

# **Propuesta de Proyecto: Detección de Anomalías en Imágenes Radiográficas MURA.**

---

## **Grupo de trabajo:**

### **Santiago Garcia Solarte**

Código estudiante: 2238081

Correo electrónico: [santiago.garcia\\_sol@uao.edu.co](mailto:santiago.garcia_sol@uao.edu.co)

### **Santiago Loaiza Cardona**

Código estudiante: 2238094

Correo electrónico: [santiago.loaiza@uao.edu.co](mailto:santiago.loaiza@uao.edu.co)

## **1. Introducción**

Las radiografías musculoesqueléticas son una herramienta diagnóstica crucial en la medicina moderna, utilizada para evaluar y detectar anomalías en los huesos y articulaciones sin necesidad de procedimientos invasivos. Sin embargo, el análisis manual de estas imágenes requiere un gran conocimiento técnico en medicina y mucha experiencia para llegar a un diagnóstico preciso. Con el avance de la inteligencia artificial y los algoritmos de machine learning, se ha abierto una oportunidad significativa para automatizar la detección de anomalías, mejorando la precisión diagnóstica y permitiendo intervenciones médicas más oportunas.

Este proyecto tiene como objetivo desarrollar un sistema de detección automática de anomalías en imágenes radiográficas del sistema musculoesquelético utilizando el dataset MURA. El sistema se basará en técnicas de machine learning para identificar patrones que indican la presencia de lesiones, facilitando un diagnóstico más preciso y una atención médica más eficiente.

## **2. Objetivo General**

Desarrollar un modelo de machine learning para la detección de anomalías en imágenes radiográficas del sistema musculoesquelético, utilizando el dataset MURA, e implementarlo en una interfaz gráfica de usuario funcional.

## **3. Objetivos Específicos**

- 3.1.** Investigar y seleccionar las partes relevantes del dataset MURA que contengan imágenes radiográficas del sistema musculoesquelético con anomalías específicas.
- 3.2.** Preprocesar los datos del dataset MURA para optimizar el rendimiento del modelo.
- 3.3.** Desarrollar y entrenar un modelo de machine learning para la detección de anomalías en las imágenes radiográficas.

- 3.4. Implementar una interfaz gráfica de usuario que permita la carga de imágenes y la visualización de los resultados de la detección de anomalías.

#### 4. Alcance del Proyecto

El alcance de este proyecto incluye el desarrollo de una aplicativo sencillo para la detección de anomalías en el sistema musculoesquelético mediante machine learning, desde la recopilación y preprocesamiento de datos hasta la implementación y validación del sistema. El proyecto abarca la selección y adaptación de modelos de reconocimiento de imágenes, la implementación de algoritmos de detección de anomalías y el desarrollo de una interfaz simple. Además, se llevarán a cabo pruebas y validaciones para asegurar el buen funcionamiento y la fiabilidad del sistema.

Este proyecto se centrará en detectar una posible lesión ósea en ciertas partes del cuerpo según el conjunto de datos disponible, cabe aclarar que no se cubrirán todas las posibles condiciones médicas, sólo un conjunto de estas.

Cabe aclarar que este proyecto se hace con fines educativos y en ningún momento se pretende que sea usado en el campo clínico.

#### 5. Opciones de Datasets

Las siguientes son cinco opciones de datasets disponibles para la detección de anomalías en imágenes TAC:

- 5.1. MURA (Musculoskeletal Radiographs): Contiene radiografías del sistema musculoesquelético con diversas anomalías.
- 5.2. LIDC-IDRI (Lung Image Database Consortium and Image Database Resource Initiative): Contiene múltiples escaneos TAC con anotaciones de nódulos pulmonares.
- 5.3. RSNA Pneumonia Detection Challenge Dataset: Incluye imágenes TAC con anotaciones para la detección de neumonía.
- 5.4. MosMedData: Chest CT Scans with COVID-19 Related Findings: Dataset de imágenes TAC de tórax con hallazgos relacionados con COVID-19.
- 5.5. CURE-TSD (CURE-Traumatic Skeletal Dataset): Contiene imágenes de traumatismos esqueléticos, incluidas fracturas.

Para el desarrollo de este proyecto será utilizado el DataSet: **MURA** el cual tiene anotaciones para detección de fracturas.

#### 6. Cronograma del Proyecto

El desarrollo de este proyecto a grandes rasgos se tiene planeado de la siguiente manera para un presupuesto de tiempo de 3 semanas:

## → Semana 1

- Investigación y selección del dataset adecuado, que contenga imágenes y sus respectivos diagnósticos.
- Descarga y preprocesamiento inicial de los datos del dataset seleccionado, asegurando que las imágenes y diagnósticos correspondan entre sí
- Crear un dataset listo para usarse en los algoritmos de Machine Learning de identificación de imágenes.

