

Estimación del Número de Mentas en una Caja Mediante Análisis de Sonido y Aprendizaje Automático

Sánchez María, Sánchez Mateo, Rodriguez Oscar

March 6, 2025

Índice General

Contents

1	Introducción	2
2	Planteamiento del problema	2
3	Objetivos	3
3.1	Objetivos Generales	3
3.2	Objetivos Específicos	3
4	Metodología	3
4.1	Recolección de Datos	3
4.2	Procesamiento de Audio	3
4.3	Entrenamiento del Modelo	4
4.4	Cronograma	4
5	Resultados Esperados	4
6	Conclusión	4

Abstract

Este proyecto propone un método para estimar la cantidad de mentas dentro de una caja utilizando el sonido generado al agitarla. Se entrenará una red neuronal para diferenciar los sonidos desde que hay una menta hasta que la caja está llena. La recolección de datos se realizará grabando el sonido producido al agitar la caja diez veces en cinco segundos. Posteriormente, se aplicará la Transformada Rápida de Fourier (FFT) para extraer la frecuencia fundamental de cada muestra. Además, se apilarán los picos de frecuencia obtenidos para eliminar el ruido y mejorar la calidad de los datos. Los datos procesados se utilizarán para entrenar un modelo de aprendizaje automático en Python.

1 Introducción

Determinar la cantidad de objetos dentro de un recipiente sin abrirlo es un problema con diversas aplicaciones en logística y manufactura. En este estudio, se explorará el uso de redes neuronales para analizar los patrones de sonido generados por una caja de metal al ser agitada con distintas cantidades de mentas en su interior.

2 Planteamiento del problema

En diversas situaciones, la estimación de la cantidad de objetos dentro de un contenedor cerrado es un desafío cuando no es posible inspeccionarlos visualmente. Métodos tradicionales como el pesaje pueden ser imprecisos si el contenido varía en densidad o si se desconoce el peso de cada elemento. En este contexto, el sonido generado al agitar un recipiente ofrece una alternativa potencial para inferir la cantidad de objetos en su interior, dado que las características acústicas de la señal dependen de la cantidad y distribución de los elementos.

Este proyecto busca determinar si es posible estimar la cantidad de mentas en una caja a partir del sonido producido al agitarla. Para ello, se explorará el uso de técnicas de procesamiento de señales, específicamente la Transformada Rápida de Fourier (FFT), para extraer información relevante de los sonidos registrados. Además, se emplearán modelos de aprendizaje automático, particularmente redes neuronales, para reconocer patrones en las señales acústicas y realizar estimaciones precisas.

El problema central radica en evaluar la viabilidad y precisión de este enfoque, así como en definir las condiciones óptimas para su implementación. Se espera que esta investigación aporte conocimientos aplicables a otros contextos donde la detección de cantidad a partir de señales acústicas pueda ser útil, como en control de inventarios, monitoreo de envases sellados o incluso en el ámbito biomédico.

3 Objetivos

3.1 Objetivos Generales

Desarrollar un método basado en el análisis de sonido y aprendizaje automático para estimar la cantidad de mentas en una caja plástica, utilizando patrones acústicos como fuente de información.

3.2 Objetivos Específicos

- Diseñar y ejecutar un experimento de recolección de datos mediante la grabación de sonidos de una caja con diferentes cantidades de mentas.
- Procesar los datos obtenidos aplicando la Transformada Rápida de Fourier (FFT) para extraer características relevantes.
- Implementar y entrenar una red neuronal capaz de predecir la cantidad de mentas en la caja a partir de los datos procesados.
- Evaluar el desempeño del modelo mediante métricas de error y optimizar su rendimiento.

4 Metodología

4.1 Recolección de Datos

El proceso de recolección de datos consta de los siguientes pasos:

- Se utilizará una caja de (Tic Tac) como recipiente.
- Cada grabación tendrá una duración de 4 segundos: el primer segundo en silencio y los siguientes tres segundos con 5 repeticiones del movimiento de sacudida.
- Se grabará el sonido producido al agitar la caja con diferentes cantidades de mentas, desde 1 hasta 29, repitiendo el experimento 30 veces por cada cantidad.

4.2 Procesamiento de Audio

Cada grabación será procesada mediante la Transformada Rápida de Fourier (FFT) para extraer las características de frecuencia más relevantes. Se analizarán las frecuencias naturales del sistema y se apilarán los picos obtenidos para eliminar el ruido y obtener una representación más limpia de los patrones sonoros característicos de cada cantidad de mentas.

4.3 Entrenamiento del Modelo

Los datos procesados se ingresarán a un modelo de aprendizaje automático implementado en Python. Se entrenará una red neuronal para correlacionar los patrones de sonido con la cantidad de mentas dentro de la caja.

4.4 Cronograma

Actividad	Tiempo (Meses)		
	Marzo	Abril	Mayo
1. Selección de caja y mentas			
2. Diseño del protocolo de grabación			
3. Grabación de sonidos			
4. Almacenamiento y organización de datos			
5. Aplicación de FFT a los datos			
6. Identificación de frecuencias clave			
7. Filtrado y normalización de datos			
8. Selección del tipo de red neuronal			
9. Implementación del modelo de red neuronal			
10. Entrenamiento del modelo			
11. Ajuste de hiperparámetros			
12. Definición de métricas de evaluación			
13. Pruebas con nuevas grabaciones			
1.4 Elaboración del informe final			

5 Resultados Esperados

Se espera que la red neuronal pueda identificar con precisión la cantidad de mentas en la caja a partir del sonido producido al agitarla, incluso en presencia de ruido.

6 Conclusión

Este estudio propone una solución innovadora basada en procesamiento de señales y aprendizaje automático para estimar la cantidad de objetos en un contenedor cerrado. Si tiene éxito, podría aplicarse a otros escenarios donde la inspección visual no sea viable.

References

- [1] Ronan Collobert, Koray Kavukcuoglu, and Clément Farabet. Torch7: A matlab-like environment for machine learning. In *BigLearn, NIPS workshop*, 2011.
- [2] Hugo Fastl and Eberhard Zwicker. *Psychoacoustics: Facts and Models*, volume 22 of *Springer Series in Information Sciences*. Springer, 3rd edition, 2007.
- [3] James F O’Brien, Perry R Cook, and Georg Essl. Synthesizing sounds from physically based motion. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 529–536, 2001.
- [4] Zhoutong Zhang, Qiujia Li, Zhengjia Huang, Jiajun Wu, Josh Tenenbaum, and Bill Freeman. Shape and material from sound. *Advances in Neural Information Processing Systems*, 30, 2017.