**Logotipo

Descripción generada automáticamente**

**Desarrollo de un sistema de diagnóstico complementario para apendicitis pediátrica utilizando redes neuronales convolucionales: Un análisis de datos clínicos del Hospital St. Hedwig**

Alejandro Martínez Hernández

Avance de Monografía

Asesor  
David Manuel Villanueva Valdés, Magister

Universidad de Antioquia  
Facultad de Ingeniería

Especialización en Analítica y Ciencia de Datos

Medellín, Antioquia, Colombia

2023

|  |  |
| --- | --- |
| **Cita** | (Martínez Hernández, 2023) |
| **Referencia**  **Estilo APA 7 (2020)** | Martínez Hernández, A. (2023). *Desarrollo de un sistema de diagnóstico complementario para apendicitis pediátrica utilizando redes neuronales convolucionales: Un análisis de datos clínicos del Hospital St. Hedwig.* [Trabajo de grado especialización]. Universidad de Antioquia, Medellín, Colombia. |

**** 

Especialización en Analítica y Ciencia de Datos, CohorteVI.

Centro de Investigación Ambientales y de Ingeniería (CIA).

|  |  |
| --- | --- |
|  | Diagrama  Descripción generada automáticamente con confianza media |

Centro de Documentación Ingeniería (CENDOI)

**Repositorio Institucional:** http://bibliotecadigital.udea.edu.co

Universidad de Antioquia - www.udea.edu.co

Rector: John Jairo Arboleda Céspedes.

Decano: Julio Cesar Saldarriaga Molina

Jefe departamento: Diego José Luis Botia Valderrama

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Antioquia ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

**Tabla de contenido**

[Resumen 6](#_Toc151246594)

[Abstract 7](#_Toc151246595)

[1. Descripción del problema 8](#_Toc151246596)

[1.1. Problema de negocio 8](#_Toc151246597)

[1.2. Aproximación desde la analítica de datos 8](#_Toc151246598)

[1.3. Origen de los datos 9](#_Toc151246599)

[1.4. Métricas de desempeño 9](#_Toc151246600)

[2. Objetivos 11](#_Toc151246601)

[2.1. Objetivo general 11](#_Toc151246602)

[2.2. Objetivos específicos 11](#_Toc151246603)

[3. Datos 12](#_Toc151246604)

[3.1. Datos originales 12](#_Toc151246605)

[3.2. Datasets 12](#_Toc151246606)

[3.3. Analítica descriptiva 15](#_Toc151246607)

[4. Referencias 24](#_Toc151246608)

**Lista de figuras**

**Figura 1** *Imagen 1.4 del dataset original donde se ve una ecografía hecha con un transductor lineal***16**

**Figura 2***. Imagen 74.3 del dataset original donde se ve una ecografía hecha con un transductor convexo***17**

**Figura 3***. Imagen 1.1 del dataset original donde se ve una ecografía hecha con un transductor lineal y varias marcas de medición azules***18**

**Figura 4***. Imagen 74.2 del dataset original donde se ve una ecografía hecha con un transductor lineal y varias marcas de colores, un rectángulo verde y una línea de medición azul***18**

**Figura 5***. Imagen 14.1 del dataset original donde se ven 6 ecografías hechas con un transductor lineal.***19**

**Figura 6***. Imagen 619.2 del dataset original donde se ven 6 ecografías hechas con un transductor lineal*19

**Figura 7***. Muestra del set de entrenamiento para el modelo. Ecografías recortadas cuadradas y transformadas*20

**Figura 8***. Distribución de clases de las imágenes según su tipo antes de aplicar un balanceo***20**

**Figura 9***. Distribución de clases de las imágenes según su tipo después de aplicar un balanceo.* **21**

**Figura 10***. Distribución de pacientes diagnosticados con y sin apendicitis antes del balanceo***21**

**Figura 11***. Distribución de pacientes diagnosticados con y sin apendicitis después del balanceo* **22**

**Figura 12***. Distribución de edades de pacientes antes del balanceo***22**

**Figura 13***. Distribución de edades de pacientes después del balanceo***23**

# Resumen

En el proyecto propuesto, se diseñará un sistema de diagnóstico médico complementario para la detección de apendicitis pediátrica, utilizando redes neuronales convolucionales y análisis de datos estructurados. Se dispondrá de una base de datos de ecografías del abdomen del Hospital St. Hedwig, enriquecida con anotaciones clínicas. Se empleará la técnica de transferencia de aprendizaje con modelos avanzados como MobileNetV2, ResNet-50, Inception y VGG16, ajustando la última capa para tres categorías diagnósticas. Se explorarán algoritmos para datos estructurados y se evaluará si la fusión de múltiples modelos mejora la precisión del diagnóstico.