→ **Semana 2**

- Desarrollo y definición de la arquitectura del modelo de machine learning, usando modelos pre-entrenados para la detección de imágenes por medio de Hugging Faces.
- Entrenamiento del modelo con los datos preprocesados, ajustando hiperparametros, para que se genere el reconocimiento y detección de las anomalías.
- Desarrollo de una interfaz gráfica de usuario para la carga de imágenes y visualización de resultados.

→ **Semana 3**

- Optimización y validación del modelo final.
- Pruebas finales y ajustes necesarios en la interfaz y el modelo.

Para detallar más el desarrollo del proyecto, el tiempo y el recurso dispuesto para las labores se presenta el siguiente diagrama de Gantt:

## Detección anomalías en radiografías: MURA

**Empresa desarrollo: IA Solutions SAS.**

**Santiago García Solarte- Santiago Loaiza Cardona**

**Fecha de inicio del proyecto:**

1/08/2024

Leyenda:

Según lo previsto

Riesgo bajo

Riesgo medio

Riesgo alto

Sin asignar

Descripción del hito	Categoría	Asignado a	Progreso	Inicio	Días
Investigación y Preparación de Datos					
Buscar y evaluar datasets disponibles	Objetivo		<div><div></div></div> 50%	1/08/2024	3
Seleccionar el dataset (LIDC-IDRI).	Hito		<div><div></div></div> 15%	4/08/2024	1
Descargar el dataset seleccionado	Riesgo bajo		<div><div></div></div> 10%	5/08/2024	2
Verificar la integridad de los datos descargados	Riesgo alto		<div><div></div></div> 0%	5/08/2024	2
Preprocesamiento Inicial de Datos					
Conversión de formatos de imagen	Riesgo bajo		<div><div></div></div> 0%	7/08/2024	1
Normalización de los datos	Riesgo bajo		<div><div></div></div> 0%	7/08/2024	1
Segmentación y etiquetado de las imágenes	Riesgo medio		<div><div></div></div> 0%	8/08/2024	1
Análisis Exploratorio de Datos					
Visualización de las imágenes	Riesgo medio		<div><div></div></div> 0%	9/08/2024	1
Análisis de las etiquetas y distribución de las anomalías	Riesgo medio		<div><div></div></div> 0%	9/08/2024	1
Identificación de patrones y posibles problemas en los datos	Objetivo		<div><div></div></div> 0%	10/08/2024	1
Desarrollo y Entrenamiento del Modelo					
Selección del tipo de modelo (CNN, ResNet, etc.)	Hito		<div><div></div></div> 0%	10/08/2024	1
Diseño de la arquitectura del modelo	Riesgo alto		<div><div></div></div> 0%	10/08/2024	1
Implementación del modelo en un framework adecuado.	Riesgo alto		<div><div></div></div> 0%	11/08/2024	2
Preparación del Entorno de Entrenamiento					
Configuración del entorno de desarrollo	Hito		<div><div></div></div> 0%	13/08/2024	1
Instalación de dependencias y herramientas necesarias	Riesgo alto		<div><div></div></div> 0%	13/08/2024	1
Entrenamiento Inicial del Modelo	Riesgo alto		<div><div></div></div> 0%	13/08/2024	2
División del dataset en conjuntos	Riesgo bajo		<div><div></div></div> 0%	15/08/2024	1
Entrenamiento del modelo con el conjunto de entrenamiento	Riesgo bajo		<div><div></div></div> 0%	15/08/2024	1
Monitoreo del rendimiento y ajustes iniciales	Riesgo medio		<div><div></div></div> 0%	15/08/2024	1
Evaluación del modelo en el conjunto de validación	Riesgo alto		<div><div></div></div> 0%	16/08/2024	1
Ajuste de hiperparámetros (learning rate, batch size, etc.)	Riesgo alto		<div><div></div></div> 0%	16/08/2024	1
Implementación de técnicas de mejora	Riesgo alto		<div><div></div></div> 0%	16/08/2024	1
Optimización y Desarrollo de la Interfaz					
Validación cruzada del modelo	Riesgo alto		<div><div></div></div> 0%	17/08/2024	1
Optimización del modelo para mejorar el rendimiento	Riesgo alto		<div><div></div></div> 0%	17/08/2024	1
Evaluación final en el conjunto de prueba	Riesgo alto		<div><div></div></div> 0%	17/08/2024	1
Implementación de la interfaz (Tkinter, Flask)	Riesgo medio		<div><div></div></div> 0%	10/08/2024	6
Integración del modelo con la interfaz	Riesgo alto		<div><div></div></div> 0%	16/08/2024	2
Pruebas Finales y Ajustes					
Pruebas exhaustivas (modelo + interfaz)	Riesgo medio		<div><div></div></div> 0%	18/08/2024	1
Identificación y corrección de errores	Riesgo medio		<div><div></div></div> 0%	19/08/2024	1
Documentación del proyecto y preparación para la entrega	Riesgo bajo		<div><div></div></div> 0%	8/08/2024	12
Imprevistos del proyecto	Riesgo alto		<div><div></div></div> 0%	20/08/2024	2

