es orientar atención, ritmo y persistencia, no cuantificar capacidad pulmonar. El valor potencial de la interfaz depende de que el usuario comprenda esa relación y de que el profesional pueda establecer una tarea compatible con las capacidades y objetivos individuales. Esta visión es coherente con la rehabilitación pulmonar como intervención individualizada que combina entrenamiento, educación y cambio de comportamiento, más que como ejecución aislada de una maniobra (Spruit *et al.*, 2013).

Desde el análisis fisioterapéutico, la principal tensión del prototipo se encuentra entre la simplicidad de la interacción y la complejidad del fenómeno respiratorio. Una exhalación terapéutica puede requerir control de ritmo, coordinación toracoabdominal, manejo de secreciones, tolerancia al esfuerzo y adaptación a síntomas, dimensiones que no pueden inferirse únicamente a partir de los MFCC (Bott *et al.*, 2009; Osadnik *et al.*, 2012). Por ello, incrementar la precisión del clasificador no elimina la necesidad de una capa clínica de interpretación. Incluso un detector acústico perfecto de respiración seguiría siendo insuficiente para certificar una técnica correcta si no se observan flujo, movimiento, esfuerzo y respuesta del paciente. El diseño debe mantener esta asimetría visible para evitar una falsa sensación de medición fisiológica.

La supervisión de pertinencia terapéutica también orienta el desarrollo futuro del sistema. Antes de ampliar la mecánica con puntuaciones, niveles o recompensas, cada elemento gamificado debería evaluarse por su efecto sobre la conducta que se desea facilitar. Recompensar la intensidad podría inducir esfuerzo innecesario; recompensar la velocidad podría acortar una fase que se pretende sostener; y penalizaciones excesivas podrían aumentar la frustración o la ansiedad. En cambio, metas temporales comprensibles, retroalimentación inmediata y registro de cumplimiento pueden ser compatibles con un enfoque educativo si se mantienen subordinados a la prescripción y a la tolerancia individual. La evaluación con usuarios debe estudiar estas respuestas y no asumir que una mayor interacción equivale automáticamente a mejor adherencia o mejor técnica.

Alcance y limitaciones

El conjunto positivo se restringió a sujetos sanos de Coswara para disponer de una referencia pública, homogénea y reproducible durante la construcción inicial del detector. UrbanSound8K se utilizó para aportar una clase negativa amplia de sonidos ambientales reales. Estas decisiones son apropiadas para una primera evaluación técnica controlada, aunque no pretenden representar toda la población clínica ni todos los eventos acústicos de un hogar. Por ello, la siguiente etapa incorpora como objetivo un conjunto externo recolectado con la propia aplicación, incluyendo población y condiciones de uso más próximas al escenario terapéutico.

El modelo identifica actividad respiratoria y no fue entrenado para separar fases respiratorias. En el protocolo actual, esta salida se utiliza como indicador de continuidad durante una exhalación guiada, lo que permite estudiar la interacción y el procesamiento local sin presentar el sistema como detector validado de exhalación. Una fase posterior incorporará etiquetas específicas de inhalación/exhalación y anotación profesional para determinar si una tarea más especializada mejora la correspondencia entre la señal acústica y el ejercicio terapéutico.

La evaluación de rendimiento computacional presentada se realizó con el entorno de ejecución TFLite del entorno experimental, lo que permite una comparación controlada entre FP32 e INT8 bajo las mismas condiciones. Para caracterizar el comportamiento del producto en hardware móvil, la fase de evaluación por dispositivo debe añadir *benchmarks* en modelos concretos de teléfonos inteligentes y mediciones de memoria, energía, batería y temperatura. De forma análoga, la validación con participantes se reserva para una etapa posterior con protocolo ético y anotación profesional.

La trazabilidad del experimento registra Python 3.12.10, la semilla aleatoria 42, las versiones de TensorFlow/Keras, las restricciones de Dart/Flutter y las dependencias móviles. Estos datos, junto con los artefactos de entrenamiento, normalización, partición del conjunto de datos y evaluación, permiten reconstruir las principales condiciones computacionales del experimento.

