Kashf Journal of Multidisciplinary Research
Vol: 02 - Issue 12 (2025)
P-ISSN: 3007-1992 E-ISSN: 3007-200X
https://kjmr.com.pk

COMPREHENSIVE DETECTION OF MULTI-TYPE WRIST FRACTURES USING IMPROVED YOLOv8 MODEL

Hina Mohsin[1], Ghazanfar Ali[2], Imsal Shabbir Mirza[3], Iqra Hameed[4], Tuba Younas[5]

[1]Department of Computer Science , University of engineering and Technology, Lahore, Pakistan.

[2]Department of Computer Science, Muhammad Nawaz Sharif University of Agriculture, Multan, Pakistan

[3]Department of Computer Science, Government College University, Lahore, Pakistan.

[4]Department of Computer Science, University of Agriculture, Faisalabad, Pakistan.

[5]Department of Computer Science, Lahore College for Women University, Lahore

[*]***Corresponding Author:*** (hinamohsin915@gmail.com)

Article Info

Abstract

Bone fractures, particularly those affecting the wrists, shoulders, and arms, are common and significantly impact patient care. This study investigates the utility of YOLOv8, a deep learning model, in detecting ulna and radius fractures which are the crucial components of wrist injuries. Through collaboration with phys- iotherapists and comprehensive data collection, an annotated dataset is curated for precise wrist fracture localization. Advanced data augmentation techniques are used including Mosaic, Mix-up, and copy paste, to enhance dataset diver- sity and model robustness. Along with baseline YOLOv8 and iYOLOv8 + GC
which currently reports high precision (97.2%), we also consider the most opti- mized variant. We also focus on the baseline YOLOv8 results relative to this newer version to demonstrate the focus on high precision vs the balance within recall, F1 score, and precision. The model excelled in detecting various wrist fracture types, advancing fracture detection in medical practice. The improved evaluation metrics, including accuracy, precision, recall, and F1-score, highlighted the robustness of the YOLOv8 model in identifying ulna and radius fractures. YOLOv8 achieved high scores with accuracy and precision score of 0.87 each recall 0.88, and F1-score 0.86, which indicates its proficiency in accurate fracture detection. The comparative analysis has highlighted the balanced performance of YOLOv8, as the best models like iYOLOv8 + GC also record higher values of accuracy (up to 97.2) whereas our strategy is robust in various measures. These findings highlight YOLOv8 as a promising diagnostic tool for expedited diagnosis and improved patient care in wrist injury management.

Keywords: *Wrist injuries; ulna and radius fractures; YOLOv8; deep learning; fracture detection.*

1. Introduction

Fractures present a considerable challenge for medical practitioners, especially in emer- gency departments where these injuries are commonly treated. Regions such as the wrists, shoulders, and arms are particularly prone to fractures due to their susceptibil- ity to injury. Untreated wrist fractures involving the radius and ulna are very common and critical in Emergency Departments, yet faint lines on low-quality x-rays are hard to detect. Good motivation for computer-aided detection and the use of sophisticated detectors. Specifically, we are focused on ulna/radius fractures on X-rays, with detailed bounding box annotations, the baseline YOLOv8 model, and we will explicitly analyze the trade-offs of the optimized iYOLOv8 derivatives.

Figure 1 indicates that our wrist consists of many types of bones which are called Carpal bones. These are Scaphoid, Lunate, Triquetrum (also known as the triangular or triquetral bone), Pisiform, Trapezium, Trapezoid, Capitate, and Hamate. Beneath these carpal bones, you'll find the radius and ulna bones Which is our main focus in this study.

Awareness of the symptoms associated with bone fractures can aid in prompt diagnosis and appropriate medical care, ultimately facilitating the recovery process from these injuries [2]. The consequences of radiologist fatigue and workload pressures can be profound. Errors in diagnosis or missed fractures can have significant ramifications for patient outcomes. As a result, healthcare institutions are increasingly exploring innovative solutions, such as artificial intelligence (AI)-assisted diagnostic tools, to help radiologists manage their caseloads more efficiently and accurately. In summation, while the advent of radio graphics, particularly X-rays, has undeniably advanced the early detection of fractures, it has concurrently spotlighted the challenge posed by the shortage of skilled radiologists. Issues with backlog in diagnosis of fractures are still a concern in healthcare thus finding solutions to fill the gap is a priority for the

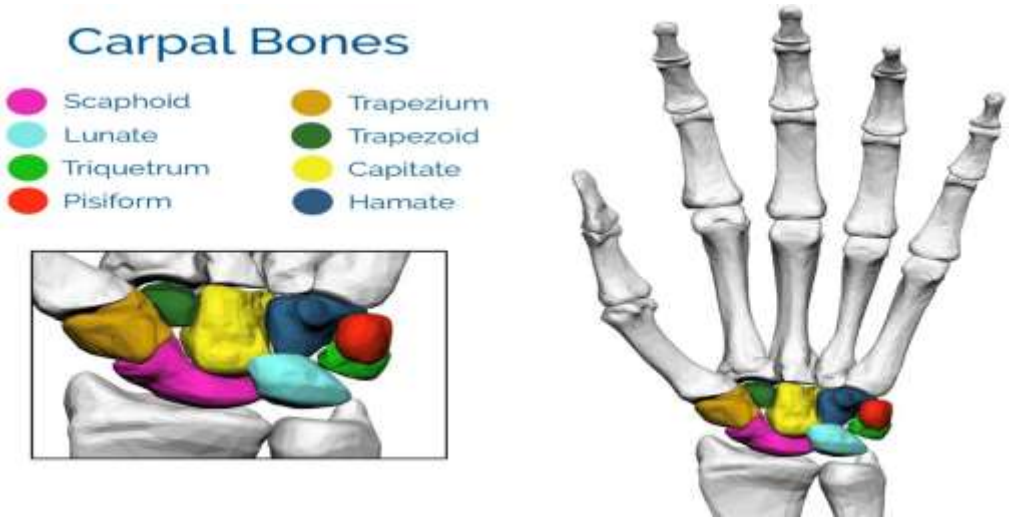

Fig. 1: Hand anatomy representing carpal bones [1].

Health system. Radiologists remain an indispensable asset in this endeavor and thus finding ways to assist them should be a priority [3]. In this regard, healthcare has employed computer-aided diagnostic (CAD) technology that will help physicians in observing the medical images for interpreting medicine. Research nowadays shows that the combination of machine learning (ML) and deep learning (DL) produces greater efficiency in CAD systems and helps the practitioners of healthcare.

ML, a branch under the artificial intelligence (AI) family, operates by utilizing different models including support vector machines (SVM) as well as other techniques.[4] For the diagnostic, categorization, and prognosis of bone fracture through automated means, the KNN [5], NN [6] and RF [7] algorithms are used. Therefore, in a situation where the count of features tangibly exceeds that of the number of available data points, such algorithms then precipitate in the provision of non-optimal solutions that are below standard medical requirements. Today, deep learning enjoys the title of the primary AI technology in the health sector [7].

The DL approach serves as an efficient means of passing through simulators where multiple stages are engulfed in the process of pulling the features and weighing them [8]. Its extensive utility is evident in numerous medical domains, pharmacy, radiology, ophthalmology, pathology [9], and many other various applications like ulnar and radial fracture detection [10, 11], classification of pulmonary tuberculosis [12], and detection of malignancy [13]. In addition, various types of fractures are detected related to ribs [14], hip [15], wrist [16], and ankle [17], as well as, their recognition [18]. Convolutional neural networks (CNNs) have emerged as a particularly efficient approach rapidly gaining prominence in the field of computer vision in recent years. To address clinical challenges, CNN-based trained models such as InceptionV3 [19], ResNet [20], U-Net [21], Xception [22], and DenseNet [23].

The wrist consists of many types of bones, called Carpal bones. Beneath these carpal bones, the radius and ulna bones are found [24]which is the main focus of this study. Given the limited availability of annotated data, techniques like data augmentation and other data generation methods become essential to expand the dataset's size. However, it is worth noting that the analysis of profound learning algorithms remains a challenge. Their accuracy relies heavily on the characteristics of the data, encom- passing diagnostic precision and the extent of fracture detection. In practice, specific data and modifications are often required for previously developed DL methods prior so they can be applied effectively in clinical advancements.

The main contributions of this paper are as follows

- Collaborated with physiotherapists and collected diverse data from online sources and hospitals.
- Carefully annotated the dataset for precise localization of wrist fractures in X-ray images.
- Developed a YOLOv8 model tailored for fracture detection with bounding box localization.

- Applied advanced data augmentation techniques to enhance model robustness.
- Achieved superior accuracy compared to existing models, demonstrating state-of-the-art performance.
- Achieved balanced F1-scores, precision, and accuracy with YOLOv8, which is a good baseline and does not overstate the better results compared to existing optimized models
- Explored detection of multiple types of wrist fractures, contributing to a compre- hensive understanding of fracture detection in medical practice.

Further, Section 2 presents the overall performances of deep learning models and their results section 3 represents the working of the YOLOv8 model and its preprocessing techniques and how the model is employed. Section 4 discusses the comparative performance of various models. Finally, Section 5 concludes this study and gives suggestions for future work.

2. Related Work

Fractures are common injuries, especially in places like wrists and arms. They hap- pen when bones are hit too hard, often due to accidents or overuse. Conditions like osteoporosis and cancer can make bones weaker, increasing the risk of fractures. There are two main types of fractures: open, where the bone breaks through the skin, and closed, where it does not. Doctors need to recognize fractures by looking for signs like pain, swelling, and changes in the bone's appearance. Understanding fractures helps doctors provide better treatment and promote healing.

