The Third International Conference on AI-Health
AIHealth 2026

Toward a Unified Public X-ray Dataset Integrating Multiple Databases to Advance Complex Fracture Analysis

* *Fatma Atitallah , Assem Thabet , Johannes C. Ayena, Neila Mezghani*

Presented by : ***Fatma ATITALLAH***
atitallah.fatma@outlook.fr

PLAN

PLAN

1 Introduction

2 Methods

3 Results and discussion

4 Conclusion and future work

Medical Context & Current

- Diagnosing upper limb fractures is still difficult because of anatomical overlap and subtle fracture patterns.
- X-ray interpretation is subjective and can vary between clinicians, which may lead to errors or delays in treatment.

AI as a Solution… and Its Limitation

- Artificial Intelligence and Deep Learning offer a promising solution to support clinicians.
- However, their development is limited by the lack of large, high-quality public datasets, which are essential to train robust models.

?

AI as a Solution… and Its Limitation

Existing public datasets are fragmented, heterogeneous (different formats, annotations, and quality), and relatively small, which hinders research progress.

Main Objective

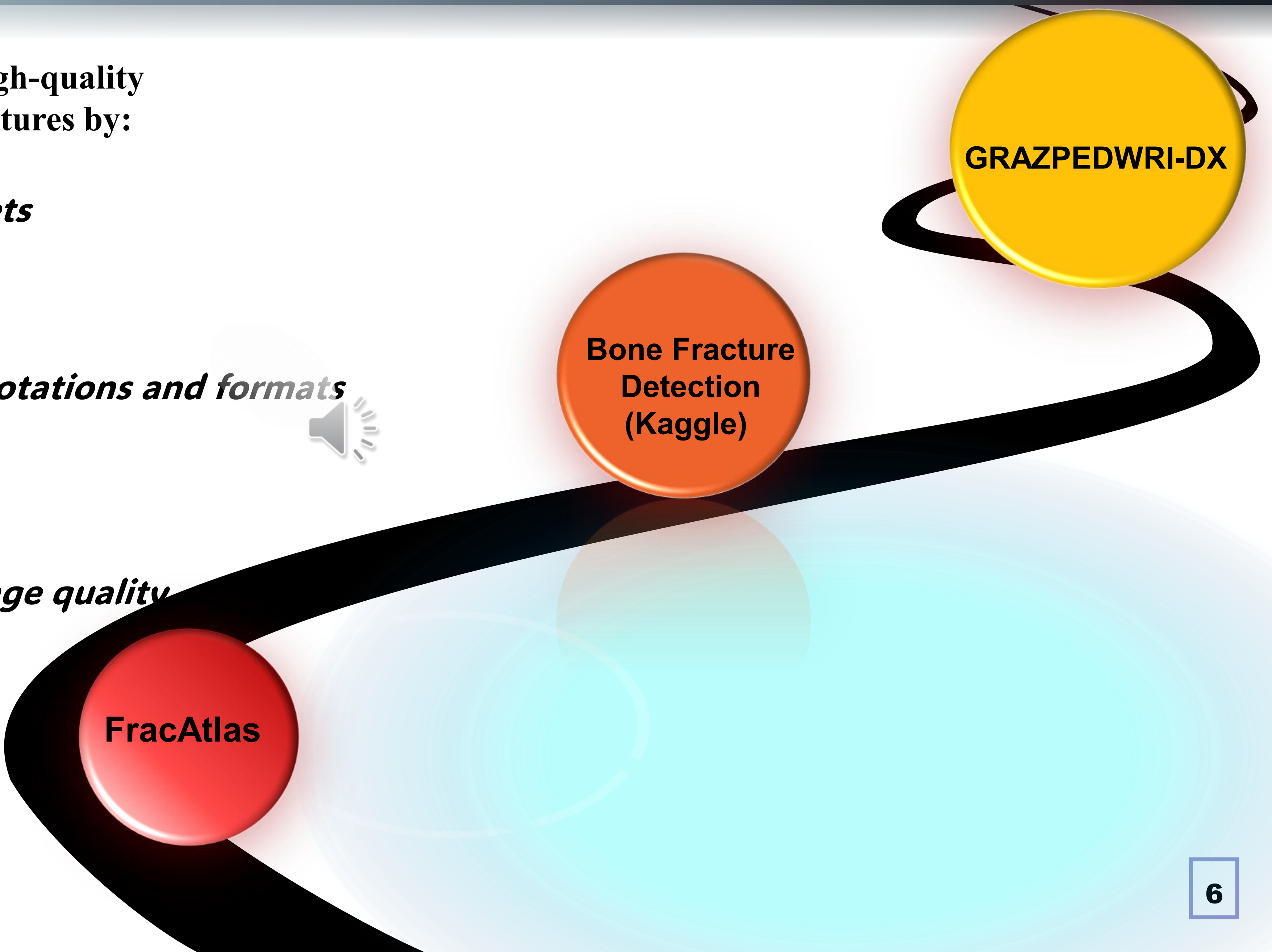
To create a unified, standardized, and high-quality resource for upper limb fracture analysis by merging, harmonizing, and validating three existing public datasets.