Conclusiones

Pulneo integra un *pipeline* de entrenamiento y evaluación con una aplicación Flutter para inferencia acústica local. La aplicación interpreta la salida del modelo como P(respiración), utiliza el umbral INT8 calibrado en validación y gestiona los resultados mediante códigos de estado explícitos.

En el conjunto de prueba, el TFLite INT8 alcanzó 97,64 % de *accuracy*, F1 de 97,63 % y ROC-AUC de 99,65 %. La cuantización redujo el tamaño del modelo en un 64,29 % y mantuvo un 99,55 % de acuerdo de decisiones respecto del TFLite FP32, aunque no produjo una mejora de latencia en el entorno evaluado.

Estos resultados respaldan la viabilidad técnica del detector dentro del dominio evaluado y justifican avanzar a una validación externa más próxima al uso previsto. La siguiente etapa está orientada a cuantificar la generalización con datos propios del ejercicio, *benchmarks* en teléfonos concretos y evaluación con anotación profesional y un protocolo ético. En paralelo, la conceptualización fisioterapéutica define un marco de uso en el que la señal acústica se interpreta como indicador de proceso, la dosificación permanece bajo supervisión profesional y la validación progresa desde concordancia técnica hasta estudios clínicos posteriores. Este planteamiento permite presentar el prototipo de acuerdo con la evidencia disponible: como una arquitectura técnica reproducible y clínicamente delimitada, con una ruta explícita para comprobar su desempeño en el escenario domiciliario antes de atribuirle efectos terapéuticos.

Reconocimientos y declaraciones

Los autores agradecen a las comunidades científicas que mantienen bases de datos abiertas de audio respiratorio y ambiental y a las comunidades de software libre utilizadas durante el desarrollo experimental.

Declaración sobre uso de herramientas de IA generativa

El manuscrito fue elaborado íntegramente por los autores, quienes asumieron la responsabilidad total del trabajo sin recurrir a herramientas de inteligencia artificial generativa para su preparación o redacción. Para garantizar la integridad intelectual y el rigor metodológico, la selección bibliográfica, la revisión crítica y la verificación de contenidos se realizaron de forma manual y autónoma, manteniendo un control absoluto sobre el producto final.

Contribución de autores

Los autores concibieron y desarrollaron el trabajo, revisaron el manuscrito y asumen la responsabilidad íntegra de su contenido. La metodología, los resultados, las referencias, la interpretación y las conclusiones fueron revisados y aprobados por los autores.

Consideraciones éticas

El trabajo presentado corresponde al diseño y evaluación técnica de un prototipo entrenado con bases de datos públicas. No se realizó reclutamiento de pacientes, intervención clínica ni recolección de datos personales de salud. Cualquier evaluación futura con participantes humanos requerirá aprobación ética, consentimiento informado, criterios de inclusión y exclusión, protocolo de seguridad, manejo de eventos adversos y política de protección de datos.

Conflicto de intereses

Los autores declaran no tener conflictos de intereses relacionados con el presente trabajo.

Disponibilidad de datos y código

Los conjuntos de datos públicos utilizados corresponden a Coswara y UrbanSound8K, citados en las referencias. El *pipeline* y la aplicación se encuentran en los repositorios públicos Pulneo RAD y pulneo_app (Valladarez Muñoz, 2026a, 2026b). La ejecución genera artefactos de auditoría del conjunto de datos y de la partición, métricas Keras/TFLite y una comparación FP32 e INT8, modelos TFLite y el contrato normalizer_mfcc.json.