In recent years, Myint et al. [25] conducted research on utilizing ML algorithms, including K-nearest neighbor (KNN) and decision tree (DT), to detect tibia fractures in X-ray images. Yao et al. [26] utilized chest CT images from 1707 patients to classify rib fractures. Employing a three-step algorithm, they achieved an accuracy score of 0.869 when applied to a total of 4496 CT images. Rashid et al. [27] used DL for wrist fracture detection using X-ray images. Their fused model, combining CNN and long short-term memory (LSTM) networks, achieved improved accuracy (86.21% to 88.24%) compared to existing methods. However, adjustments may be needed with more data. This study highlights the potential of CAD systems for fracture detection in healthcare.

Abigael Cohen et al. [28] examined morphological wrist traits linked to a higher risk of scaphoid fractures. They identified these risk factors through linear metrics and statistical shape modeling (SSM), which documented variations in wrist morphology between poster-anterior and lateral radiographs. Sarbjit Kaur et al. [29] conducted a study on children under 16 years old who experienced acute post-traumatic wrist injuries and were treated following the unit's protocol. The research aimed to eval- uate injury patterns and clinical management. They predicted that early magnetic resonance imaging (MRI) could shorten the immobilization period for scaphoid distal pole fractures, facilitating prompt diagnosis and tailored treatment. Similarly, Hemke et al. [30]

concentrated on utilizing computer-assisted algorithms for bone segmenta- tion. The objective was to use it for diagnosis and treatment strategy planning for wrist fractures.

The segmentation of wrist bones has been investigated by N. Handrix et al. [31] as a preliminary step in wrist fracture classification, by G. Manos et al. [32] for bone age assessment, and by Justyna Wl-odarczyk et al. [33] for rheumatoid arthritis diagnosis. Amanpareet et al. [34] proposed a CNN model which demonstrated high performance. For binary class, the model achieved sensitivity and specificity of 82%, and 85%, respectively while the accuracy is 80%. For multi-class problems, the model obtained an overall sensitivity of 85%, specificity of 85%, and accuracy of 80%.

Amir et al. [35] noted limited studies on wrist bone segmentation from wrist radiographs using DL techniques. The authors observed that low contrast, different distances between carpal bones, and irregularities in bone shapes can contribute to this scarcity. Additionally, according to Daniela Giordano et al., [36], previous studies often focused on segmenting fewer than ten wrist bones. This narrow scope was largely influenced by the research emphasis on children's bone assessment, as their wrist bones have not fully matured. Kang et al. [37] developed a CNN for wrist bone segmentation in retrospective radiographs of older patients. They utilized the Fine Mask R-CNN architecture with single-shot multi-book detector (SSD) layers to extract wrist regions of interest (ROI).

Chen et al. [38] introduced a novel approach called "generating adversarial UNet," which tackles overfitting during training by allowing the generation of diverse medical images. Franko et al. [39] employed a YOLOv4-based model to notably enhance wrist fracture detection in X-ray images, surpassing both radiologists and the U-Net model. By merging YOLOv4's bounding boxes with U-Net's pixel-wise detection, they cre- ated a CAD system. Additionally, integrating natural language processing for patient

history further improved accuracy. This study highlights the potential of YOLOv4- based models in assisting radiologists with wrist fracture diagnosis. Revel et al. [40] developed a finite element (FE) model of the radius to differentiate between fractured and non-fractured bones from low-trauma falls. Using clinical CT scans and cadaveric radii, they achieved strong agreement between simulated and experimental strain val- ues. The FE models, with Pistoia's failure criterion, demonstrated 82% accuracy in classifying radius fractures.

Cohen et al. [41] conducted a retrospective study comparing AI with radiologists in detecting wrist fractures. AI showed a higher sensitivity of 83% than non-specialized radiologists 76%, with a similar specificity of 96%. Combining AI and radiologist analysis improved sensitivity of 88% but reduced specificity of 92%. AI's sensitivity for scaphoid fractures was 84%, but only 41% for other carpal bones. The study suggests AI outperforms non-specialized radiologists, but challenges remain in detecting carpal bone fractures, indicating potential improvement with diverse AI training datasets. Tai-Hua Yang et al. [42] proposed the method Faster R-CNN network to segment the scaphoid bone from X-ray images. Subsequently, the ResNet model is utilized for feature

extraction in the second stage. Detection using rotational bounding boxes resulted in a recall of 0.789 and a precision of 0.894. Fracture classification achieved a recall of 0.735 and a precision of 0.898. Other recent developments are iYOLOv8 optimized variants, like iYOLOv8 + GC, with a 97.2% precision, though with a 67% F1-score (Amirouche et al., 2025)[43].Compared to the iYOLOv8 + GC, our baseline purposefully favors a balanced precision–recall–F1 profile on a smaller data set instead of maximizing a single metric. This reflects similar research in emphasizing when to use a baseline configuration, i.e., clinically useful (less false negatives, stable F1), or other optimized variants (screening cases where ultra-high precision is paramount. Table 1 gives us information about recent studies that have been published for wrist fracture detection with their proposed method, dataset information, and limitations.

Table 1: Exploring the landscape of knowledge: a comprehensive overview of literature.

Reference	**Model**	**Data Set**	**Accuracy**	**Limitations**
Tooba Rashid et al., [3]	CNN-DCNN	Mura	86.21%	Limited dataset for experiments
G. Hajianfar et al., [44]	DenseNet 121 network	Medical College	73%	Only few images are used
M. Revel et al., [40]	Credible finite element model	University Department of Anatomy Rockefeller	82%	- Cost of validation - Small dataset - Complex geometries and loading conditions
Tai-Hua Yang et al., [42]	Faster R-CNN network	National Cheng Kung University Hospital	82%	- High false-negative rate - Limited data
Emre Ozkaya et al., [45]	ResNet50 network	Local Hospital	84%	- Retrospective study - Small sample size - Used only front-to-back (AP) wrist X-rays.

3. Proposed Methodology for Multi Type Wrist Fracture Detection Using YOLOv8

The proposed methodology aims to develop a robust DL model using YOLOv8 archi- tecture for the detection of various types of wrist fractures in X-ray images. With millions affected annually, accurate diagnosis is crucial due to accidents and severe injuries leading to wrist fractures. To achieve this, a dataset of labeled X-ray images capturing different wrist fractures is collected. Leveraging YOLOv8's efficiency and accuracy, the model will be trained and validated on this dataset, known for handling complex detection tasks effectively. The training process involves pre-processing X- ray images to standardize dimensions and enhance clarity. Augmentation techniques enrich the dataset, enabling effective generalization in diverse scenarios. During train- ing, YOLOv8-specific hyperparameters are optimized and fine-tuning strategies are employed to prevent overfitting and improve model stability. Upon successful training, the model will be implemented in clinical settings to assist healthcare professionals in accurate diagnosis of wrist fractures, empowering informed treatment decisions and improving patient outcomes. Figure 2 shows the proposed MultiFractYolov8 that how the detection of fracture along the bounding box on the fractured area of writs is done using the YOLOv8 model. Pipeline outline: (i) assembling annotating of ulna/radius fractures; ii) resizing to 640×640; iii) applying augmenta- tion (Mosaic, MixUp, Copy-Paste); iv) separating into train/val/test; v) baseline YOLOv8 is trained; and vi) evaluating the held-out test set with Accuracy, Precision, Recall, F1, and mAP.

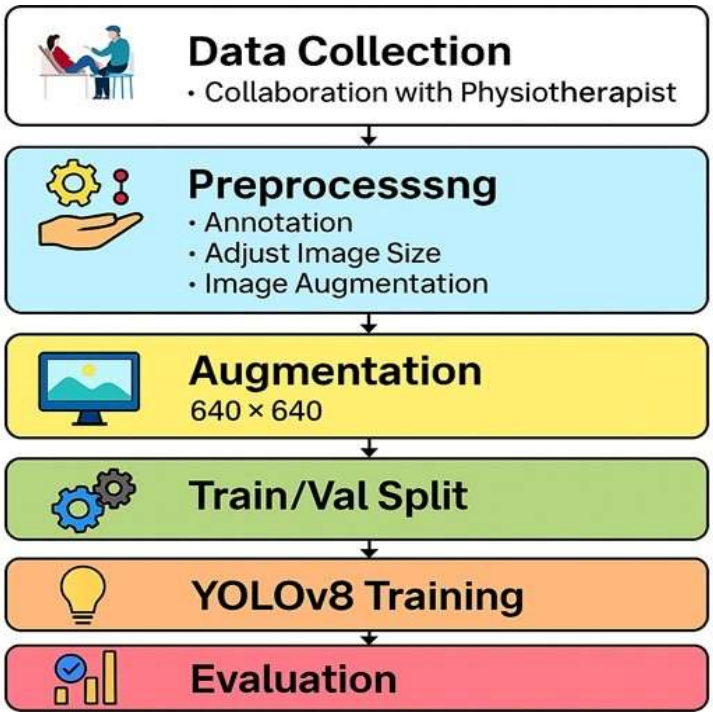

Fig. 2: Proposed methodology: charting the course for research excellence.

In Figure 2 the proposed methodology for wrist fracture detection begins with collecting X-ray images containing ulna and radius fractures. These images are pre- processed and divided into three parts: training, validation, and testing generations. Further, we use YOLOv8 architecture for real-time damage detection, which is sup- ported by advanced human-designed abnormal data augmentation tricks. For the YOLOv8 model, the training step serves the purpose of distinguishing ulna and radius breakages, obtaining precision to work in different conditions of fracture patterns and radiography. After training, it develops an ability to automatically detect fractures on X-ray images with the exact localization of reduced fragments in a clinic for informative treatment choices and successful patient results.