*Palabras clave*: Apendicitis pediátrica, ecografía abdominal, redes neuronales convolucionales, transferencia de aprendizaje.

**Repositorio:** <https://github.com/Dezarti/monografia_individual>

# Abstract

In the proposed project, a complementary medical diagnostic system for pediatric appendicitis detection will be designed, using convolutional neural networks and analysis of structured data. A database of abdominal ultrasounds from St. Hedwig Hospital, enriched with clinical annotations, will be available. The technique of transfer learning with advanced models such as MobileNetV2, ResNet-50, Inception, and VGG16 will be employed, adjusting the final layer for three diagnostic categories. Algorithms for structured data will be explored, and whether the fusion of multiple models improves diagnostic accuracy will be evaluated.

*Keywords***:** Pediatric appendicitis, abdominal ultrasound, convolutional neural networks, transfer learning.

# Descripción del problema

## Problema de negocio

A pesar de los avances en las metodologías de diagnóstico de apendicitis, que comprenden el uso de sistemas de puntaje y análisis de imágenes (Rentea et al., 2017), en este último mencionado la interpretación de las ecografías (un método estándar de bajo costo) puede ser altamente dependiente de la experiencia del radiólogo. Esta situación se complica con la considerable variación y limitación en la precisión de la evaluación, con índices de errores diagnósticos que oscilan entre el 28% y el 57% en niños de 2 a 12 años, y cercanos al 100% en infantes menores a ese rango (Almaramhy, 2017; Haxhirexha et al., 2018).

Esta situación no solo afecta a los pacientes, que aún en países desarrollados pueden enfrentar esperas mayores a 5 horas en recibir tratamiento adecuado (Abbas et al., 2018), sino también a hospitales, clínicas y centros de diagnóstico que en su búsqueda por brindar soluciones rápidas y acertadas que puedan apoyar y optimizar los diagnósticos médicos, luchan contra una creciente demanda en la solicitud de este tipo de servicios y un déficit en especialistas en imágenes diagnósticas que pueda suplirlo (Gómez-Mendoza, 2010).

## Aproximación desde la analítica de datos

Dentro del proceso de diagnóstico médico complementario se propone el desarrollo de un modelo de inteligencia artificial basado en redes neuronales convolucionales pre-entrenadas, cuyas últimas capas serán entrenadas con datos de pacientes pediátricos, incluyendo ecografías y síntomas clínicos y cuya salida sea la probabilidad que el paciente posea dicha patología con base en los datos de entrada. Este modelo predictivo estará diseñado para identificar patrones y características en las imágenes médicas que indiquen la presencia o ausencia de apendicitis, aumentando así la precisión en el diagnóstico. En casos de precisión subóptima, se ajustará la arquitectura del modelo y se enriquecerá la calidad de los datos.

En situaciones donde los recursos médicos son limitados o los radiólogos están sobrecargados, estos modelos proporcionan una herramienta valiosa de apoyo a la decisión, permitiendo a los profesionales médicos con menos experiencia realizar diagnósticos más precisos y rápidos, mejorando así la eficiencia y efectividad del tratamiento y contribuyendo directamente a una mejor gestión en entornos con alta demanda de servicios de diagnóstico por imágenes.

## Origen de los datos

Se hará uso de una base de datos pública proporcionada por el Hospital St. Hedwig en Regensburg, Alemania. Esta base de datos fue creada a partir de un estudio que recopiló y organizó información sobre una cohorte de pacientes pediátricos que presentaban dolor abdominal entre 2016 a 2021.

Es importante aclarar que la información fue validada por el personal del hospital alemán anteriormente mencionado; asimismo, el estudio que publica los datos fue aprobado por el comité de ética de la Universidad de Regensburg, se llevó a cabo según las regulaciones y pautas de dicha universidad y fue finalmente publicado a través del repositorio de datos de acceso abierto respaldado por el CERN, Zenodo, bajo la licencia de Creative Commons para usos no comerciales.