PLAN

1. Introduction
2. **Methods**
3. Results and discussion
4. Conclusion and future work

To create a unique, consistent, and high-quality resource dedicated to upper limb fractures by:

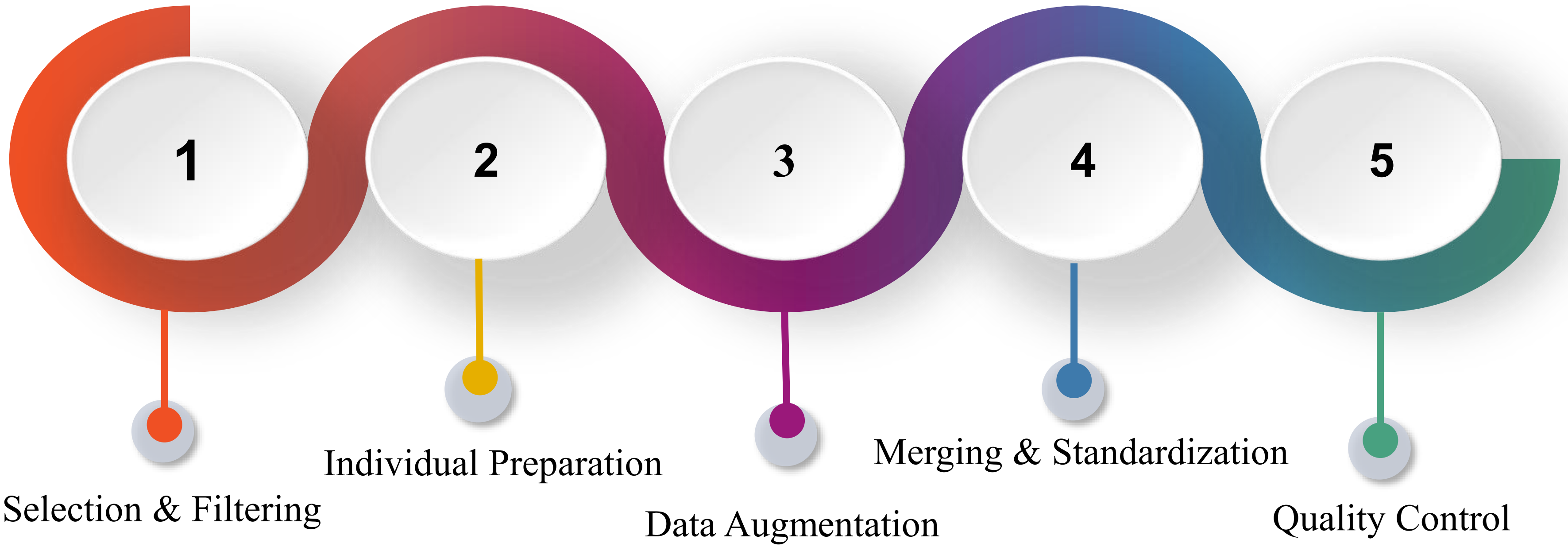
- ***Merge 3 public datasets***
- ***Harmonize annotations and formats***
- ***Standardize image quality***



