International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 06 | Jun 2024 www.irjet.net p-ISSN: 2395-0072

Detection and Classification of Femoral Neck Fracture using YOLOv8

Mousab Abibi Abdi[1], Rafet Akdeniz[2],

[1]MS Student, Computer Engineering, Istanbul Aydin University, Istanbul, Turkey
[2]Professor, Computer Engineering, Istanbul Aydin University, Istanbul, Turkey

---***---

Abstract - : *Femoral neck fractures represent a critical orthopedic problem to their complexity and possible complications. For this reason, it is crucial to detect these fractures quickly and accurately for proper clinical management. This study use YOLOv8 model to detect and classify femoral neck fractures in X-ray images. The performance of YOLOv8 is 97.9% in mAP50, 93.5 in precision and 62.5% in mAP50-95. Our proposed system consist data collection, preprocessing, training and testing the model and model deployment. The proposed model shows potential for automated detection and classification of femoral neck fractures which provides valuable assistance to radiologists.*

***Key Words*: Femoral neck fracture, deep learning, YOLO, YOLOv8, X-Ray images.**

1.INTRODUCTION

Femoral neck fractures are a significant cause of death and disability in the elderly [1], [2]. While examining pelvis radiographs, slipping femur fractures are sometimes missed, and their delayed diagnosis means an increase in costs as well as harmful outcomes [3]. Consequently, the effectiveness of any recommendation for treatment may depend on a prompt and accurate diagnosis. In practice, doctors and radiologists use X-ray images to find fractures. Detecting these fractures through manual checks or with the help of a conventional X-ray machine is a laborious and time-consuming process.

In four non-Level-1 trauma hospital emergency rooms, 975 radiography patients with subsequent CT1 showed 68 false-negative cases. The greater trochanter, ilium, and pubis are the areas with the most often overlooked fractures [3]. Osteoporotic hip fractures account for 14% of those occurring in the United States of America (USA), but they represent 72% of costs related to healthcare caused by osteoporotic fractures [3]. Costs are the other part of the story. Patients who cannot walk by themselves are 40% after a hip fracture; the other 60% cannot do at least one daily living activity, while 80% have additional limitations such as no driving anymore after these kinds of fractures have affected their pelvises.

Additionally, it is widely accepted that current technologies such as x-ray, CT scan, or MRI are not always enough to diagnose severe healing abnormalities [4] like delayed unions or non-unions in patients with femur fractures. There is a necessity to introduce AI-based solutions to revolutionize the approach by allowing for more uniform and algorithm-driven analysis. These advances in AI-driven approaches offer great ways for better-automated image segmentation and local feature extraction, which may replace subjective radiographic union scoring with a standard and portable algorithm-based evaluation of severe injuries like femur fractures.

Studies have found that the use of deep learning techniques in femoral neck fracture detection is promising. In recent research, advanced deep-learning techniques were implemented to automate the detection and classification process for femoral neck fractures, with an accuracy rate of 92.3% in two-class prediction cases and 86.% in three-class prediction [5]. In another study, deep learning and genetic algorithm approaches were used, in which a sensitivity value of 83% was recorded, a specificity value of 73%, and an F1 score value of.78, respectively [6]. The level of accuracy achieved in this particular study shows that combining genetic algorithms with deep learning can enhance fracture detection.

Furthermore, the YOLOv8 algorithm is applied for the detection of fractures in pediatric wrist trauma x-ray images, attaining a mean average precision (mAP) of.638% under an overlap of 50% [7]. This is an indication that YOLOv8 performs well in identifying fractures within pediatric cases accurately. Another deployed YOLOv8 model on femur fracture detection gave out an mAP score of 0.842 at a 50% overlap threshold for both precision 0.85 and recall 0.83 [8].

2. PROPOSED SYSTEM

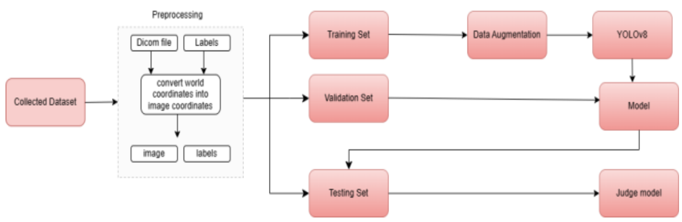

Fig-1: Proposed System

Our proposed femoral neck fracture detection system using YOLOv8 is based on a well-structured machine learning pipeline. The first step involves collecting data by acquiring original radiographs of femoral necks. This is then followed by preprocessing, where images undergo various enhancement procedures, such as image augmentation. In the training and testing phases, the YOLOv8 model is trained on the prepared dataset. The model learns to detect and classify fracture and normal patterns using its advanced deep learning architecture, initiating iterative tests for fine-tuning its parameters and validating applicable accuracy levels at each stage. Finally, the model was deployed to differentiate between normal femoral necks and those with fractures.