Según lo anterior, se estima que los datos son confiables, además de ser la única base de datos pública de este tipo que se pudo encontrar; sin embargo, en una exploración exhaustiva sobre el dataset se encontraron ciertas fallas y/o falencias que se explicaran a fondo más adelante.

## Métricas de desempeño

Hay muchas métricas de negocio que se podrían evaluar, sin embargo, la gran mayoría de ellas implica necesariamente la evaluación del sistema de diagnóstico en un entorno clínico, saliéndose del alcance de esta monografía. Es por ello por lo que solo se evaluará una métrica de la cual se tengan datos para comparar.

* **Exactitud (Accuracy):** Mide el porcentaje total de predicciones correctas. Un valor mínimo del 80% es deseable, pues indicaría una mejora frente a los métodos tradicionales cuyo rango de exactitud en ciertos pacientes pediátricos ronda de 43-72% (Almaramhy, 2017; Haxhirexha et al., 2018).
* **Sensibilidad (Sensitivity):** Crucial para minimizar los falsos negativos. Un valor mínimo del 85% es recomendable para garantizar que la mayoría de los casos reales de apendicitis sean identificados, aunque estará sujeto a modificación según los resultados de la exactitud.
* **Especificidad (Specificity):** Importante para reducir los falsos positivos. Un mínimo del 85% asegura que los casos no apendicitis sean correctamente excluidos. Igualmente estará sujeto a modificación según los resultados de la exactitud.
* **Tiempo de Procesamiento:** En análisis de imágenes y síntomas podría llegar a ser un factor decisivo, pues para mejorar el proceso este debe ser ágil. Por las capacidades de procesamiento actual, se espera que está métrica no sea mayor a 5 minutos.

# Objetivos

## Objetivo general

Diseñar un proceso de diagnóstico médico complementario para el soporte en la detección de apendicitis pediátrica, utilizando redes neuronales convolucionales entrenadas a partir de una base de datos de ecografías de abdomen del Hospital St. Hedwig en Regensburg, Alemania, correspondiente al período de 2016 a 2021.

## Objetivos específicos

* Identificar, limpiar y etiquetar la información relevante de la base de datos del Hospital St. Hedwig para el entrenamiento y evaluación de modelos de IA en la detección de apendicitis.
* Diseñar, entrenar y evaluar una red neuronal convolucional, para identificar signos de apendicitis pediátrica a partir de datos de ecografías de abdomen.
* Establecer protocolos para integrar las predicciones del modelo de IA como una herramienta complementaria en el proceso de diagnóstico clínico de apendicitis pediátrica.

# Datos

## Datos originales

En este dataset se encuentran varios archivos en diferentes formatos que según sus creadores, Marcinkevičs et al. (2023), corresponden a imágenes de ecografía abdominal en modo B que “muestran diversas regiones de interés, como el cuadrante inferior derecho del abdomen, el apéndice, los intestinos, los ganglios linfáticos y los órganos reproductores”, además de conjuntos de datos que “incluyen información sobre pruebas de laboratorio, resultados de la exploración física, puntuaciones clínicas, como las puntuaciones de Alvarado y de apendicitis pediátrica, y resultados ecográficos elaborados por expertos”.

Nuevamente, citando directamente a Marcinkevičs et al. (2023), el dataset posee la siguiente estructura:

* “US\_Pictures/: carpeta con las imágenes originales de ecografía en modo B en formato BMP; las imágenes se denominan como <subject #>.<view #> \*.bmp”. Se poseen 2092 sonogramas de 780 pacientes.
* “app\_data.xlsx: Archivo MS Excel con datos tabulares (la pestaña ‘Resumen de datos’ contiene una explicación de las variables); los números de asunto correspondientes de la carpeta US\_Pictures/ se encuentran en la columna US\_Number”
* “multiple\_in\_one\_: una lista de nombres de imágenes de ultrasonido que contienen varias instantáneas”
* “test\_set\_codes\_csv: una lista de puntos de datos del conjunto de prueba”

## Datasets

Recordando que para el procesamiento de imágenes se utilizarán redes neuronales convolucionales, es importante tener en cuenta que estos modelos comienzan con parámetros aleatorios que se ajustan durante el entrenamiento. Dado que se dispone de un conjunto limitado de datos, se empleará la técnica de 'transferencia de aprendizaje', usando modelos como MobileNetV2, ResNet-50, Inception (GoogleNet) y VGG16. Se eliminará la capa final del modelo será reemplazada por una nueva con tres categorías ('sin apéndice', 'apendicitis', 'no apendicitis'). Inicialmente, solo se entrenarán las últimas capas del modelo, y posteriormente, se ajustarán todas las capas con una tasa de aprendizaje reducida.