The Three Source Datasets

Feature	FracAtlas	Bone Fracture Detection (Kaggle)	GRAZPEDWRI-DX
Region	Upper & lower limbs	Mainly upper limbs	Wrist only
Population	Adults & children	Not specified	Pediatric (0–19 years)
Annotations	Masks, boxes, labels	Boxes, few masks	Masks, polygons, boxes
Strength	Clinical exams, subtle fractures	Large variety of upper limb cases	Large-scale pediatric dataset
Limitation	Class imbalance	Missing metadata	Single-center, uneven distribution

From Heterogeneity to Harmonization

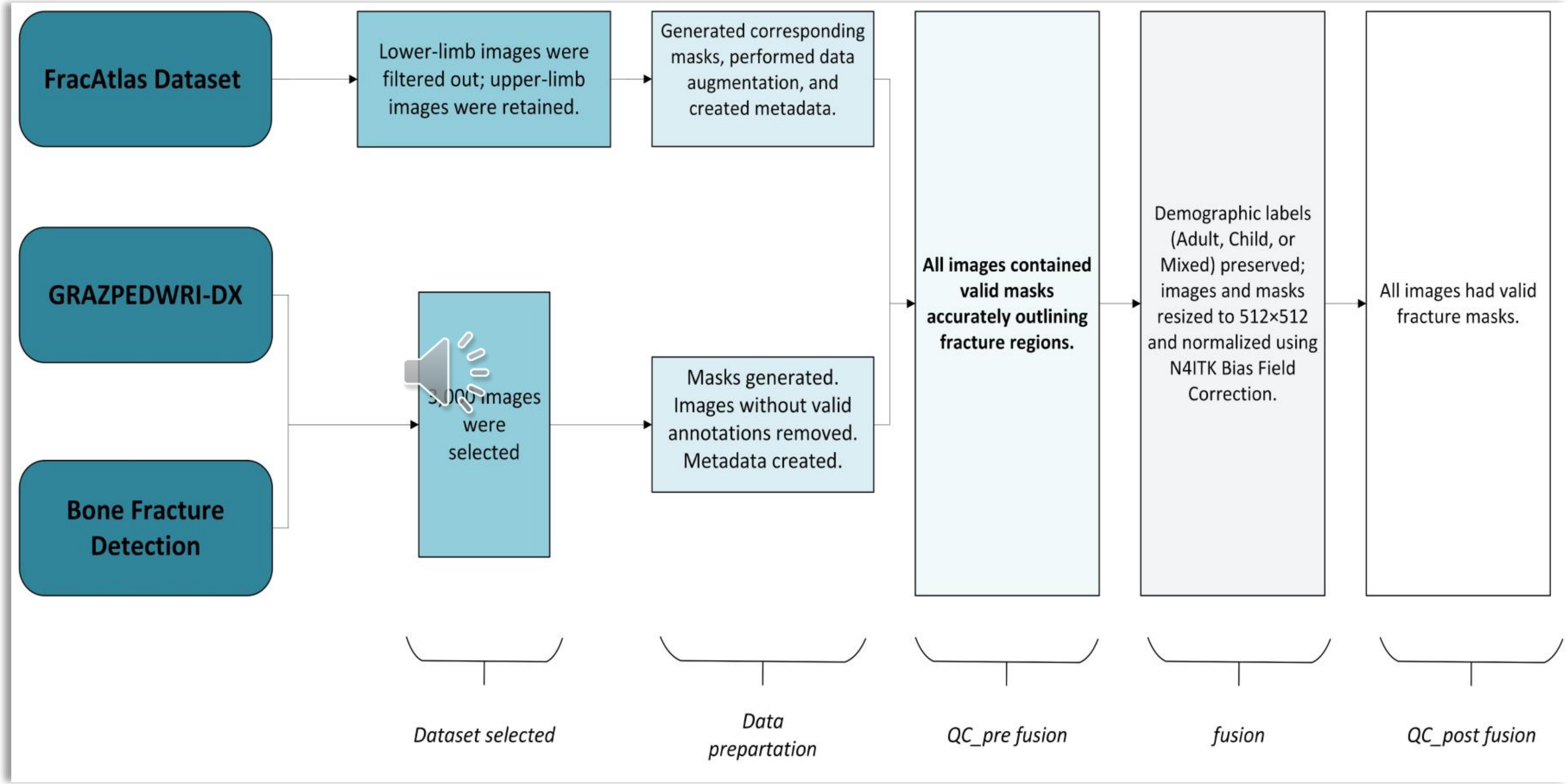


Three heterogeneous sources → One standardized unified dataset

Double quality control (before and after merging)