Algorithm 1 represents the methodology of our YOLOv8 model, in line 1 the process starts by gathering and annotating a dataset, labeling each image with the necessary features. In line 3 The images are then resized to 640x640 pixels for unifor- mity and in line 4 normalized to standardize the data. In line 5 This dataset is divided into training and testing sets to evaluate model performance. In lines 7,8 and 9 Aug- mentation techniques like 'Mosaic', 'Mix-up', and 'Copy Paste' are optionally applied to increase data variability. In lines 11,12 and 13 data is prepared for training. In line 16 The YOLOv8 model is trained using the prepared dataset and in line 17 optimized by fine-tuning hyper-parameters. Finally, in line 18 the model's performance is evaluated using the test set, resulting in a fully trained YOLOv8 model ready for object detection tasks. In a bid to ensure transparency and consistency with the reviewer comments, we highlight that this methodology is a base YOLOv8 architecture with additions but is not an optimized or a tailored model.

Algorithm 1 Proposed methodology for training YOLOv8.

Require: Raw dataset $D = \{(I_i, A_i)\}_{i=1}^{700}$ with images I_i of size 640×640

Ensure: Trained detection model *YOLOv8*

1: Collect and annotate the dataset to obtain $D = \{(I_i, A_i)\}_{i=1}^{700}$

2: **for** each (I_i, A_i) in D **do**

3: $I_i \leftarrow \text{Resize}(I_i, 640 \times 640)$

4: $I_i \leftarrow \text{Normalize}(I_i)$

5: **end for**

6: Split D into training set D_{train} and testing set D_{test} 7:
Define augmentation techniques:

8: $T_{\text{Mosaic}} \leftarrow$ Mosaic

9: $T_{\text{Mix-up}} \leftarrow$ Mix-up

10: $T_{\text{CopyPaste}} \leftarrow$ CopyPaste

11: **if** augmentation is needed **then**

$N \leftarrow$ number of images for augmentation
if $N > 0$ **then**
  $(A_{\text{images}}, A_{\text{annotations}}) \leftarrow \text{Augment}(D_{\text{train}}, T_{\text{Mosaic}}, T_{\text{Mix-up}}, T_{\text{CopyPaste}})$
else
  $(A_{\text{images}}, A_{\text{annotations}}) \leftarrow D_{\text{train}}$
end if
else
  $(A_{\text{images}}, A_{\text{annotations}}) \leftarrow D_{\text{train}}$
end if
Train model *Y OLOv*8 using augmented dataset $(A_{\text{images}}, A_{\text{annotations}})$
Optimize hyperparameters θ for *Y OLOv*8 on D_{train}
Evaluate *Y OLOv*8 on D_{test}
return: Trained model *Y OLOv*8

3.1 Data Collection

The data sample consisted of 700 images of radius and ulna fractures, collected from a reliable online platform named GRAZPEDWRI-DX [46], while diversity in fracture types, degree of severity, and orientation was taken into account for data collection. Along with these 50 images were also collected from a hospital, and also the dual- mode training interface comprises 50 highly realistic background images to ensure an appropriate non-fracturing environment. All images were annotated, with polygons and bounding boxes showing the radius and ulna-order injuries. This labeled collection was used to build and test the YOLOv8 model to detect wrist fractures on the different bones. The addition of background images as support to the model helps the model since it is able to distinguish between fracture and non-fracture shots, which contributes to the extent to which the model can be trusted when it is involved in real scenarios. Nonetheless, it is true that, the dataset size (700 images) is also quite small in comparison to large-scale medical imaging dataset, and this small size is a major limitation to generalization. We clearly indicate this as a limitation of our study, which contributes to the explanation of why performance is not as high as the very high levels of precision of optimized YOLOv8 variants that are reported in the literature.

Table 2: Dataset details and configuration for fracture detection.

Item	Value	Notes
Total X-ray images	650	GRAZPEDWRI-DX dataset and Hospital
Background images	50	Non-fracture, dual-mode interface

Preprocessing Parameters	Details	Value
Rotation	In the image rotation process, the image is rotated on a specified angle to make the model more robust and capable of handling objects at different locations. Its value ranges from 0-360	90
Translate	In translation object is shifted or moved within the image for better results during training and its value ranges from 0 to 1	0.1
Flipping images	Images are flipped along vertical and horizontal axes, so the model becomes more adaptable to accurately detecting objects regardless of their orientation	0.5
Classes	Ulna fracture; Radius fracture	Expert-annotated
Train / Val / Test	420 / 140 / 140	70/10/20 split
Input size	640×640	Standardized

3.2 Data Preprocessing

To train the YOLOv8 model, the images of wrist X-rays were made to conform to an identical size of 640×640 pixels according to the specifications of the provided architecture. The uniform resizing hereby resulted in the consistency of the dataset and saved processing during both the training and inference stages. Moreover, with the usage of different scales as well as some other improvement techniques including rotation, translation, flipping, and color variations to extend the dataset and improve model generalization.

Table 3: Data preprocessing parameters with details and values.

Table 3 describing details of data preprocessing parameters like rotation, translation, and flipping images is done in the preprocessing steps of data images, along with their values which are used for better results during training. Although these preprocessing and augmentation actions enhance robustness, the fact that they use resized and relatively low-resolution images is another weakness that can limit the capability of the model to recognize small fracture patterns. In our study, this limitation is clearly recognized. The overall, training and evaluation configuration is summarized in Table

Table 4: Training & Evaluation Setup.

3.2.1 Annotation

To train the model effectively for wrist fracture detection which was the focus, so intensively annotated the collected pictures with bounding boxes delineating the shape of the fracture precisely. The latter instance dealt with a detailed clinical study, which included qualified physiotherapist Dr. Komal Shahzadi for annotation of images with an annotation tool and resulting bounding boxes served as a clear indication to the model on what to concentrate on in the radiological imaging. As a result, the model accurately localized the regions of fracture.

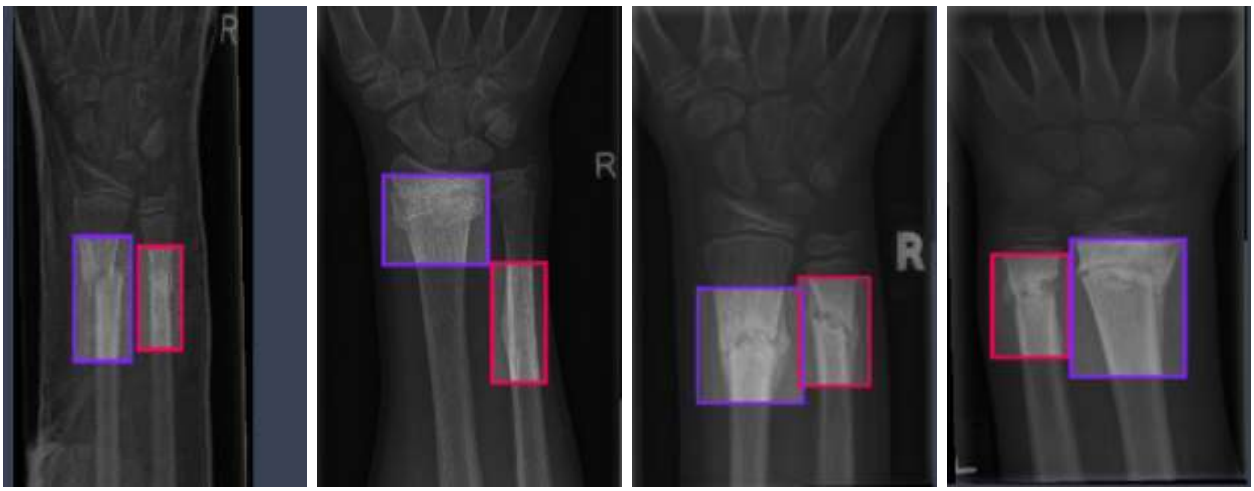

Fig. 3: Presenting annotated data of radius and ulna fracture: A professional approach to visualizing insights.

Figure 3 illustrates specific segments of the upcoming data. Moreover, it is also done using tools.

Item	Value	Notes
Detector	YOLOv8 (anchor-free head)	Selected for robust fracture detection
Input size	640×640	Standardized input dimension
Augmentations	Mosaic, Mix-up, Copy-Paste; flips; rotation; translation	Applied to improve generalization
Metrics	Accuracy, Precision, Recall, F1, AP, mAP	Used for model evaluation
Test protocol	Held-out test set (n=140)	Ensures unbiased performance assessment

In wrist fracture detection annotations play a very significant role they help to define a label for

the model which not only helps training but also for model validation, and model performance evaluation. They are the foundation for the development of effective and correct algorithms used for fracture detection, which in turn lead to the enhancement of the diagnosis and outcomes in patient care in clinical practice.

3.2.2 Data Splitting

Data splitting	Images Number
Training	420
Validation	140
Testing	140

Training, validation, and testing data sets were separated out of the overall dataset. Usually, about 70% are applied for training, 10% for validation, and then 20% is used for testing.

Table 5: Distribution of images for training, validation, and testing.

Table 5 explains those data that have been utilized for the purpose of training and testing. An effective method for the design and validation of neuron networks in a wrist fracture diagnosis application is the data splitting technique. It makes sure that the model is trained on a diverse set of examples; therefore it is able to generalize well for new cases and to provide true predictions when the model is tested with authentic data.

3.2.3 Image Resizing and Flipping

The Image Size parameter in preprocessing determines the dimensions of input images during training. By default, setting it to 640 pixels standardizes the image format for model training. Larger sizes can capture finer details but require more computational resources, so 640 pixels strike a balance between detail and efficiency. Flipping parameters facilitate horizontal and vertical flipping of images, creating mirror reflections that diversify the dataset. This augmentation technique aids the model in learning to recognize fractures from various orientations, enhancing its ability to generalize across different image perspectives.

3.2.4 Image Augmentation Parameters

The first step within our pipeline AutoHueSatValue(HueSaturationValue) that completely improves the dataset is in the sense that the system should overcome its general limitations. The cultural value of hsv h is established as 0.015, hsv s as 0.7, and hsv v as 0.4. color space is being changed this way to make the model scan more fractures on seeing the relative color and light.