Para los datos estructurados se puede desarrollar un modelo convencional haciendo uso de técnicas como Random Forest, maquina de vectores de soporte (SVM), K vecinos más cercanos (KNN) u otros sugeridos por la literatura en este campo como gradient boosting (XGBoost o LightGBM) (Rufo et al., 2021)

Es importante señalar que, antes de finalizar este proyecto, no se puede determinar con certeza cuál sería el enfoque más efectivo, ya sea seleccionar un único modelo del grupo entrenado o aplicar una fusión de modelos. Esta última ha mostrado resultados prometedores en algunos estudios (Park et al., 2021), se sugieren técnicas como el voto mayoritario, que promedia las predicciones de múltiples modelos, o el apilamiento, que emplea las salidas de varios modelos como entradas para un meta-modelo. Se explorarán más a fondo estas metodologías en las próximas secciones del trabajo.

* + 1. **Procesamiento de imágenes.**

Las imágenes de ultrasonido ubicadas en ‘US\_Pictures/’ son capturas de pantalla del software de visualización. Estas son generalmente creadas por radiólogos para realizar anotaciones diagnósticas. Aunque todas las imágenes son útiles, algunos elementos del software en las capturas pueden ser irrelevantes para el entrenamiento del modelo, pudiendo incluso ocultar detalles cruciales. Por ello, se utilizó OpenCV (cv2) para detectar y recortar los bordes relevantes de cada imagen. Posteriormente, se clasificaron según su nombre y la información de la base de datos estructurada, organizándolas en diferentes carpetas.

La metodología consta de cuatro pasos:

* 1. **Selección:** Según la información ‘app\_data.xlsx’, se organiza todo el conjunto de imágenes según sus características en 3 carpetas:
     1. ***con\_info:*** Guarda las imágenes que si tienen diagnóstico
     2. ***no\_apendice:*** Guarda las imágenes donde no es posible ver el apéndice
     3. ***no\_paciente:*** No se poseen los datos necesarios para clasificar estas imágenes. Son descartadas ya que no tienen utilidad en el proceso y facilita el procesamiento de las siguientes imágenes.

* 1. **Recorte:** Toma imágenes de las subcarpetas 'con\_info' y 'no\_apendice', aplicando la identificación de bordes mediante OpenCV (cv2). Se establecen criterios adicionales para asegurar la calidad del recorte: ningún borde debe ser menor a 250 píxeles y la relación entre el borde más largo y el más corto no debe exceder 1.5 veces. Estos criterios garantizan que las imágenes recortadas contengan la máxima información posible de la ecografía. Los resultados se mantienen ordenados según la clasificación inicial."
  2. **Ordenar:** Para el entrenamiento del modelo, se establecen tres categorías clave: imágenes donde no es visible el apéndice (y por tanto, no se puede emitir un diagnóstico), imágenes con diagnóstico positivo de apendicitis, y aquellas con diagnóstico negativo. Esta clasificación tridimensional es esencial para entrenar al modelo de manera efectiva, permitiéndole distinguir entre estos escenarios clínicos distintos.