Referencias

- Abadade, Y., Benamar, N., Bagaa, M., & Chaoui, H. (2024). Empowering healthcare: TinyML for precise lung disease classification. *Future Internet*, 16(11), Article 391. <https://doi.org/10.3390/fi16110391>
- Bott, J., Blumenthal, S., Buxton, M., Ellum, S., Falconer, C., Garrod, R., Harvey, A., Hughes, T., Lincoln, M., Mikelsons, C., Potter, C., Pryor, J., Rimington, L., Sinfield, F., Thompson, C., Vaughn, P., & White, J. (2009). Guidelines for the physiotherapy management of the adult, medical, spontaneously breathing patient. *Thorax*, 64(Suppl 1), i1–i51. <https://doi.org/10.1136/thx.2008.110726>
- Brown, C., Chauhan, J., Grammenos, A., Han, J., Hasthanasombat, A., Spathis, D., Xia, T., Cicuta, P., & Mascolo, C. (2020). Exploring automatic diagnosis of COVID-19 from crowdsourced respiratory sound data. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 3474–3484). Association for Computing Machinery. <https://doi.org/10.1145/3394486.3412865>
- Chen, J., Yang, T., He, Q., Pang, M., Cao, Y., Liu, Z., Li, L., Liu, H.-I., & Bullen, C. (2025). The impact of gamified interventions on the management of chronic obstructive pulmonary disease: Systematic literature review. *JMIR Serious Games*, 13, Article e69510. <https://doi.org/10.2196/69510>
- Cox, N. S., Dal Corso, S., Hansen, H., McDonald, C. F., Hill, C. J., Zanaboni, P., Alison, J. A., O'Halloran, P., Macdonald, H., & Holland, A. E. (2021). Telerehabilitation for chronic respiratory disease. *Cochrane Database of Systematic Reviews*, 2021(1), Article CD013040. <https://doi.org/10.1002/14651858.CD013040.pub2>
- Davis, S., & Mermelstein, P. (1980). Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 28(4), 357–366. <https://doi.org/10.1109/TASSP.1980.1163420>
- Frownfelter, D., & Dean, E. (2012). *Cardiovascular and pulmonary physical therapy: Evidence to practice* (5th ed.). Elsevier.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Holland, A. E., Mahal, A., Hill, C. J., Lee, A. L., Burge, A. T., Cox, N. S., Moore, R., Nicolson, C., O'Halloran, P., Lahham, A., Gillies, R., & McDonald, C. F. (2017). Home-based rehabilitation for COPD using minimal resources: A randomised, controlled equivalence trial. *Thorax*, 72(1), 57–65. <https://doi.org/10.1136/thoraxjnl-2016-208514>
- Larson, E. C., Goel, M., Boriello, G., Heltshe, S., Rosenfeld, M., & Patel, S. N. (2012). SpiroSmart: Using a microphone to measure lung function on a mobile phone. In *Proceedings of the 2012 ACM conference on ubiquitous computing* (pp. 280–289). Association for Computing Machinery. <https://doi.org/10.1145/2370216.2370261>
- McCarthy, B., Casey, D., Devane, D., Murphy, K., Murphy, E., & Lacasse, Y. (2015). Pulmonary rehabilitation for chronic obstructive pulmonary disease. *Cochrane Database of Systematic Reviews*, 2015(2), Article CD003793. <https://doi.org/10.1002/14651858.CD003793.pub3>
- Osadnik, C. R., McDonald, C. F., Jones, A. P., & Holland, A. E. (2012). Airway clearance techniques for chronic obstructive pulmonary disease. *Cochrane Database of Systematic Reviews*, 2012(3), Article CD008328. <https://doi.org/10.1002/14651858.CD008328.pub2>
- Oppenheim, A. V., & Schaffer, R. W. (2010). *Discrete-time signal processing* (3rd ed.). Pearson.

