

Proyecto semestral (primera versión)

Detección de Fracturas Óseas con el modelo de detección de objetos basado en transformadores YOLO.

Equipo:

- Jesus Flores Lacarra
- Sebastián Rodríguez Serrano
- Orlando Lopez Roque
- Michell Berenice Altamirano Ocejo

La detección temprana y precisa de fracturas óseas es esencial para mejorar el diagnóstico clínico y reducir complicaciones derivadas de errores humanos. Este estudio propone utilizar el modelo YOLO para la identificación automática de fracturas en imágenes radiográficas, con el objetivo de evaluar qué tan eficiente resulta para resolver este problema. Se adopta un enfoque cuantitativo basado en el entrenamiento del modelo con un conjunto de radiografías etiquetadas y su posterior validación mediante métricas de precisión, sensibilidad y especificidad. Asimismo, se analiza el tiempo de procesamiento y la capacidad de localización de las lesiones. Los resultados permitirán comparar su desempeño frente a la evaluación tradicional realizada por especialistas y estimar su viabilidad como herramienta de apoyo clínico. Se espera demostrar que YOLO ofrece alta exactitud y rapidez, contribuyendo al desarrollo de sistemas de diagnóstico asistido por inteligencia artificial en el ámbito médico, mejorando la eficiencia y calidad asistencial en hospitales.

1. Introducción

Las fracturas óseas son lesiones musculoesqueléticas muy comunes que requieren un diagnóstico rápido y preciso, generalmente utilizando radiografías convencionales (rayos X). La interpretación manual de los rayos X es difícil, especialmente para fracturas sutiles o difíciles de ver (ocultas), lo que lleva a errores, diagnósticos perdidos o retrasados debido a la presión de tiempo, la gran carga de trabajo clínico, la variabilidad entre observadores (diferentes médicos interpretan de manera diferente) y la fatiga del clínico. Existe una clara necesidad de sistemas de diagnóstico asistido por computadora (CAD) automatizados, fiables y eficientes para ayudar a los médicos en la detección y clasificación de fracturas.

Algunas limitaciones para satisfacer esta necesidad giran en torno a la heterogeneidad de los *datasets* (datos médicos). Los datos públicos y privados son muy diferentes en calidad de imagen, esquemas de etiquetado (cómo se nombran las fracturas) y cobertura anatómica (qué parte del cuerpo se incluye).

Al entrenar modelos por separado en estos conjuntos de datos fragmentados, se obtienen soluciones "específicas para una tarea" (soluciones *ad hoc*) que funcionan bien en pruebas limitadas, pero fallan al aplicarse en diferentes entornos clínicos (falta de generalización).

Al ofrecer una herramienta de soporte que combine la velocidad de la computación con la alta precisión, se alivia la carga de trabajo de los radiólogos, y también se eleva la calidad de la atención diagnóstica, asegurando que las fracturas, incluso las menos evidentes, sean identificadas de manera oportuna.

Proyecto semestral (primera versión)

Con esto, se pueden obtener grandes beneficios, como:

- Acelerar y, posiblemente, automatizar el proceso de diagnóstico.
- Facilitar el diagnóstico para el experto y mejorar la precisión.
- Cuantificar datos y almacenarlos de manera eficiente.

2. Justificación

La importancia de esta investigación reside en efficientizar los tiempos de diagnóstico de médicos especializados.

Se espera que los resultados obtenidos por el modelo puedan ser interpretados de manera rápida y sencilla por dichos médicos, disminuyendo la carga cognitiva de esta tarea para así concentrarse en fases más importantes para la rehabilitación del paciente.

3. Revisión literaria

La aplicación de la Inteligencia Artificial (IA) y, específicamente, del aprendizaje profundo (Deep Learning) en el diagnóstico por imágenes ha emergido como un campo de investigación crucial, particularmente en la radiología. El desafío de la detección de fracturas óseas ha sido abordado históricamente mediante sistemas de diagnóstico asistido por computadora (CAD) que empleaban técnicas de procesamiento de imágenes más tradicionales. Sin embargo, la revolución de las Redes Neuronales Convolucionales (CNN) ha permitido alcanzar niveles de precisión que superan consistentemente a los métodos anteriores.