3. MODEL ARCHITECTURE

YOLOv8 consist of backbone for feature extraction, followed by a neck for refinement and a head for predicting bounding boxes, class probabilities, and confidence scores. Post-processing involves filtering and non-maximum suppression to produce the final object detection results.

A. Backbone

The backbone constitutes a pre-trained Convolutional Neural Network (CNN) that extracts features from the input image. In YOLOv8, a modified version of CSPDarknet53 architecture consist of 53 convolutional layers is used. It utilizes cross-stage partial connections which help improve the flow of information between different layers enhancing its ability to capture low or high-level features from an image.

B. Neck

The neck is responsible for merging feature maps that are extracted from the output of backbone through various methods like Path Aggregation Blocks or Feature Pyramid Network (FPN) at various scales so that it can be able to detect objects of different sizes.

C. Head

The head is composed of multiple convolutional layers followed by several fully connected layers. It is on this part that we predict bounding boxes, object classes as well as confidence scores for all detected objects.

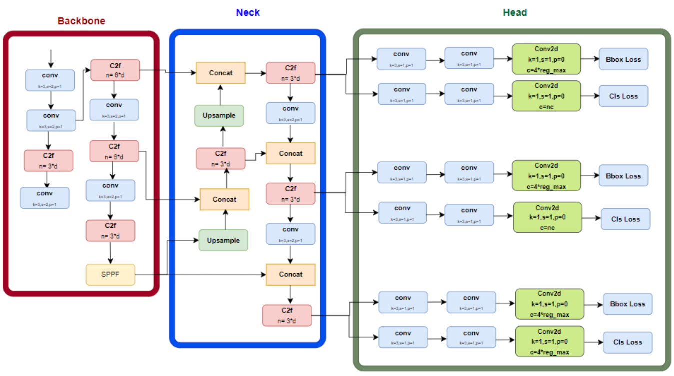

Fig 2: YOLOv8 architecture

4. EXPERIMENT AND RESULT

4.1 Experiment

Initialy we used a dataset of 475 images divided into two classes: "broken" and "normal." To ensure reliable training, we separated the original dataset into three subsets: training, validation, and test. A Python script was used to achieve this split, with 70% of the data (332 images) allocated randomly to the training set, 20% (95 images) to the validation set, and 10% (47 images) to the test set. Following the split, we focused solely on data augmentation for the training set.

4.2 Evaluation Metrics

A. Precision

Evaluate the capacity or performance of object detection models by means of Precision. Precision is how well the model can avoid giving false positives.

$$Precision = \frac{TP}{TP + FP}$$

B. Recall

Recall or sensitivity refers to the portion of right predicted items (true positives) from all actual items found in a dataset; this shows that every pertinent occurrence of an object within an image was covered by the model. To calculate recall use

$$Recall = \frac{TP}{TP + FN}$$

C. MAP50

mAP50 evaluates the average precision over all classes while assessing its ability at 50% IoU criterion. Intersection over Union (IoU) measures the overlap between ground truth bounding boxes and predicted bounding boxes

$$AP50 = \frac{1}{N}\sum_{i=1}^{N} APi$$

D. MAP50-95

Its score is obtained by taking the mean average precision for different IoU thresholds from 50% through 95%. These thresholds embody the sensitivity of the model in detecting objects that have variable levels of overlap with the ground-truth bounding boxes.

$$AP50-95 = \frac{1}{N}\sum_{i=1}^{N} AP50-95\ i$$

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 06 | Jun 2024 www.irjet.net p-ISSN: 2395-0072

4.3 Result

During the training of the model, we initially set 300 epochs but had an early stoppage that shows us that our best epoch are within 150 epoch. We trained the model with Adam optimizer known for its effectiveness with small datasets and set the image size to be 800*800 to allow the model capture detailed features.

Table 1. Validation results of YOLOv8 for each classes on the dataset

Class	Optimizer	mAP50	mAP50-95	Precision	Recall
All	Adam	97.9%	62.5%	93.5%	96.8%
Normal	Adam	97.8%	61.9%	94.2%	96.7%
Broken	Adam	97.9%	63.1%	92.9%	96.8%

Examining the validation results of our model for each class shows similarities in the model's performance across classes as shown in **Table 1**. Across all classes, the Adam optimizer consistently exhibits effective optimization, with a parameter count of 25.8 million. The model's mean Average Precision (mAP50) of 97.9% indicates how well it can detect objects across the dataset. The Broken class has better performance metrics than the Normal class, with a higher mAP50-95% of 63.1% and a precision of 92.9%.