Figure 4 shows augmented data which is used for fracture detection with the help of augmentation parameters. Apart from the classical approaches in data augmentation, the preprocessing pipeline that employs novel techniques will provide more data, improve the model generalization ability increase the fracture detection accuracy across more complicated environments, and support doctors make more accurate diagnoses.

Table 6 shows the specified values of hue saturation value. Hue shifts colors along the spectrum, saturation controls the color intensity, and value adjusts brightness.

Fig. 4: Presenting annotated data of radius and ulna fracture: A professional approach to visualizing insights.

Table 6: Understanding color: exploring hue, saturation, and value.

Parameter	Parameter Variable	Value
Hue	hsv h	0.015
Saturation	hsv s	0.7
Value	hsv v	0.4

3.2.5 Image Augmentation Techniques

Through the pre-processing pipeline, phenomena like ‘Mosaic’, ‘Mix-Up’, and ‘Copy- Pasting’ are used to enhance the dataset and enhance fracture detection. Mosaic is a combination of four unique images that take the viewer to diverse scenes. Mix-up stitches photographs and markings together in a linear fashion hence, it is good for augmentation of data during training. The neural network module Copy Paste takes areas from one image and then adds them to others, this procedure visualizes many different scenes. This way, the strategies lead to the enriched dataset, and thereby the model can learn intricate relations and cover-ups efficiently. Developing a specific

point of this YOLOv8 model does not only require a dataset but also considering carefully the parameters to extend the variety of data and ensure great accuracy in fracture detection.

• **Mosaic**: Mosaic enhancement turns four pictures into one, which makes training pictures more useful and varied.

• **Steps**: Pick four images at random from your information.

• **Resize and Place**: Make each picture smaller and put it in one of the four corners of a new picture frame. To create a mosaic image, begin by arranging the four pictures as follows: Picture 1 in the top left, Picture 2 in the top right, Picture 3 in the bottom left, and Picture 4 in the bottom right. Next, adjust the bounding boxes in each image to correctly position the objects in the mosaic. **Finally, combine the four pictures on a new board to form the complete mosaic image.**

• **Mixup:** MixUp takes two pictures and combines them by mixing their pixel values and titles.

• **Steps**: It will take two pictures and blend them with some coefficient values which are taken from Beta Distribution.

• **Blend Images and Labels:**

$$x\tilde{} = \lambda x_i + (1 - \lambda)x_j \; y\tilde{} = \lambda y_i + (1 - \lambda)y_j$$

where, x˜ and y˜ are the mixed image and label. xi and xj are two randomly selected images from the dataset. yi and yj are the corresponding labels of xi and xj. λ is a random value drawn from a Beta distribution, $\lambda \sim \text{Beta}(\alpha, \alpha)$, where α is a hyperparameter that determines the strength of the mix. This technique helps improve the model's generalization by generating new training samples through linear interpolation of pairs of images and their labels.

• **Copy Paste Augmentation:**

Steps:

To create a composite image for training, it starts by randomly selecting a source and a target image from our dataset. Identify objects in the source image to copy, then extract these objects along with their masks, if available. Paste the extracted objects onto random positions within the target image, ensuring they fit within its boundaries. Update the bounding boxes to reflect the new positions of the pasted objects. Finally, use this modified target image, now containing the pasted objects, for training purposes.

3.3 Proposed YOLOv8 Model

YOLOv8 transforms object detection into a fast single-pass architecture so an object can handle that instantaneous processing in real-time. By way of its uniform system, the tensor flow takes care of object bounding boxes and class probabilities in one pass which allows for high efficiency and accuracy. Localize objects through the usage of anchor boxes in YOLO v8, whether their sizes are small or large, and have a great effect on accurate object detection. Feature pyramid network (FPN) is another precise tool of the network that gives the cognition of multi-scale features which are very important to be able to see objects of varying sizes.

Model selection with rationale: We chose baseline YOLOv8 as the model for localizing ulna/radius fractures, because it is anchor-free and box assignment is easier than there are non-linear cortex edges. Also, its FPN is improved for sensitivity of small lesions on low-res films. Conversely, we would point to iYOLOv8 + GC (Amirouche et al, 2025) which reported 97.2% precision but only 67% f1. Moreover, and equally important, our baseline is best because balances case detection more successfully with smaller and more heterogeneous datasets. By recognizing when it is the case that baseline detection is clinically meaningful, baseline detection versus a better version can rely on understanding and calibrating benchmarks, in cases.

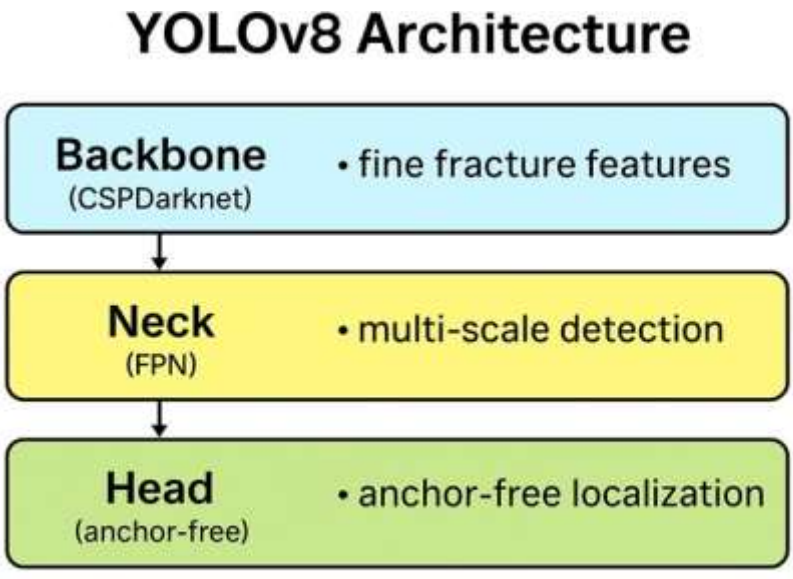

Fig. 5: Detailed architecture of the YOLOv8 model. The backbone, neck, and head are parts of the model, while the rest are modules of it [47].

3.3.2 Feature Pyramid Network

The FPN within YOLOv8 constructs a pyramid of feature maps, with each tier rep- resenting features at varying scales. Initially, features are extracted from the input image using a CNN

backbone CSPDarkNet53. These features undergo up-sampling and lateral connections at each pyramid level.

Up-sampling increases the resolution of feature maps to align with higher-level maps, while lateral connections facilitate information exchange between neighboring levels. This integration of detailed information and high-level semantic content allows the model to effectively detect objects of diverse sizes and shapes. The resulting multi-scale feature maps are then utilized for object detection, with subsequent layers predicting bounding boxes and class probabilities. Therefore, the FPN improves the object detection ability of YOLOv8 in relation to object size, allowing it to be balanced in its accuracy and flexibility to detect a wrist fracture without purporting to be better than other optimized models.

Figure 6 gives a detailed architecture of the Feature Pyramid Network, in which up-sampling is done for the feature that was obtained from the backbone.

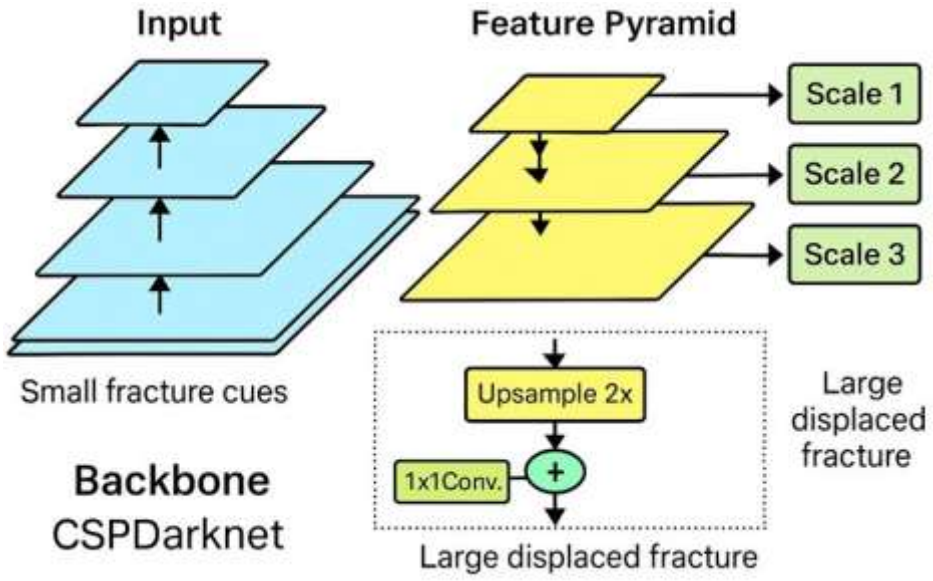

Fig. 6: Detailed architecture of FPN. The neck part of the YOLOv8 model which is used for up-sampling the features was obtained from the backbone [48].

4 Results and Discussion

This part of the work presents the results after the DL model and discusses the comparison with other models for wrist fracture detection. The outcomes are displayed as tables and these functions are evaluated on performance metrics such as accuracy, precision, recall, and f1-score.

4.1 Evaluation Metrics

An evaluation matrix is a structured framework used to assess the performance or quality of something based on specific criteria which are given as follows:

Precision refers to the accuracy of a measurement, indicating how close multiple measurements are to one another. It is calculated as the ratio of true positive predictions to the sum of true positive and false positive predictions.

Recall, also known as sensitivity or true positive rate, measures the model's ability to correctly identify positive instances from the total actual positive instances.

The importance of the F1 score is that it compresses both precision and recall into one number and so it is a comprehensive metric of the model's performance. One of the important advantages is that it effectively corrects the unbalanced class situation in data (class imbalance).