La detección de fracturas se conceptualiza como un problema de detección de objetos, donde el "objeto" de interés es la línea de fractura. En este contexto, la literatura reciente se centra en el uso de arquitecturas de *Deep Learning* que se dividen principalmente en dos categorías:

1. Modelos de dos etapas (Two-Stage Detectors): Incluyen arquitecturas como R-CNN, Fast R-CNN y Faster R-CNN. Estos modelos primero proponen regiones de interés (*Region Proposal Networks* - RPN) y luego clasifican y refinan los *bounding boxes*. Son conocidos por su alta precisión (*mAP*), pero suelen ser más lentos en la inferencia, lo que puede limitar su utilidad en escenarios clínicos de alta demanda (triaje).
2. Modelos de una etapa (One-Stage Detectors): Arquitecturas como la familia YOLO (You Only Look Once) y SSD (Single Shot MultiBox Detector). Estos modelos realizan la propuesta de región y la clasificación simultáneamente, ofreciendo una velocidad de inferencia significativamente mayor, a menudo en tiempo real o casi real, sacrificando mínimamente la precisión en comparación con los modelos de dos etapas más robustos.

La familia de modelos YOLO ha demostrado ser particularmente prometedora en aplicaciones médicas que requieren rapidez. Desde su introducción, YOLO ha pasado por

Proyecto semestral (primera versión)

varias iteraciones (YOLOv1 a YOLOv8, incluyendo variantes como YOLO-LITE y YOLO-NAS), cada una mejorando la velocidad y precisión del modelo predecesor.

Varios estudios han aplicado versiones específicas de YOLO para la detección de fracturas en diferentes regiones anatómicas (muñeca, tobillo, codo). Por ejemplo, se ha reportado que YOLOv4 y YOLOv5 han logrado métricas de precisión y *recall* comparables o superiores a los de radiólogos *junior* y con un tiempo de inferencia de milisegundos. Esta eficiencia subraya la viabilidad de YOLO como herramienta de triaje o como segundo lector en entornos hospitalarios. La capacidad de YOLO para procesar imágenes rápidamente es un factor diferenciador clave para su adopción clínica. Además, su arquitectura de una sola etapa que analiza la imagen completa facilita la captura de contexto global, lo que es esencial para diferenciar artefactos de fracturas sutiles, una limitación común en los modelos de dos etapas que se enfocan en parches de imagen.

A pesar de los avances, existen desafíos persistentes en el uso de *Deep Learning* para el diagnóstico de fracturas:

- Heterogeneidad de Datos: La falta de grandes datasets estandarizados y etiquetados por expertos sigue siendo la principal barrera. La mayoría de los *datasets* públicos son limitados en tamaño y pueden tener sesgos de distribución (p. ej., predominancia de un tipo de fractura o región anatómica).
- Generalización: Los modelos entrenados en un *dataset* específico a menudo muestran un rendimiento reducido cuando se aplican a datos de otra institución (falta de generalización), lo que destaca la necesidad de técnicas como el *transfer learning* o el entrenamiento federado.
- Fracturas Sutiles/Ocultas: La detección de fracturas de tipo *hairline* o aquellas superpuestas a otras estructuras sigue siendo un cuello de botella, requiriendo a menudo resoluciones
- de imagen más alta o arquitecturas de modelo más sensibles.

Metodología

Para abordar este problema, evaluamos diversos algoritmos de detección de objetos, desde los clásicos hasta Redes Neuronales Convolucionales de dos etapas como Faster R-CNN. Decidimos utilizar YOLO (You Only Look Once, por sus siglas en inglés) porque su enfoque unificado de etapa única nos pareció interesante e idóneo para la aplicación médica. Creemos que es importante porque ofrece la velocidad de inferencia cercana al tiempo real necesaria para un posible triaje clínico manteniendo una alta precisión gracias a su capacidad de analizar el contexto global de la imagen.

La visión por computadoras nos ofrece la extracción de información de imágenes, que es relevante para nuestro proyecto porque nos brinda una descripción de la radiografía ingresada. Convirtiendo esta tecnología en una herramienta de diagnóstico de alta precisión y con capacidad de identificar patrones sutiles que, a veces, pueden pasar desapercibidos por el ojo humano. De modo que sirve como apoyo adicional al médico al momento de tomar decisiones clínicas, acelerando diagnósticos y confirmándolos.

Cronograma

[illegible]