Las imágenes procesadas en el paso anterior se ordenan según su clasificación como:

* + 1. ***apendicitis:*** Guarda las imágenes recortadas donde se ve el apéndice y el diagnóstico fue apendicitis.
    2. ***no\_apendicitis:*** Guarda las imágenes recortadas donde se ve el apéndice y el diagnóstico fue NO apendicitis.
    3. ***no\_apendice:*** Guarda directamente las imágenes recortadas que sabemos que no muestra el apéndice.
  1. **Balanceo de imágenes:** se optó por eliminar aleatoriamente instancias de la clase mayoritaria (Undersampling), pues, aunque se cuenta con pocos en cantidad, otras metodologías podrían representar un deterioro en la calidad de los datos.
  2. **Transformación:** Considerando la diversidad en la colección de imágenes, que incluye variaciones en anotaciones de radiólogos, orientación y tamaño, se implementaron técnicas de aumento de datos para mitigar el riesgo de sobreajuste. Este proceso involucra la normalización de los valores de píxeles (ajustándolos entre 0 y 1), y la aplicación aleatoria de transformaciones como rotaciones, traslaciones horizontales y verticales, recortes y zoom. Estas técnicas enriquecen el conjunto de datos, proporcionando una representación más amplia y variada de posibles escenarios clínicos.
  3. **Entrenamiento y validación:** De los datos transformados, se usará un 80% para entrenamiento y un 20% para validación.
     1. **Procesamiento de datos estructurados.**
  4. **Valores Faltantes:** Se detectó la presencia de valores faltantes en varias columnas, lo que requiere una estrategia de imputación o eliminación.
  5. **Variables Categóricas:** Se identificaron columnas categóricas que necesitarán ser convertidas a un formato numérico para su procesamiento por modelos de IA.
  6. **Datos Numéricos:** Las columnas numéricas serán normalizadas o estandarizadas para asegurar la consistencia en la escala de los datos.
  7. **Análisis Exploratorio:** Las estadísticas descriptivas revelan la distribución general y las características de los datos numéricos. Aquí junto a la ayuda de la literatura se seleccionan las siguientes columnas, pues reflejan factores clínicamente relevantes para el diagnóstico de apendicitis:
     1. **Edad y BMI:** Indicadores importantes del estado general de salud del paciente.
     2. **Sexo:** Puede influir en la presentación de síntomas.
     3. **Altura y Peso:** Datos básicos de salud.
     4. **Diagnóstico:** Variable objetivo para la clasificación.
     5. **Puntuación de Alvarado:** Un índice clínico utilizado para evaluar la probabilidad de apendicitis.
     6. **Temperatura Corporal y WBC Count:** Indicadores de inflamación o infección.
  8. **Balanceo de los datos:** Se balancean los datos de tal manera que el número de entradas para pacientes diagnosticados con apendicitis y sin apendicitis sea parecido.