The Intersection over Union (IoU) measures the overlap between an anchor box and a ground-truth box, expressed as the ratio of their intersection area to their union area. IoU values range from 0.0 to 1.0. An IoU of 0.0 indicates no overlap, while an IoU of 1.0 signifies that the two boxes are perfectly aligned and identical. As the boxes become more similar, the IoU value increases accordingly.

The evaluation of an object detection model involves calculating the average precision (AP) for each class of objects within the dataset and subsequently averaging these AP scores across all classes. This process provides a comprehensive measure of the model's accuracy across different object categories.

4.2 Results

Following thorough training and evaluation, the YOLOv8 model showcases strong performance in detecting various types of wrist fractures. It achieves high levels of accuracy and precision, even across different imaging settings and fracture patterns. According to Table 7, our YOLOv8 model had an accuracy of 0.87, precision of 0.87, recall 0.88 and F1-score of 0.86, indicating comparable and consistent performance in all measures. These equitable values indicate that though the model is not better than the best precision rates in the reported optimized iYOLOv8 versions, the model gives stable accuracy, recall, and F1-scores, which can be considered clinically viable in general fracture detection tasks. These values indicate a low miss-rate (recall 0.88) without too many false alarms (precision 0.87). This allows busy EDs to triage questionable films to radiologists more readily while not overstating superiority. Our baseline YOLOv8 provides a more balanced result when compared to iYOLOv8 + GC by Amirouche et al. (2025) which is more precise (97.2) but has a significantly lower F1-score (67). This means that optimized models can be good at one measure, but can compromise the overall stability, but our methodology proves to be stable when

considered by several measures of assessment. In practice, these metrics make YOLOv8 a promising baseline model to be used by radiologists and other medical workers who specialize in wrist injuries. It has a balanced detection capability that can enable rapid diagnosis and treatment planning, which can improve patient care. We note though that these findings are constrained by the size of the data sets (600 images), restriction of image resolutions and the fact that we are using YOLOv8 as a baseline and not a more optimized version.

Evaluation parameter	Score
Accuracy	0.87
Precision	0.87
Recall	0.88
F1 score	0.86

When considered by several measures of assessment. In practice, these metrics make YOLOv8 a promising baseline model to be used by radiologists and other medical workers who specialize in wrist injuries. It has a balanced detection capability that can enable rapid diagnosis and treatment planning, which can improve patient care. We note though that these findings are constrained by the size of the data sets (600 images), restriction of image resolutions and the fact that we are using YOLOv8 as a baseline and not a more optimized version.

Table 7: Evaluating YOLOv8 model performance: metrics for assessment.

Table 7 showcases the exceptional performance of our YOLOv8 model in accurately identifying ulna and radius fractures. With an accuracy score of 0.87, the model demonstrates proficiency in fracture classification, which is vital for diagnosing wrist injuries. Moreover, its high precision (0.87) indicates a minimal rate of false positives, while a recall of 0.88 signifies its ability to identify a substantial number of true posi- tives. The balanced F1 score of 0.86 underscores its consistent and accurate fracture detection.

Practically, these metrics validate the YOLOv8 model's potential as a valuable tool for radiologists and healthcare practitioners specializing in wrist injuries. Its precise and reliable fracture detection facilitates expedited diagnosis, and treatment planning, and ultimately, enhances patient care. Overall, the outstanding performance of the YOLO v8 model underscores its importance in medical imaging, particularly for ulna and radius fractures in wrist injuries.

Figure 7 shows the YOLOv8 model's precision for precise ulna and radius fracture detection. The model achieves a 0.87 precision score, which shows its performance metrics are well. It demonstrates a very high accuracy level in the identification of True Positive cases and it also minimizes the rate of False Positives. Therefore, this protocol's capability highlights its ability to distinguish and localize areas with and without fracture in wrist X-ray pictures which significantly contribute to the process of clinical diagnosis for developing an appropriate clinical plan.

In Figure 8, the recall outcomes from the YOLO v8 model illustrate its capability to accurately detect ulna and radius fractures. With a recall score of 0.88, the model show cases strong sensitivity by identifying around 88% of genuine fractures while reducing false negatives. This highlights its effectiveness in thorough fracture detection across various imaging scenarios and categories, ultimately enhancing diagnostic accuracy in clinical settings.

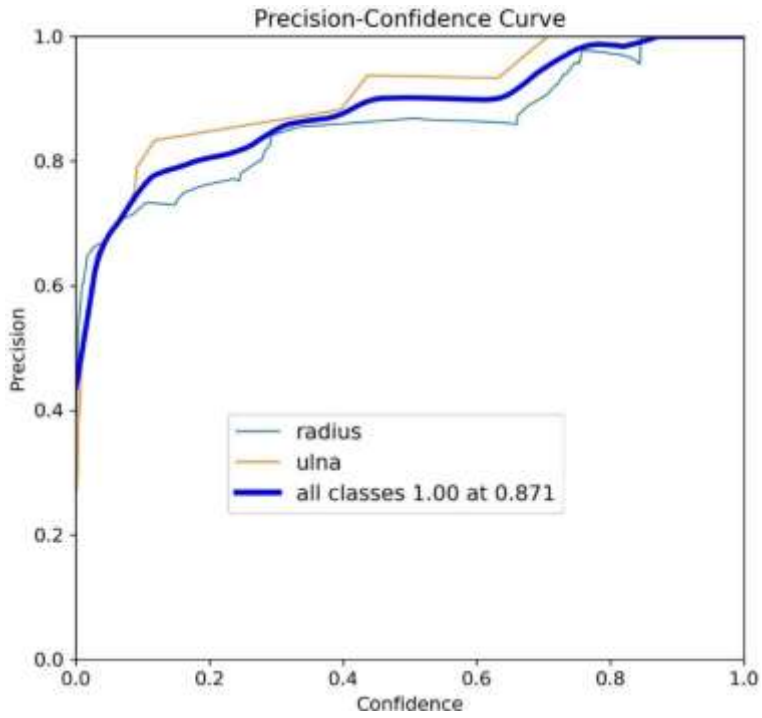

Fig. 7: Visualizing precision: confidence curve graph for YOLOv8 model analysis.

Figure 9, the F1 score, which is a combination of precision, and recall, determines how the YOLOv8 model performs in detecting ulna and radius fractures. Having achieved the F1 score of 0.86 the model has met the task of avoiding false positives and vice versa education of true positives, thus preventing mistakes in the diagnosis of fractures. This score may reflect the stability as well as the quality of the model, which is in turn considered as a prerequisite for clinicians in clinical practice to achieve better diagnostic processes and quality of patient care.

Figure 10 shows the detection of ulna and radius fracture with bounding boxes on the fractured part. It is worth mentioning that this is a feature that has not been observed in most of the previous papers.

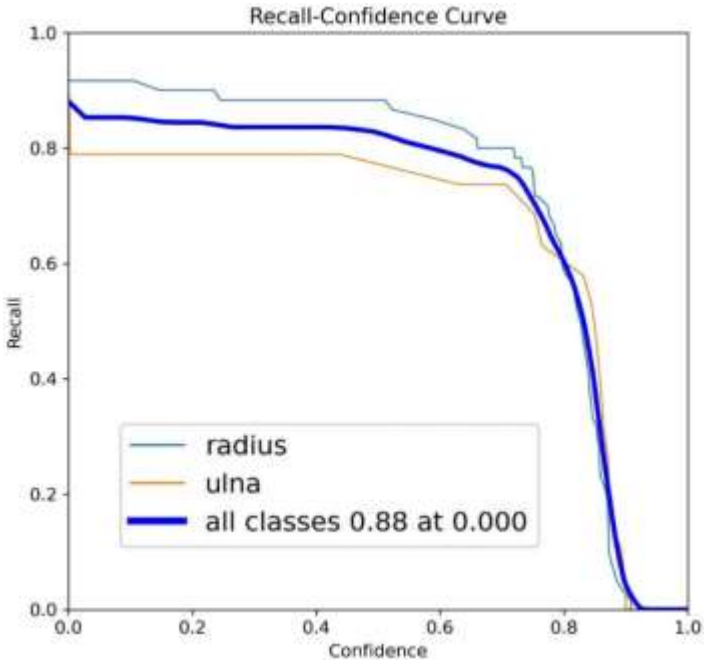

Fig. 8: Analyzing recall: confidence curve graph for YOLOv8 model assessment.

4.3 Comparative Analysis

Comparing the performance of different models is crucial for identifying the most effective approach. Scientists can appraise the effectiveness of their methods including accuracy, precision, recall, and F1-score through rigorous testing. This can reveal the strengths and weaknesses of the methods to researchers and practitioners. In the analysis presented below, YOLOv8 proves to be equally effective against other models, achieving 87 percent accuracy, and the values of precision, recall, and F1-scores are consistent. The next in line is the deep CNN (DCNN) [3] that data-wise drew about 84.38% accuracy and proved highly consistent in precision and recall as well. After this Faster R-CNN [42] performed very well on given data, its accuracy was 82%, another model fine mask recurrent neural network (RCNN) [37] with an accuracy of 83% performed better than faster RCNN in detecting and localizing the ulna and radius fractures. All these models were used already on different datasets but their accuracy was low as compared to the dataset on which they are trained now. Some of the above models were only used for the classification of fracture without localization of fracture and other few models were used for delineation of wrist bones not for fracture detection. Thus, the comparison is based on trade-offs among various approaches, and the YOLOv8 provides balanced detection on various metrics instead of superiority.