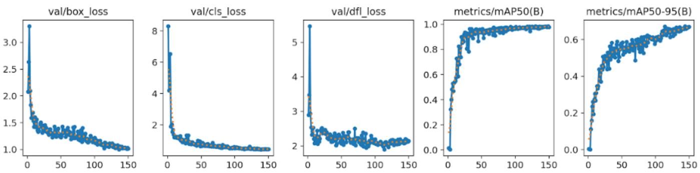

Fig 3. Validations graphs for YOLOv8m

Fig 3. Displays graphs depicting the results for both training and validation sets. This figure consists of three categories of losses (box*loss, cls*loss, dfl*loss*) as well as Precision, Recall, mAP50, mAP50-95 values.

4.4 Comparison

We compared the performance of a YOLOv8m model and a Faster R-CNN (Region-based Convolutional Neural Network) variation. Faster R-CNN is a traditional object detection framework that uses a region proposal network (RPN) to create region proposals before moving on to a detection network for classification and bounding box regression.

Table 2. Comparison of YOLOv8 and Faster RCNN

Models	Optimizer	Param-Eter (M)	mAP50	mAP50-95	Precision	Recall
Yolov8	Adam	25.8	97.9%	62.5%	93.5%	96.8%
Faster rcnn	Adam	4.4	89.4%	56.3%	67.3%	65.5%

Table 2. Shows how our YOLOv8 outperforms Faster R-CNN due to its single-stage architecture. In addition, our model reduces compute load by predicting boxes as well as class scores concurrently instead of using a multi-stage technique like Faster-RCNN.

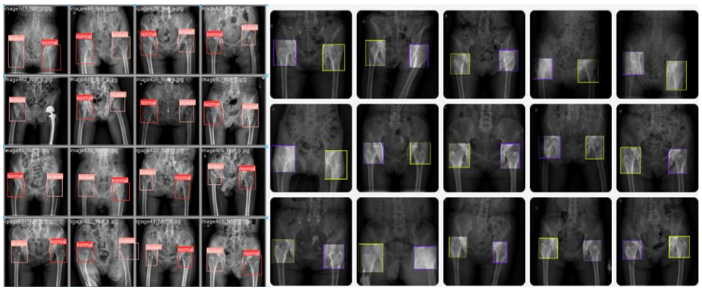

Fig 4. Example of femoral neck fracture detection on X-ray images left side predicted images right manually labeled images.

5. CONCLUSIONS

This study aims to support physicians and intern doctors by introducing the usage of YOLOv8 model to detect and classify femoral neck fracture from X-ray images. The results of model shows that it detect better than other deep learning models. Based on our result we have shown that integrating AI based solution into clinical practice, doctors and radiologists can benefit from a reliable tool that aids in quicker and more accurate diagnosis, potentially leading to better patient care and outcomes.

The research could expand on this foundation by incorporating larger datasets and additional imaging techniques, paving the way for even more accurate and versatile diagnostic tools.

REFERENCES

[1] R. Ju and W. Cai, “Fracture detection in pediatric wrist trauma X-ray images using YOLOv8 algorithm,” *Sci Rep*, vol. 13, p. 20077, 2023. doi: 10.1038/s41598-023-47460-7.

[2] R. Marks, J. P. Allegrante, C. R. MacKenzie, and J. M. Lane, “Hip fractures among the elderly: Causes, consequences, and control,” ScienceDirect, 2003.

[3] M. Weitz, C. Schwartz, and M. H. Scheinfeld, “Radiologic blind spots in hip and pelvic radiographs,” Emerg Radiol, vol. 30, pp. 569–575, 2023. doi: 10.1007/s10140-023-02156-3.

[4] P. Schwarzenberg, S. Darwiche, R. S. Yoon, et al., “Imaging modalities to assess fracture healing,” Curr Osteoporos Rep, vol. 18, pp. 169–179, 2020. doi: 10.1007/s11914-020-00584-5.

[5] S. Mutasa, S. Varada, A. Goel, et al., "Advanced deep learning techniques applied to automated femoral neck fracture detection and classification," J Digit Imaging, vol. 33, pp. 1209–1217, 2020. doi: 10.1007/s10278-020-00364-8.

[6] S. Beyaz, K. Açıcı, and E. Sümer, "Femoral neck fracture detection in X-ray images using deep learning and genetic algorithm approaches," Joint Diseases and Related Surgery, vol. 31, no. 2, pp. 175-183, 2020.

[7] R.-Y. Ju and W. Cai, "Fracture detection in pediatric wrist trauma X-ray images using YOLOv8 algorithm," arXiv preprint arXiv:2304.05071, 2023.

[8] D. Xie, H. Yu, Z. Du, H. Wang, X. Shen, and Z. Wang, "Femur fracture detection based on deep learning model YOLOv8," in Proc. 2nd Int. Seminar Artif. Intell., Networking Inf. Technol. - Vol. 1: ANIT, SciTePress, 2023, pp. 190-193. doi: 10.5220/0012277300003807.