- Pryor, J. A., & Prasad, S. A. (2008). *Physiotherapy for respiratory and cardiac problems: Adults and paediatrics* (4th ed.). Churchill Livingstone.
- Rabiner, L. R., & Schafer, R. W. (2011). *Theory and applications of digital speech processing*. Pearson.
- Ravi, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., & Yang, G.-Z. (2017). Deep learning for health informatics. *IEEE Journal of Biomedical and Health Informatics*, 21(1), 4–21. <https://doi.org/10.1109/JBHI.2016.2636665>
- Rivas-Navarrete, J. A., Pérez-Espinosa, H., Padilla-Ortiz, A. L., Rodríguez-González, A. Y., & García-Camero, D. C. (2025). Edge computing system for automatic detection of chronic respiratory diseases using audio analysis. *Journal of Medical Systems*, 49, Article 33. <https://doi.org/10.1007/s10916-025-02154-7>
- Rocha, B. M., Filos, D., Mendes, L., Serbes, G., Ulukaya, S., Kahya, Y. P., Jakovljevic, N., Turukalo, T. L., Vogiatzis, I. M., Perantoni, E., Kaimakamis, E., Natsiavas, P., Oliveira, A., Jácome, C., Marques, A., Maglaveras, N., Paiva, R. P., Chouvarda, I., & de Carvalho, P. (2019). An open access database for the evaluation of respiratory sound classification algorithms. *Physiological Measurement*, 40(3), Article 035001. <https://doi.org/10.1088/1361-6579/ab03ea>
- Salamon, J., Jacoby, C., & Bello, J. P. (2014). A dataset and taxonomy for urban sound research. In *Proceedings of the 22nd ACM international conference on multimedia* (pp. 1041–1044). Association for Computing Machinery. <https://doi.org/10.1145/2647868.2655045>
- Sharma, N., Krishnan, P., Kumar, R., Ramoji, S., Chetupalli, S. R., Nirmala, R., Ghosh, P. K., & Ganapathy, S. (2020). Coswara—A database of breathing, cough, and voice sounds for COVID-19 diagnosis. In *Interspeech 2020* (pp. 4811–4815). International Speech Communication Association. <https://doi.org/10.21437/Interspeech.2020-2768>
- Shih, C.-H., Tomita, N., Lukic, Y. X., Hernández Reguera, Á., Fleisch, E., & Kowatsch, T. (2019). Breeze: Smartphone-based acoustic real-time detection of breathing phases for a gamified biofeedback breathing training. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 3(4), Article 152. <https://doi.org/10.1145/3369835>
- Spruit, M. A., Singh, S. J., Garvey, C., ZuWallack, R., Nici, L., Rochester, C., Hill, K., Holland, A. E., Lareau, S. C., Man, W. D. C., Pitta, F., Sewell, L., Raskin, J., Bourbeau, J., Crouch, R., Franssen, F. M. E., Casaburi, R., Vercoulen, J. H., Vogiatzis, I., ... Wouters, E. F. M. (2013). An official American Thoracic Society/European Respiratory Society statement: Key concepts and advances in pulmonary rehabilitation. *American Journal of Respiratory and Critical Care Medicine*, 188(8), e13–e64. <https://doi.org/10.1164/rccm.201309-1634ST>
- Valladarez Muñoz, D. F. (2026a). *Pulneo: Pipeline RAD para respiración* [Código fuente]. GitHub. <https://github.com/DavidValladarez95/pulneo>
- Valladarez Muñoz, D. F. (2026b). *pulneo_app: Aplicación móvil gamificada para fisioterapia respiratoria domiciliaria* [Código fuente]. GitHub. https://github.com/DavidValladarez95/pulneo_app
- Warden, P., & Situnayake, D. (2019). *TinyML: Machine learning with TensorFlow Lite on Arduino and ultra-low-power microcontrollers*. O'Reilly Media.
- Watson, J. S., Jordan, R. E., Gardiner, L., Adab, P., & Jolly, K. (2023). A systematic review of the effectiveness of interventions to promote referral, adherence, and uptake of pulmonary rehabilitation for patients with chronic obstructive pulmonary disease. *International Journal of Chronic Obstructive Pulmonary Disease*, 18, 1637–1654. <https://doi.org/10.2147/COPD.S396317>
- Zhantleuova, A. K., Makashev, Y. K., & Duzbayev, N. T. (2025). Optimizing MFCC parameters for breathing phase detection. *Sensors*, 25(16), Article 5002. <https://doi.org/10.3390/s25165002>