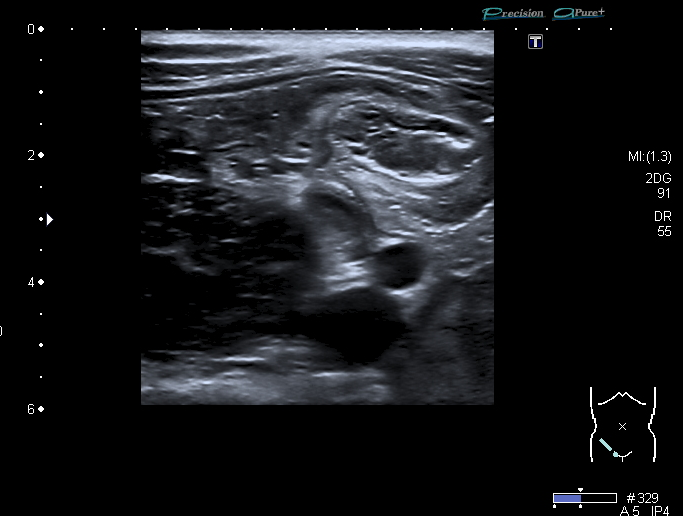
## Analítica descriptiva

* + 1. **Procesamiento de imágenes**

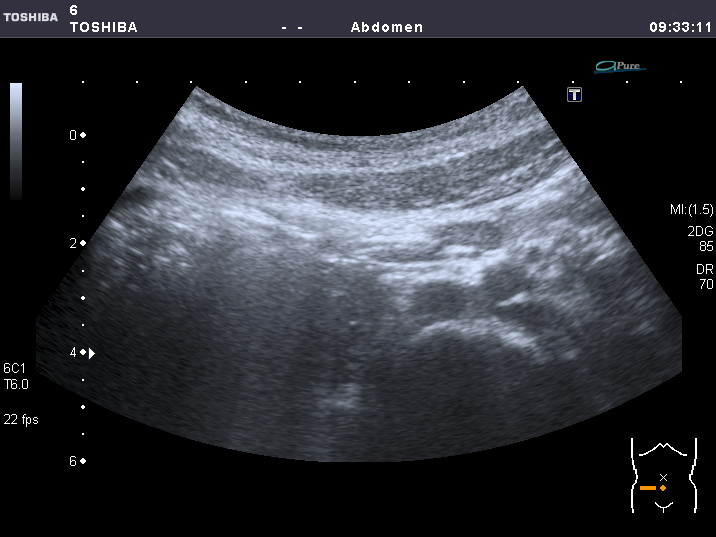
Como se mencionó anteriormente, el conjunto de datos consistía principalmente en capturas de pantalla del software de visualización de ecografías. Estas ecografías fueron realizadas con distintos tipos de transductores, desde lineales (figura 1) hasta convexos (figura 2). Además, se enfrentaron desafíos relacionados con la orientación de las imágenes (transversales y longitudinales) y, en varios casos, con una incorrecta etiquetación, ya sea por no seguir las guías de nomenclatura establecidas o por contener datos erróneos. Asimismo, se observaron variadas anotaciones por parte de los radiólogos (figura 3-4) y algunas imágenes incluían múltiples capturas en una sola (figura 5). Todos estos elementos se homogeneizaron mediante técnicas de recorte descritas previamente (figura 6), buscando estandarizar el conjunto de datos para su análisis.

Con las imágenes transformadas (figura 7) se analizó su distribución de clases entre ‘Apendicitis’, ‘No Apendicitis’ y ‘No Apéndice’ para descubrir un desbalance entre estos (figura 8) el cual fue posteriormente corregido (figura 9).

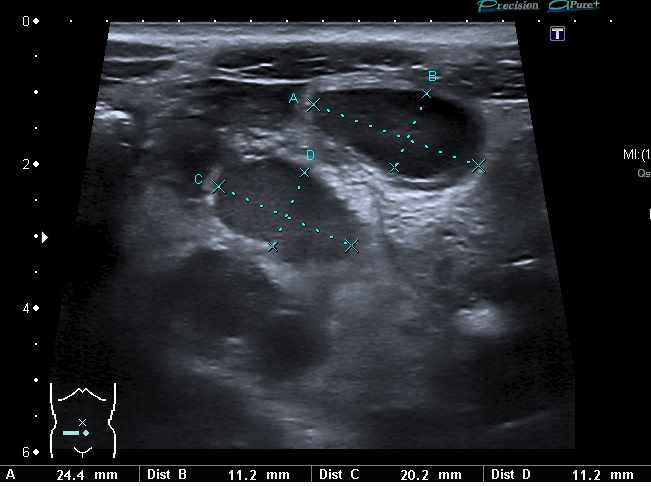
**Figura 1** *Imagen 1.4 del dataset original donde se ve una ecografía hecha con un transductor lineal.*



**Figura 2** *Imagen 74.3 del dataset original donde se ve una ecografía hecha con un transductor convexo.*



**Figura 3** *Imagen 1.1 del dataset original donde se ve una ecografía hecha con un transductor lineal y varias marcas de medición azules.*