Targeted data augmentation for FracAtlas

Random but reproducible selection for GRAZPED and Kaggle



Data Preparation – FracAtlas Dataset

Filtering Upper Limb Fractures:

$$\text{UpperLimbFra} = \left\{ i \;\middle|\; (\text{hand}_i = 1 \lor \text{shoulder}_i = 1) \land (\text{leg}_i = 0 \land \text{hip}_i = 0) \land (\text{fractured}_i = 1) \right\}$$

Binary Mask Definition:

$$M_i(x,y) = \begin{cases} 255, & \text{if the pixel } (x,y) \text{ belongs to a fracture,} \\ 0, & \text{otherwise.} \end{cases}$$

Data Augmentation (A = 3):

- Flip: $p_H = 0.5$,$p_V = 0.3$
- Rotation: $\theta \in \{-, 15°, \ldots 15°\}$
- Brightness/Contrast: $I_{aug} = \alpha I_{orig} + \beta$
- Gaussian noise: $\mathcal{N}(0, \sigma^2), \sigma = 10$

Output: $N_{final} = N_{fil} \times (A + 1)$

Bounding Box to Mask Conversion

$$x_{\min} = \max(0, x_c \cdot W - \frac{w \cdot W}{2})$$
$$x_{\max} = \min(W - 1, x_c \cdot W + \frac{w \cdot W}{2})$$
$$y_{\min} = \max(0, y_c \cdot H - \frac{h \cdot H}{2})$$
$$y_{\max} = \min(H - 1, y_c \cdot H + \frac{h \cdot H}{2})$$

Polygon Rasterization

$$M_i(x,y) = \begin{cases} 255, & (x,y) \in \text{polygon interior} \\ 0, & \text{otherwise} \end{cases}$$

Mask Binarization

$$M_i(x,y) = \begin{cases} 255, & M_i(x,y) > 127 \\ 0, & \text{otherwise} \end{cases}$$

Metadata Creation (per image)

- Image ID and file name
- Fracture class labels
- Annotation coordinates (x, y)
- Fracture area estimation

$$A_j = (x_{\max} - x_{\min}) \times (y_{\max} - y_{\min})$$

Dataset Statistics

Distribution of fracture types

$$p_k = \frac{\text{count}(c_j = k)}{N_{\text{final}}} \times 100$$

Data Preparation – GRAZPEDWRI-DX Dataset

Subset Selection

- Random selection of 3000 images/masks with fixed seed for reproducibility

Mask Generation:

$$M_i(x,y) = \begin{cases} 255, & \text{if the pixel } (x,y) \\ & \text{belongs to a fracture,} \\ 0, & \text{otherwise.} \end{cases}$$

Preprocessing Steps

- Parse Pascal VOC annotations
- Polygon & bounding box extraction
- Resize and clamp masks
- Harmonized filenames
- Metadata CSV creation

Output: Pediatric wrist fracture segmentation dataset

Dataset Fusion

- ✓ Three datasets merged into a unified database
- ✓ Demographic labels preserved (Adult, Child, Mixed)

<u>Spatial Standardization</u>

$$I_i^{\text{res}} = \text{Res}(I_i, 512, 512, \text{interp} = \text{cubic}),$$
$$M_i^{\text{res}} = \text{Res}(M_i, 512, 512, \text{interp} = \text{nearest})$$

- ✓ Cubic interpolation for images
- ✓ Nearest-neighbor for masks

<u>Bias Field Correction (N4ITK)</u>

$$\hat{I}_i = I_i^{\text{res}} \cdot B_i^{-1}, \quad B_i = \text{BiasField}(I_i^{\text{res}})$$

<u>Intensity Normalization</u>

$$I_i^{\text{norm}} = \frac{\hat{I}_i - \mu_{\hat{I}_i}}{\sigma_{\hat{I}_i}} \cdot 255$$