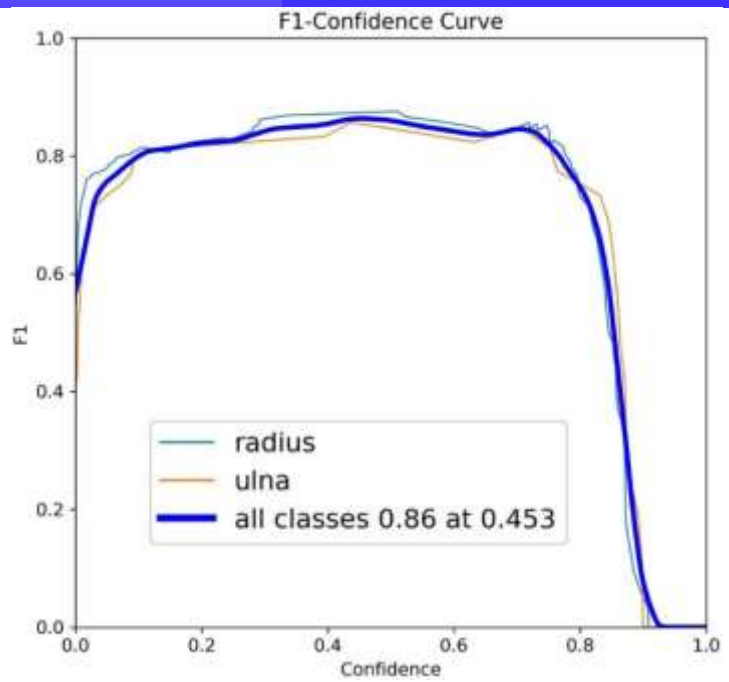

Fig. 9: Measuring F1 Score: confidence curve graph for YOLOv8 model evaluation.

Table 8: Comparative Analysis: YOLOv8 model versus other models concerning pre- cision, recall, average precision, mAP, etc.

Model	Accuracy	Precision	Recall	F1- Score	AP	mAP
DCNN	84.38%	0.875	0.840	0.85	0.730	0.730
Faster R-CNN	82.00%	0.780	0.730	0.75	0.560	0.560
Fine Mask RCNN	83.00%	0.800	0.820	0.80	0.656	0.656
YOLOV8	87.00%	0.870	0.880	0.86	0.760	0.760

Table 8 shows YOLO V8 leads with 87% accuracy, 87% precision, 88% recall,86% F1-score,76% average precision, and 76% mean average precision indicating equaliza- tion in detection of various metrics. DCNN follows at 84.38% accuracy, with 87.5% precision, 84% recall, 85% F1-score 73% average precision, and 73% mean average pre- cision.More to the point, the baseline YOLOv8 can be contrasted with the optimized version of iYOLOv8 + GC by Amirouche et al. (2025) with 97.2% precision and 67% F1-score compared to our variant that displays more balanced and consistent behavior. In that way, YOLOv8 is to be treated with cautious eye on a

prospective contribu- tion to the standards, but not the most efficient or ultimate solution. Its trade-off in terms of accuracy, precision, recall, and F1-score is balanced, which facilitates its clinical applicability in detecting wrist fractures, although optimized models can prioritize one of the measurements at the expense of another.

In addition to the numerical comparison provided in Table 8, graphical representations in Figure 12 offer a visual understanding of the performance differences among the evaluated models. Bar charts or radar charts can effectively illustrate the varying performance metrics, including accuracy, precision, and recall, for each model.

Bar charts display the accuracy, precision, and recall scores of each model side by side, allowing for easy comparison. Each model is represented by a distinct color, and the height of the bars indicates the corresponding metric value. This visual representation enables stakeholders to quickly identify which models excel in specific metrics and where trade-offs may occur. The bar chart is a great tool for getting a better under- standing of general performance characteristics at once, by comparing each model metric by metric. Each provided model is depicted by a polygon whose vertices are at the accuracy, precision, and recall values. The shape and size of the polygons played constitute the fact that shows the overall performance and shape of each model. It is digital that models with higher scores in all metrics will show more symmetrical and broader polygons which is the only medium to denote their high performances in overall performance.

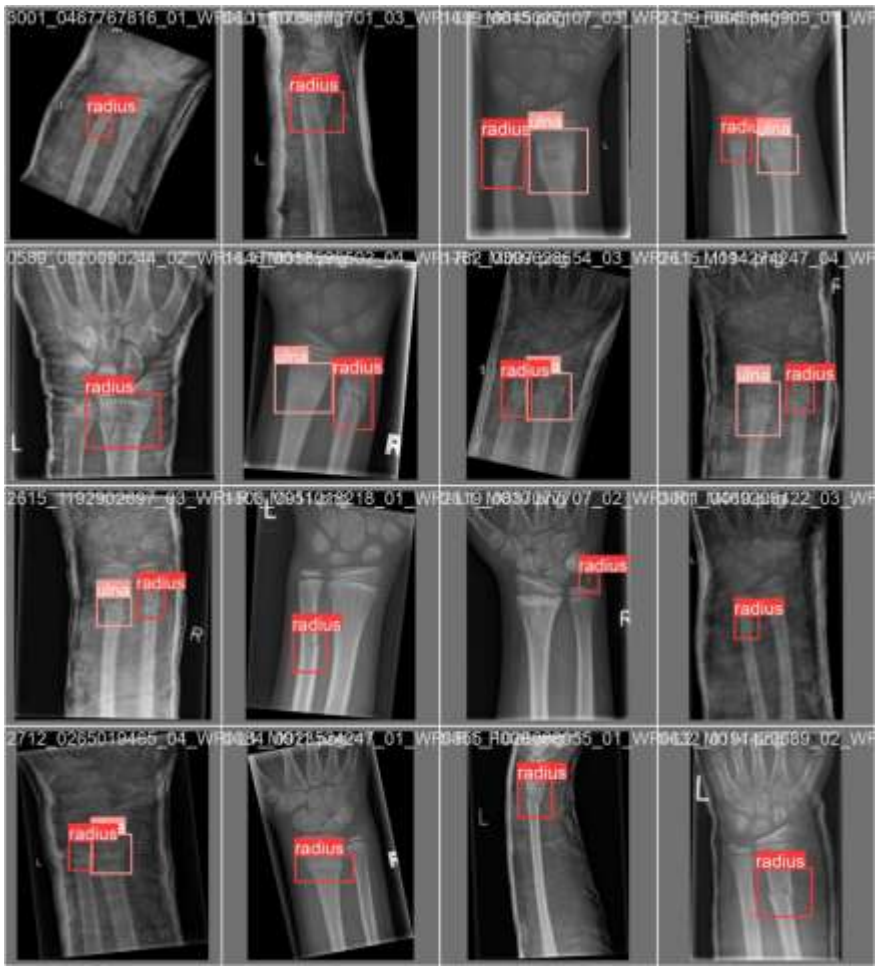

Fig. 10: Enhanced Diagnosis: YOLOv8 model output displays bounding boxes with Ulna and Radius classification for fracture detection.

Fig. 11: Enhanced Diagnosis: YOLOv8 model output displays bounding boxes with Ulna and Radius classification for fracture detection.

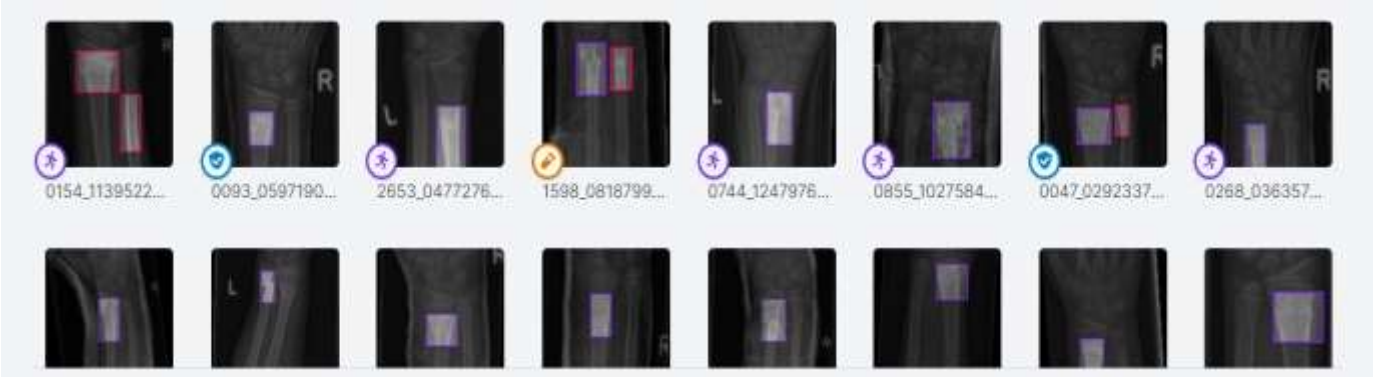

Fig. 12: Visualizing Model Performance: comparative graphical analysis of YOLOv8 with other models.

4.4 Limitation

1. Dataset size diversity – only 600 images, limited generalizability.

2. Resolution - resizing to 640×640 may have masked micro-fractures.

3. Baseline scope - only YOLOv8 baseline, not optimized variants.

4. Evaluation – single held out split, no multicenter validation.

4.5 Future Work

In conclusion, the YOLOv8 neural network model demonstrates balanced and clinically relevant performance in the detection of ulna and radius fractures, achieving stable values across precision, recall, and F1-score. Instead of making it a decisive or better tool, we highlight its role as a foundational framework that can be used to aid the clinicians by helping them in the diagnosis and treatment planning of wrist fractures. Looking ahead, future work in this domain should focus on several directions. To begin with, the model architecture and hyper parameters optimization can be further refined, potentially leading to a subsequent increase in accuracy. Second, the strategy should be applied to other forms of wrist fractures besides radial and ulna to increase clinical applicability. Third, it is critical to expand the dataset to more and more diverse and larger-scale clinical data with better robustness and generalization. Moreover, multi-modal fusion (e.g. X-rays with CT or patient history) and prospective validation in the real clinical setting should also be integrated in order to increase the credibility and acceptance of the system. Integrating the explain ability and interpretability features might also enhance the clinician trust and present the system as a helpful diagnostic aid but not a substitution. With these purposes in mind, YOLOv8 can be considered as a promising base of ulna and radius fracture detection. It will have to undergo future improvement by adding data, improving this, and clinical validation before it can be incorporated fully into clinical decision support systems.