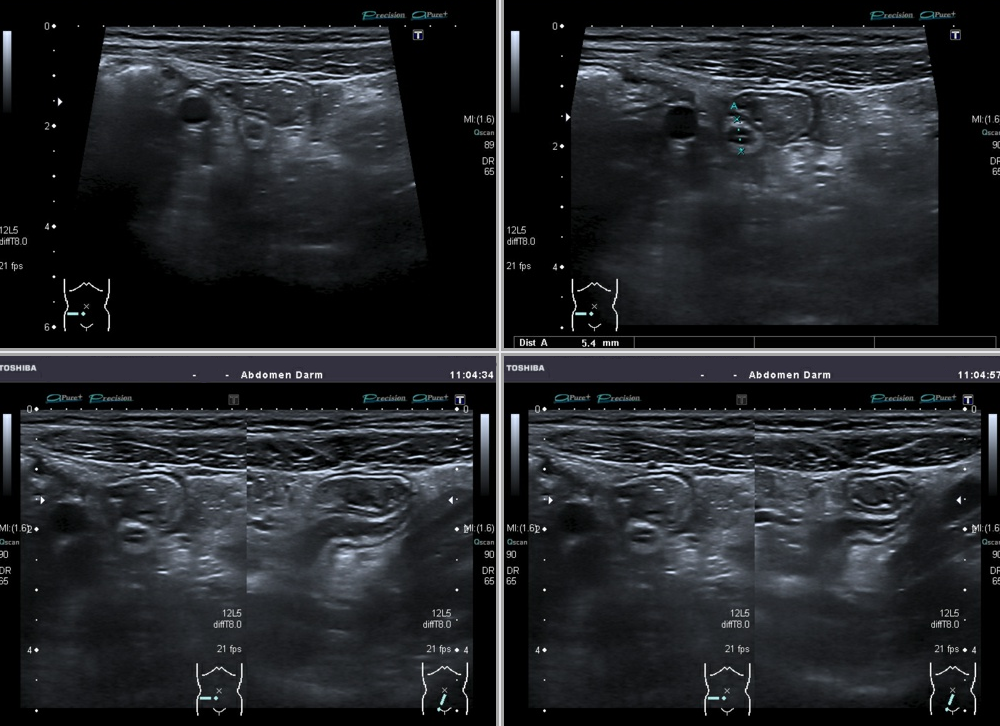


**Figura 4** *Imagen 74.2 del dataset original donde se ve una ecografía hecha con un transductor lineal y varias marcas de colores, un rectángulo verde y una línea de medición azul.*

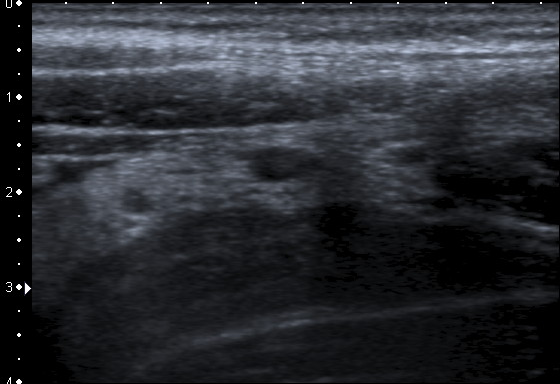
Captura de pantalla de un videojuego

Descripción generada automáticamente

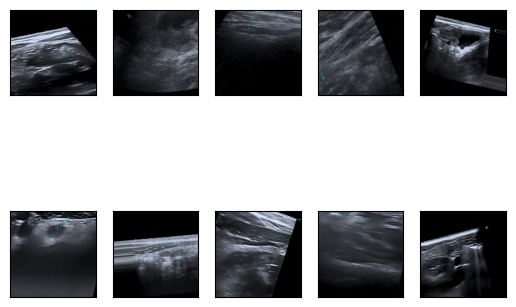
**Figura 5** *Imagen 14.1 del dataset original donde se ven 6 ecografías hechas con un transductor lineal.*



**Figura 6** *Imagen 619.2 del dataset original donde se ven 6 ecografías hechas con un transductor lineal.*



**Figura 7** *Muestra del set de entrenamiento para el modelo. Ecografías recortadas cuadradas y transformadas.*



**Figura 8** *Distribución de clases de las imágenes según su tipo antes de aplicar un balanceo.*

Gráfico, Gráfico de barras

Descripción generada automáticamente

**Figura 9** *Distribución de clases de las imágenes según su tipo después de aplicar un balanceo.*

Gráfico, Gráfico de barras

Descripción generada automáticamente

* + 1. **Procesamiento de datos estructurados**

Después de las transformaciones necesarias se disponen de 7 columnas. Una representación gráfica de dos de estas se puede observar en la figura 10 y 12. Las figuras 11 y 13 corresponden a las representaciones gráficas de esas mismas columnas, pero después del balanceo.