The resulting fused D_{fused} and normalized dataset I_{norm}

$$\mathcal{D}_{\text{fused}} = \{(I_i^{\text{norm}}, M_i^{\text{bin}}, \text{metadata}_i)\}_{i=1}^{N_{\text{total}}}$$

Dataset Quality Assurance

Automated Quality Checks

- Mask existence and size validation
- Empty mask detection
- Logging of anomalies

Quantitative Fracture Coverage

$$\text{Fracture\_\%}_i = \frac{\sum_{x,y} \mathbb{1}[M_i(x,y) > 127]}{\text{width}_i \times \text{height}_i} \times 100,$$

Fracture Morphology Complexity

$$\text{Complexity}(C_j) = 0.5 \cdot \frac{\max(w_j, h_j)}{\min(w_j, h_j)} + 0.5 \cdot \frac{P_j^2}{A_j}.$$

$$\text{Complexity}(I_i) = \frac{1}{n_i} \sum_{j=1}^{n_i} \text{Complexity}(C_j).$$

$$\text{Mean Complexity} = \frac{1}{N_{\text{final}}} \sum_{i=1}^{N_{\text{final}}} \text{Complexity}(I_i),$$

$$\text{Max Complexity} = \max_{i \le N_{\text{final}}} \text{Complexity}(I_i).$$

PLAN

1 Introduction

2 Methods

3 **Results and discussion**

4 Conclusion and future work

Tab. Unified summary of the three datasets

Dataset	***N_init***	***Filtering & Resulting Outputs***			***Surface Statistics (%)***				***Complexity***
		N_fil	A	N_final	Min	Max	Mean	Median	Mean / Max
FracAtlas	717	439	3	1,756	0.007	2.241	0.352	0.312	11.48 / 28.66
Kaggle	3,000	1,491	--	1,491	0.123	18.583	2.256	1.651	11.20 / 24.97
GRAZPED WRI-DX	3,000	2,017	--	2,017	0.320	14.986	2.622	2.288	11.56 / 24.63

Notes

- **N_init** : initial number of X-ray images selected in each dataset
- **N_fil** : number of valid image–mask pairs remaining after filtering
- **A** : number of augmentation operations applied (only FracAtlas was augmented with A = 3)
- **N_final** : final number of image–mask pairs after filtering and augmentation
- **Min, Max, Mean, Median (%)** : proportion of fracture pixels
- **Mean / Max of Complexity** : average and maximum fracture complexity based on morphological descriptors

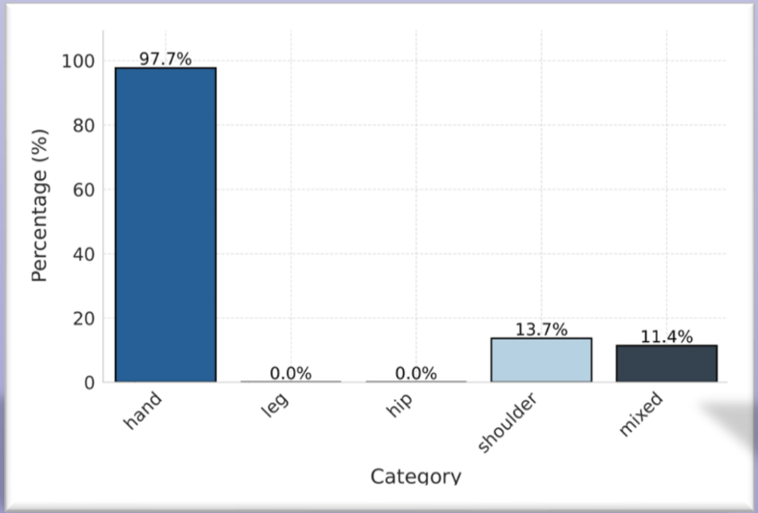


Figure 1. Class distribution in the selected FracAtlas subset.