5. Conclusion

Accurate detection of human wrist fractures is helpful in determining rapid and appropriate treatment. The computer-aided detection has proven to be both rapid and accurate, particularly involving the use of deep learning models. A simple YOLOv8 model was created in this work that will be used to detect wrist fractures. The goal was to identify Ulna type and Radius type wrist in the form of bounding box localization. The model illustrates a balanced and clinically meaningful performance in the identification of wrist fractures, as opposed to purporting to be a conclusive and better solution. Improved evaluation metrics, including average precision and mean average precision, demonstrated the robustness of the YOLOv8 model in identifying Ulna and Radius fractures with accuracy, precision, recall, and F1 scores of 0.87, 0.87, 0.88, and 0.86, respectively, indicating its proficiency in accurate fracture detection. The YOLOv8 model achieved high scores

of 0.76 and 0.76 for average precision and mean average precision, respectively, compared to deep CNN, Faster R-CNN, and Fine

REFERENCES

1. Hardalac¸, F., Uysal, F., Peker, O., C¸ ic¸eklida˘g, M., Tolunay, T., Tokg¨oz, N., Kut- bay, U., Demirciler, B., Mert, F.: Fracture detection in wrist x-ray images using deep learning-based object detection models. Sensors 22(3), 1285 (2022)

2. Rashid, T., Zia, M.S., Meraj, T., Rauf, H.T., Kadry, S.: A minority class balanced approach using the dcnn-lstm method to detect human wrist fracture. Life 13(1), 133 (2023)

3. Umadevi, N., Geethalakshmi, S.: Multiple classification system for fracture detec- tion in human bone x-ray images. In: 2012 Third International Conference on Computing, Communication and Networking Technologies (ICCCNT'12), pp. 1–8 (2012). IEEE

4. Khan, S. R., Raza, A., Waqas, M., & Raphay Zia, M. A. (2024). Efficient and Accurate Image Classification via Spatial Pyramid Matching and SURF Sparse Coding. Lahore Garrison University Research Journal of Computer Science and Information Technology, 7(4).

5. Greenspan, H., Van Ginneken, B., Summers, R.M.: Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. IEEE transactions on medical imaging 35(5), 1153–1159 (2016)

6. M. Waqas, Z. Khan, S. U. Ahmed and Asif. Raza, "MIL-Mixer: A Robust Bag Encoding Strategy for Multiple Instance Learning (MIL) using MLP-Mixer," 2023 18th International Conference on Emerging Technologies (ICET), Peshawar, Pakistan, 2023, pp. 22-26.

7. Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G., Thrun, S., Dean, J.: A guide to deep learning in healthcare. Nature medicine 25(1), 24–29 (2019)

8. Hekmat, A., et al., Brain tumor diagnosis redefined: Leveraging image fusion for MRI enhancement classification. Biomedical Signal Processing and Control, 2025. 109: p. 108040.

9. O. Bilal, Asif Raza, S. ur R. Khan, and Ghazanfar Ali, "A Contemporary Secure Microservices Discovery Architecture with Service Tags for Smart City Infrastructures ", VFAST trans. softw. eng., vol. 12, no. 1, pp. 79–92, Mar. 2024

10. M. Waqas, Z. Khan, S. U. Ahmed and Asif. Raza, "MIL-Mixer: A Robust Bag Encoding Strategy for Multiple Instance Learning (MIL) using MLP-Mixer," 2023 18th International Conference on Emerging Technologies (ICET), Peshawar, Pakistan, 2023, pp. 22-26.

11. Khan, M.A., Khan, S.U.R. & Lin, D. Shortening surgical time in high myopia treatment: a randomized controlled trial comparing non-OVD and OVD techniques in ICL implantation. BMC Ophthalmol 25, 303 (2025). https://doi.org/10.1186/s12886-025-04135-3

12. Mahmood, F., Abbas, K., Raza, A., Khan,M.A., & Khan, P.W. (2019). Three Dimensional Agricultural Land Modeling using Unmanned Aerial System (UAS). International Journal of Advanced Computer Science and Applications (IJACSA) [p-ISSN : 2158-107X, e-ISSN : 2156-5570], 10(1).

13. Waqas, M., Bandyopadhyay, R., Showkatian, E., Muneer, A., Zafar, A., Alvarez, F. R., ... & Wu, J. (2025). The Next Layer: Augmenting Foundation Models with Structure-Preserving and Attention-Guided Learning for Local Patches to Global Context Awareness in Computational Pathology. arXiv preprint arXiv:2508.19914.

14. Khan, U. S., Ishfaque, M., Khan, S. U. R., Xu, F., Chen, L., & Lei, Y. (2024). Comparative analysis of twelve transfer learning models for the prediction and crack detection in concrete dams, based on borehole images. Frontiers of Structural and Civil Engineering, 1-17.

15. Bilal, Omair, Arash Hekmat, Inzamam Shahzad, Asif Raza, and Saif Ur Rehman Khan. "Boosting Machine Learning Accuracy for Cardiac Disease Prediction: The Role of Advanced Feature Engineering and Model Optimization." The Review of Socionetwork Strategies (2025): 1-30.

16. Khan, S. U. R., & Khan, Z. (2025). Detection of Abnormal Cardiac Rhythms Using Feature Fusion Technique with Heart Sound Spectrograms. Journal of Bionic Engineering, 1-20.

17. Khan, Saif Ur Rehman, Asif Raza, Inzamam Shahzad, and Ghazanfar Ali. "Enhancing concrete and pavement crack prediction through hierarchical feature integration with VGG16 and triple classifier ensemble." In 2024 Horizons of Information Technology and Engineering (HITE), pp. 1-6. IEEE, 2024.

18. Muneer, A., Waqas, M., Saad, M. B., Showkatian, E., Bandyopadhyay, R., Xu, H., ... & Wu, J. (2025). From Classical Machine Learning to Emerging Foundation Models: Review on Multimodal Data Integration for Cancer Research. arXiv preprint arXiv:2507.09028.

19. Khan, S.U.R., Asif, S., Bilal, O. et al. Lead-cnn: lightweight enhanced dimension reduction convolutional neural network for brain tumor classification. Int. J. Mach. Learn. & Cyber. (2025). https://doi.org/10.1007/s13042-025-02637-6.

20. Khan, Z., Hossain, M. Z., Mayumu, N., Yasmin, F., & Aziz, Y. (2024, November). Boosting the Prediction of Brain Tumor Using Two Stage BiGait Architecture. In 2024 International Conference on Digital Image Computing: Techniques and Applications (DICTA) (pp. 411-418). IEEE.

21. S.ur R. Khan, Asif. Raza, Muhammad Tanveer Meeran, and U. Bilhaj, "Enhancing Breast Cancer Detection through Thermal Imaging and Customized 2D CNN Classifiers", VFAST trans. softw. eng., vol. 11, no. 4, pp. 80–92, Dec. 2023.

22. Waqas, M., Tahir, M. A., Al-Maadeed, S., Bouridane, A., & Wu, J. (2024). Simultaneous instance pooling and bag representation selection approach for multiple-instance learning (MIL) using vision transformer. Neural Computing and Applications, 36(12), 6659-6680.

23. Khan, S. U. R., Asif, S., Zhao, M., Zou, W., Li, Y., & Xiao, C. (2026). ShallowMRI: A novel lightweight CNN with novel attention mechanism for Multi brain tumor classification in MRI images. Biomedical Signal Processing and Control, 111, 108425.

24. Khan, M. A., Khan, S. U. R., Rehman, H. U., Aladhadh, S., & Lin, D. (2025). Robust InceptionV3 with Novel EYENET Weights for Di-EYENET Ocular Surface Imaging Dataset: Integrating Chain Foraging and Cyclone Aging Techniques. International Journal of Computational Intelligence Systems, 18(1), 1-26.

25. Khan, U. S., & Khan, S. U. R. (2025). Ethics by Design: A Lifecycle Framework for Trustworthy AI in Medical Imaging From Transparent Data Governance to Clinically Validated Deployment. arXiv preprint arXiv:2507.04249.

26. Khan, S. U. R., Asif, S., Zhao, M., Zou, W., Li, Y., & Xiao, C. (2026). ShallowMRI: A novel lightweight CNN with novel attention mechanism for Multi brain tumor classification in MRI images. Biomedical Signal Processing and Control, 111, 108425.

27. Waqas, M., Tahir, M. A., & Qureshi, R. (2023). Deep Gaussian mixture model based instance relevance estimation for multiple instance learning applications. Applied intelligence, 53(9), 10310-10325.

28. Al-Khasawneh, M. A., Raza, A., Khan, S. U. R., & Khan, Z. (2024). Stock Market Trend Prediction Using Deep Learning Approach. Computational Economics, 1-32.

29. HUSSAIN, S., Raza, A., MEERAN, M. T., IJAZ, H. M., & JAMALI, S. (2020). Domain Ontology Based Similarity and Analysis in Higher Education. IEEEP New Horizons Journal, 102(1), 11-16.

30. Khan, S.U.R., Zhao, M. & Li, Y. Detection of MRI brain tumor using residual skip block based modified MobileNet model. Cluster Comput 28, 248 (2025). https://doi.org/10.1007/s10586-024-04940-3

31. Shahzad, Inzamam, Asif Raza, and Muhammad Waqas. "Medical Image Retrieval using Hybrid Features and Advanced Computational Intelligence Techniques." Spectrum of engineering sciences 3, no. 1 (2025): 22-65.