**Figura 10** *Distribución de pacientes diagnosticados con y sin apendicitis antes del balanceo.*

Gráfico, Gráfico de barras

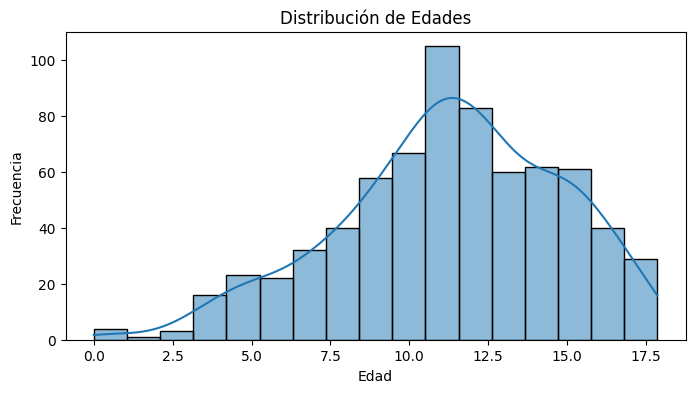
Descripción generada automáticamente

**Figura 11** *Distribución de pacientes diagnosticados con y sin apendicitis después del balanceo*

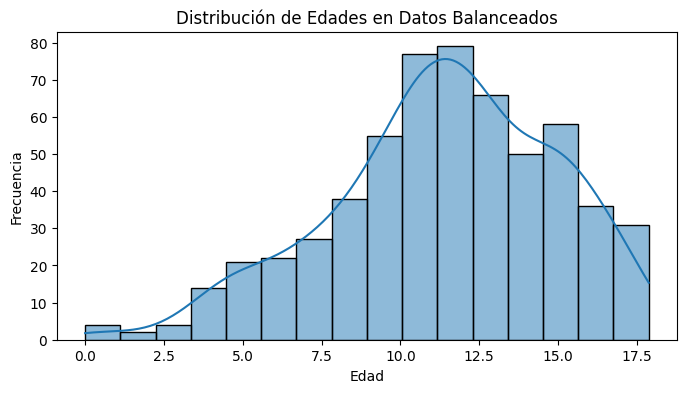
Gráfico, Gráfico de barras

Descripción generada automáticamente

**Figura 12** *Distribución de edades de pacientes antes del balanceo*



**Figura 13** *Distribución de edades de pacientes después del balanceo*



# Referencias

Abbas, P. I., Zamora, I. J., Elder, S. C., Brandt, M. L., Lopez, M. E., Orth, R. C., Bisset, G. S., & Cruz, A. T. (2018). How Long Does it Take to Diagnose Appendicitis? Time Point Process Mapping in the Emergency Department. *Pediatric Emergency Care*, *34*(6), 381–384. https://doi.org/10.1097/PEC.0000000000000720

Almaramhy, H. H. (2017). Acute appendicitis in young children less than 5 years: review article. *Italian Journal of Pediatrics*, *43*(1), 15. https://doi.org/10.1186/s13052-017-0335-2

Gómez-Mendoza, M. A. (2010, May 10). *Colombia, panorama actual del radiodiagnóstico*. Https://Diagnosticojournal.Com/Colombia-Panorama-Actual-Del-Radiodiagnostico/.

Haxhirexha, K., Dogjani, A., Zylbeari, L., & Haxhirexha, F. (2018). Misdiagnosed Appendicitis in Children. *Albanian Journal of Trauma and Emergency Surgery*, *2*(2), 99–104. https://doi.org/10.32391/ajtes.2018.2.2.005

Marcinkevičs, R., Reis Wolfertstetter, P., Klimiene, U., Ozkan, E., Chin-Cheong, K., Paschke, A., Zerres, J., Denzinger, M., Niederberger, D., Wellmann, S., Knorr, C., & Vogt, J. E. (2023). *Regensburg Pediatric Appendicitis Dataset*. Zenodo. https://doi.org/10.5281/zenodo.7669214

Park, D. J., Park, M. W., Lee, H., Kim, Y.-J., Kim, Y., & Park, Y. H. (2021). Development of machine learning model for diagnostic disease prediction based on laboratory tests. *Scientific Reports*, *11*(1), 7567. https://doi.org/10.1038/s41598-021-87171-5

Rentea, R. M., Peter, S. D. S., & Snyder, C. L. (2017). Pediatric appendicitis: state of the art review. *Pediatric Surgery International*, *33*(3), 269–283. https://doi.org/10.1007/s00383-016-3990-2

Rufo, D. D., Debelee, T. G., Ibenthal, A., & Negera, W. G. (2021). Diagnosis of Diabetes Mellitus Using Gradient Boosting Machine (LightGBM). *Diagnostics*, *11*(9), 1714. https://doi.org/10.3390/diagnostics11091714

# 