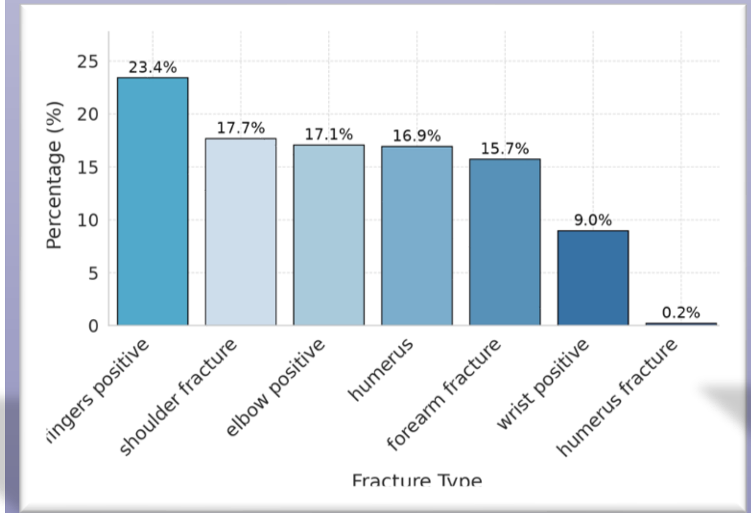


Figure 2. Class distribution in the selected Bone Fracture Detection subset.

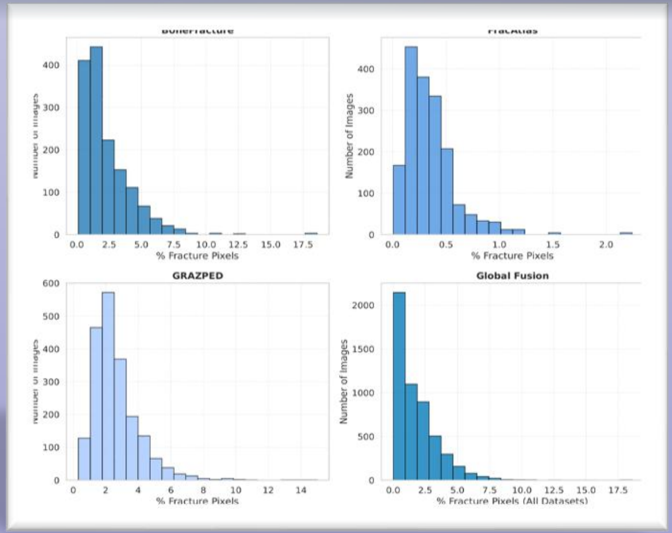


Figure 3. Distribution of fractured pixels across datasets.

the distribution of fracture types after filtering and augmentation. Hand fractures dominate, while leg and hip fractures were effectively excluded, confirming that the dataset focuses on upper-limb fractures suitable for our analysis.

The Kaggle Bone Fracture dataset underwent the same preprocessing pipeline. Out of 3,000 images, 1,491 valid image-mask pairs were retained. Masks were generated and fracture percentages calculated to support statistical analysis. As shown, there is some class imbalance, particularly for wrist-positive and humerus fractures.

For the GRAZPEDWRI-DX dataset, we retained 2,017 valid image-mask pairs. All images, masks, and metadata were harmonized, making this subset ready to integrate with the other datasets for robust training and evaluation.

PLAN

1 Introduction

2 Methods

3 Results and discussion

4 **Conclusion and future work**

Conclusion: Unified Dataset

- Unified X-ray dataset for upper-limb fractures
- Merges FracAtlas, Kaggle, and GRAZPEDWRI-DX
- Harmonized images, annotations, and preprocessing
- Suitable for training and evaluating deep learning models

Dataset Features

- Broad anatomical coverage
- Detailed segmentation masks
- Supports detection, localization, and segmentation research

Limitations & Challenges

- Class imbalance (subtle fractures overrepresented)
- Limited representation of complex fractures
- Subjective augmentation choice (A = 3)

Future Work

- Expand dataset and targeted sampling
- Benchmark state-of-the-art models
- Explore clinical deployment potential

Reproducibility & Data Access

- Dataset publicly availableà
- Synapse platform ensures reproducibility
- DOI: https://doi.org/10.7303/syn71834100

Acknowledgment

- Financial support by Mitacs Globalink program
- Fondation Docteur Sadok Besrour

Thanks for your attention

atitallah.fatma@outlook.fr