Implementing YOLOv8 with Attention Mechanisms for Pediatric Wrist Fracture Detection

Rithvik Doshi, Tianyi Wang, Yiyang Cai, Yuzhe Xu
rithvik@bu.edu, danielty@bu.edu, Yycai@bu.edu, yx8756a@bu.edu

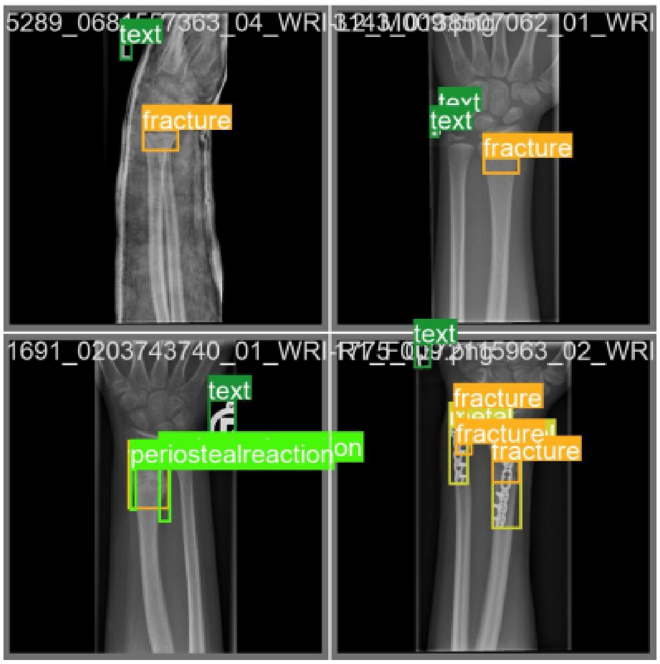

Figure 1: Pediatric Wrist Fracture Detection. Image from [1]

1. Task

The Team aims to replicate and analyze the result of the paper "YOLOv8-AM: YOLOv8 with Attention Mechanisms for Pediatric Wrist Fracture Detection" [1] which investigates an innovative approach to improving fracture detection using neural networks. The team looks to progress by taking the following steps: firstly, we will collect a suitable dataset of wrist X-ray images with labeled fracture annotations, likely using the same or a similar data source as [1]. Secondly, we'll proceed to implement the YOLOv8-AM architecture by incorporating attention mechanisms (Convolutional Block Attention Module, Global Attention Mechanism, Efficient Channel Attention, and Shuffle Attention) into the original YOLOv8, train the model using the prepared dataset and fine-tune hyperparameters, such as learning rate, batch size, and optimizer settings. Next, we will evaluate the result by calculating precision, recall, and F1-score for fracture detection. Subsequently, we will compare the performance of YOLOv8-AM with other fracture detection models, especially the original YOLOv8 without attention mechanisms as well as other models. This includes experimenting with different attention module combinations (e.g., ResBlock + CBAM, GAM, ResGAM) to find the best-performing variant. Finally, we hope to test on other real world data as well as potential application on adult fractures.

2. Related Work

The use of deep learning for fracture detection, particularly in wrist fractures and pediatric cases, has shown promising results in recent studies. The paper "Critical evaluation of deep neural networks for wrist fracture detection" [2] highlights the potential of convolutional neural networks (CNNs) in automating wrist fracture detection from X-ray images. However, it emphasizes the need for rigorous testing, especially on challenging cases that require confirmation by CT imaging.

Another study, "Fracture detection in pediatric wrist trauma X-ray images using YOLOv8 algorithm," [3] focuses on pediatric cases, where accurate fracture detection is crucial for treatment decisions. The study utilizes the YOLOv8 algorithm, a state-of-the-art object detection model, trained on a pediatric wrist trauma X-ray dataset. The results show significant improvement over previous models, demonstrating the potential for deep learning to assist in fracture detection for pediatric patients.