32. Raza, A., & Meeran, M. T. (2019). Routine of Encryption in Cognitive Radio Network. Mehran University Research Journal of Engineering and Technology [p-ISSN: 0254-7821, e-ISSN: 2413-7219], 38(3), 609-618.

33. Khan, S. U. R., Asim, M. N., Vollmer, S., & Dengel, A. (2025). FloraSyntropy-Net: Scalable Deep Learning with Novel FloraSyntropy Archive for Large-Scale Plant Disease Diagnosis. arXiv preprint arXiv:2508.17653.

34. Meeran, M. T., Raza, A., & Din, M. (2018). Advancement in GSM Network to Access Cloud Services. Pakistan Journal of Engineering, Technology & Science [ISSN: 2224-2333], 7(1).

35. Khan, S. U. R., Rehman, H. U., & Bilal, O. (2025). AI-powered cancer diagnosis: classifying viable (live) vs non-viable (dead) cells using transfer learning. Signal, Image and Video Processing, 19(15), 1326.

36. Bilal, O., Hekmat, A., Shahzad, I. et al. Boosting Machine Learning Accuracy for Cardiac Disease Prediction: The Role of Advanced Feature Engineering and Model Optimization. Rev Socionetwork Strat (2025). https://doi.org/10.1007/s12626-025-00190-w

37. Asif Raza, Inzamam Shahzad, Ghazanfar Ali, and Muhammad Hanif Soomro. "Use Transfer Learning VGG16, Inception, and Reset50 to Classify IoT Challenge in Security Domain via Dataset Bench Mark." Journal of Innovative Computing and Emerging Technologies 5, no. 1 (2025).

38. Khan, M. A., Khan, S. U. R., Rehman, H. U., Aladhadh, S., & Lin, D. (2025). Robust InceptionV3 with Novel EYENET Weights for Di-EYENET Ocular Surface Imaging Dataset: Integrating Chain Foraging and Cyclone Aging Techniques. International Journal of Computational Intelligence Systems, 18(1), 204.

39. Khan, Z., Khan, S. U. R., Bilal, O., Raza, A., & Ali, G. (2025, February). Optimizing Cervical Lesion Detection Using Deep Learning with Particle Swarm Optimization. In 2025 6th International Conference on Advancements in Computational Sciences (ICACS) (pp. 1-7). IEEE.

40. Raza, Asif, Inzamam Shahzad, Muhammad Salahuddin, and Sadia Latif. "Satellite Imagery Employed to Analyze the Extent of Urban Land Transformation in The Punjab District of Pakistan." Journal of Palestine Ahliya University for Research and Studies 4, no. 2 (2025): 17-36.

41. Khan, S. U. R., Asif, S., Zhao, M., Zou, W., Li, Y., & Li, X. (2025). Optimized deep learning model for comprehensive medical image analysis across multiple modalities. Neurocomputing, 619, 129182.

42. Asif Raza, Salahuddin, Ghazanfar Ali, Muhammad Hanif Soomro, Saima Batool, "Analyzing the Impact of Artificial Intelligence on Shaping Consumer Demand in E-Commerce: A Critical Review", International Journal of Information Engineering and Electronic Business(IJIEEB), Vol.17, No.5, pp. 42-61, 2025. DOI:10.5815/ijieeb.2025.05.04

43. M. Wajid, M. K. Abid, A. Asif Raza, M. Haroon, and A. Q. Mudasar, "Flood Prediction System Using IOT & Artificial Neural Network", VFAST trans. softw. eng., vol. 12, no. 1, pp. 210–224, Mar. 2024.

44. Khan, S.U.R., Raza, A., Shahzad, I., Khan, S. (2025). Subcellular Structures Classification in Fluorescence Microscopic Images. In: Arif, M., Jaffar, A., Geman, O. (eds) Computing and Emerging Technologies. ICCET 2023. Communications in Computer and Information Science, vol 2056. Springer, Cham. https://doi.org/10.1007/978-3-031-77620-5_20

45. Maqsood, H., & Khan, S. U. R. (2025). MeD-3D: A Multimodal Deep Learning Framework for Precise Recurrence Prediction in Clear Cell Renal Cell Carcinoma (ccRCC). arXiv preprint arXiv:2507.07839.

46. Khan, S. U. R., Asif, S., Zhao, M., Zou, W., & Li, Y. (2025). Optimize brain tumor multiclass classification with manta ray foraging and improved residual block techniques. Multimedia Systems, 31(1), 1-27.

47. Raza, A., Salahuddin, & Inzamam Shahzad. (2024). Residual Learning Model-Based Classification of COVID-19 Using Chest Radiographs. Spectrum of Engineering Sciences, 2(3), 367–396.

48. Khan, S. U. R., Asif, S., & Bilal, O. (2025). Ensemble Architecture of Vision Transformer and CNNs for Breast Cancer Tumor Detection from Mammograms. International Journal of Imaging Systems and Technology, 35(3), e70090.

49. Raza, A., Soomro, M. H., Shahzad, I., & Batool, S. (2024). Abstractive Text Summarization for Urdu Language. Journal of Computing & Biomedical Informatics, 7(02).

50. Hekmat, A., Zuping, Z., Bilal, O., & Khan, S. U. R. (2025). Differential evolution-driven optimized ensemble network for brain tumor detection. International Journal of Machine Learning and Cybernetics, 1-26.

51. Rehman, Laiba, Humera Batool Gill, Sohaib Ahmad, and Ezza Irshad. "FROM MOLECULES TO MIND: SIMULATING THE FUTURE OF DATA STORAGE WITH DNA, GRAPHENE AND NEURAL ENCODING." Kashf Journal of Multidisciplinary Research 2, no. 10 (2025): 45-60.

52. Waqas, M., Ahmed, S. U., Tahir, M. A., Wu, J., & Qureshi, R. (2024). Exploring multiple instance learning (MIL): A brief survey. Expert Systems with Applications, 250, 123893.

53. Khan, S. U. R. (2025). Multi-level feature fusion network for kidney disease detection. Computers in Biology and Medicine, 191, 110214.

54. Khan, S. R., Asif Raza, Inzamam Shahzad, & Hafiz Muhammad Ijaz. (2024). Deep transfer CNNs models performance evaluation using unbalanced histopathological breast cancer dataset. Lahore Garrison University Research Journal of Computer Science and Information Technology, 8(1).

55. Aslam, N., Meeran, M. T., Aslam, M., Maqbool, M. S., & Saeed, B. (2025). UNDERSTANDING URBAN EXPANSION THROUGH MULTI-TEMPORAL SATELLITE DATA ANALYSIS. Kashf Journal of Multidisciplinary Research, 2(09), 252-273.

56. Hiyat, Muhammad, Muhammad Usama Javed, Maria Akhtar, and Muhammad Tanveer Meeran. "A NOVEL SEMI-SUPERVISED FRAMEWORK FOR CYBERSECURITY THREAT DETECTION IN WIRELESS SENSOR NETWORKS." Kashf Journal of Multidisciplinary Research 2, no. 10 (2025): 1-16.

57. Bilal, O., Hekmat, A., & Khan, S. U. R. (2025). Automated cervical cancer cell diagnosis via grid search-optimized multi-CNN ensemble networks. Network Modeling Analysis in Health Informatics and Bioinformatics, 14(1), 67.

58. Asim, M. N., Ibrahim, M. A., Malik, M. I., Dengel, A., & Ahmed, S. (2022). LGCA-VHPPI: A local-global residue context aware viral-host protein-protein interaction predictor. Plos one, 17(7), e0270275.

59. Khan, S. U. R., Asim, M. N., Vollmer, S., & Dengel, A. (2025). Robust & precise knowledge distillation-based novel context-aware predictor for disease detection in brain and gastrointestinal. arXiv preprint arXiv:2505.06381.

60. N. Mayumu, D. Xiaoheng, P. Mukala, S. U. R. Khan and M. U. Saeed, "Omni-V2X: A Vision-Language Model for Actionable Insights in Vehicle-to-Everything Systems," 2025 International Joint Conference on Neural Networks (IJCNN), Rome, Italy, 2025, pp. 1-8, doi: 10.1109/IJCNN64981.2025.11228491.

61. Wasim, M., Mahmood, W., Asim, M. N., & Khan, M. U. (2018). Multi-label question classification for factoid and list type questions in biomedical question answering. IEEE Access, 7, 3882-3896.

62. Khalid, Ariba, Muhammad Usama Javed, Muhammad Tanveer Meeran, Naeem Aslam, and Muhammad Fuzail. "TOWARDS SAFE AND SECURE URBAN TRANSPORTATION INTRUSION DETECTION SYSTEM FOR CONNECTED VEHICLES IN SMART CITIES." Kashf Journal of Multidisciplinary Research 2, no. 10 (2025): 17-35.

63. Yang, H., Khan, S. U. R., Bilal, O., Chen, C., & Zhao, M. (2025). CEOE-Net: Chaotic Evolution Algorithm-Based Optimized Ensemble Framework Enhanced with Dual-Attention for Alzheimer's Diagnosis. Computer Modeling in Engineering & Sciences, 145(2), 2401.

64. S. Ur Rehman Khan, O. Bilal, S. Mistry, N. Deb, M. Mahmud and M. Bhuyan, "KDLight: A Lightweight Knowledge Distillation Framework for Medical Image Classification," *2025 International Joint Conference on Neural Networks (IJCNN)*, Rome, Italy, 2025, pp. 1-8, doi: 10.1109/IJCNN64981.2025.11228615.

65. O. Bilal, S. Ur Rehman Khan, S. Mistry, N. Deb, M. Mahmud and M. Bhuyan, "Towards Efficient Pruning and Multi-Scale Feature Transformations to Uncover Medical Diseases," *2025 International Joint Conference on Neural Networks (IJCNN)*, Rome, Italy, 2025, pp. 1-8, doi: 10.1109/IJCNN64981.2025.11229047.