

# BONE EDA

**Tóm tắt bài báo**

**A Radiograph Dataset for the Classification,  
Localization, and Detection of Malignant Bone  
Tumors**

# Tóm tắt

**Bài báo này giới thiệu một bộ dữ liệu X-quang được thiết kế để hỗ trợ phân loại, xác định vị trí và phát hiện các khối u xương ác tính. Các khối u xương ác tính là nguyên nhân gây tử vong do ung thư đứng thứ ba ở bệnh nhân dưới 20 tuổi. Việc chẩn đoán sớm thông qua hình ảnh X-quang đóng vai trò quan trọng trong điều trị.** Tuy nhiên, do sự hiếm gặp của các khối u này, việc thu thập đủ dữ liệu để phát triển các mô hình học máy hiệu quả gặp nhiều khó khăn. Bộ dữ liệu được giới thiệu trong bài báo bao gồm các hình ảnh X-quang được chú thích chi tiết, nhằm hỗ trợ cộng đồng nghiên cứu trong việc phát triển và đánh giá các phương pháp chẩn đoán dựa trên trí tuệ nhân tạo. Bộ dữ liệu này có tiềm năng cải thiện độ chính xác trong chẩn đoán và điều trị các khối u xương ác tính.

# Dữ liệu có trong bài báo

## Dữ liệu đầu vào:

**Hình ảnh X-quang gồm**

- 1.879 ảnh bình thường (không có ung thư)**
- 1.867 ảnh có khối u (ung thư xương)**

## Dữ liệu đầu ra

**Mỗi hình ảnh xuất ra được xác định vị trí và phân loại khối u**

# Bone EDA

**Bone EDA (Exploratory Data Analysis) là quá trình phân tích dữ liệu khám phá tập trung vào dữ liệu liên quan đến xương. Mục tiêu chính là hiểu rõ cấu trúc dữ liệu, phát hiện xu hướng, kiểm tra bất thường và mối quan hệ giữa các biến trước khi xây dựng mô hình phân tích hoặc máy học.**

# Mục tiêu

**Kiểm tra tổng quan dữ liệu:** Xem số lượng ảnh X-quang, thông tin về bệnh nhân, loại tổn thương.

**Phát hiện bất thường:** Xác định dữ liệu bị thiếu, nhiễu hoặc các giá trị ngoại lai.

**Phân tích phân bố:** Xem xét sự phân bố của các biến như tuổi bệnh nhân, vị trí khối u.

**Trực quan hóa dữ liệu:** Sử dụng biểu đồ, heatmap để hiểu mối quan hệ giữa các yếu tố.

# **Ứng dụng data**

**xác định được loại của khối u**

**xác định vị trí**

**Phân loại khối u**

**Hỗ trợ trong y học**

# Xác định khối u trên hình X - quang

```
for shape in shapes:
    points = np.array(shape.get("points", []), dtype=np.int32)
    if len(points) == 0:
        print(f'Lỗi: Không có điểm cho nhãn {shape.get('label', 'Unknown')}')
        continue
    cv2.polylines(image, [points], isClosed=True, color=color_pink, thickness=2)
    cv2.putText(image, shape.get("label", "Unknown"), tuple(points[0] - [0, 10]),
               cv2.FONT_HERSHEY_SIMPLEX, 0.6, color_pink, 2)
```

**Hàm Shape ở đây dùng để Khoanh vùng khối u trên ảnh X - quang**

# Ảnh X - quang



# Kết quả thu được



**Ở đây ta phát hiện được 1 khối u nằm  
phía bên phải góc dưới đan được  
khoanh vùng**