Additionally, "ParallelNet: multiple backbone network for detection tasks on thigh bone fracture" [4] proposes a novel two-stage R-CNN network for thigh fracture detection. The method incorporates multiple parallel backbone networks and a feature fusion connection structure to extract features with different reception fields. The model achieves high performance on a thigh fracture dataset, outperforming other state-of-the-art deep learning frameworks.

These studies collectively highlight the potential of deep learning in improving fracture detection accuracy, especially in challenging cases and pediatric patients, which can ultimately lead to better patient care and treatment outcomes.

3. Approach

In our proposal, we plan to re-implement the YOLOv8-AM model for pediatric wrist fracture detection, as detailed in the reference article[1]. Our approach is to integrate attention mechanisms into the YOLOv8 architecture to improve detection accuracy. Specifically, we will integrate the Convolutional Block Attention Module (CBAM), Global Attention Mechanism (GAM), Efficient Channel Attention (ECA), and Shuffled Attention (SA) into the model. In order to adapt the architecture to our specific dataset, we may adjust the attention mechanisms based on preliminary results. We intend to build on the existing YOLOv8 implementation, utilizing the PyTorch framework, and will develop additional code to integrate the attention module. Our dataset will undergo preprocessing steps such as normalization and augmentation to improve the robustness of the model. To adapt the architecture to our needs, we will adjust the loss function to better handle class imbalances common in medical imaging datasets. If our experiments show that the standard loss does not adequately address the class imbalance problem, we will also write custom code for a data loader that efficiently manages the augmented dataset, as well as a new loss function. The program will include rigorous evaluation using standard metrics such as mAP, with a focus on achieving high accuracy in fracture detection. Our evaluation will focus on metrics such as mean accuracy (mAP) at various intersection and union (IoU) thresholds to assess the effectiveness of the model.

4. Dataset and Metric

To reimplement the paper, we will use the GRAZPEDWRI-DX dataset, which contains 20,327 pediatric wrist trauma X-ray images with 74,459 image labels and 67,771 labeled objects. These X-ray images were collected by multiple pediatric radiologists at the Department for Pediatric Surgery of the University Hospital Graz between 2008 and 2018, involving 6,091 patients and a total of 10,643 studies. This dataset is annotated with 74,459 image labels, featuring a total of 67,771 labeled objects. The dataset will be divided into 70% training (14,204 images), 20% validation (4,094 images), and 10% test sets (2,029 images). Preprocessing will involve adjusting the contrast and brightness using OpenCV's addWeighted function to increase the robustness of the model due to the dataset's limited brightness diversity.

For evaluation metrics, you will use Mean Average Precision (mAP) to assess the object detection model's performance, focusing on the model's ability to correctly identify and locate objects within the images. mAP combines precision (the proportion of correct object detections) and recall (the proportion of actual objects detected) across all categories, providing a comprehensive measure of model accuracy.

5. Approximate Timeline

Task	**Deadline**
Data Preprocessing	March 22
Build Model	April 5
Evaluation	April 19
Optimization	April 26

References

1) Chien, Chun-Tse, et al. "YOLOv8-AM: YOLOv8 with Attention Mechanisms for Pediatric Wrist Fracture Detection." arXiv preprint arXiv:2402.09329 (2024).
2) Raisuddin, A.M., Vaattovaara, E., Nevalainen, M. et al. Critical evaluation of deep neural networks for wrist fracture detection. Sci Rep 11, 6006 (2021). https://doi.org/10.1038/s41598-021-85570-2
3) Ju, RY., Cai, W. Fracture detection in pediatric wrist trauma X-ray images using YOLOv8 algorithm. Sci Rep 13, 20077 (2023). https://doi.org/10.1038/s41598-023-47460-7
4) Wang, Mengxuan & Yao, Jinkun & Zhang, Guoshan & Guan, Bin & Wang, Xinbo & Yueming, Zhang. (2021). ParallelNet: multiple backbone network for detection tasks on thigh bone fracture. Multimedia Systems. 27. 10.1007/s00530-021-00783-9.