A Comprehensive Review of Fracture Detection and Identification from Medical Images using Deep Learning Methods

Praveen Kumar T and Dr.Suthendran K

Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education University
Virudhunagar,Tamil Nadu, India
Email : tpraveenkumarphd2023@gmail.com

Professor and Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education University
Virudhunagar,Tamil Nadu, India
Email : k.suthendran@klu.ac.in

Abstract:

The detection and identification of fractures have significant importance in the field of medical imaging, as they play a crucial role in facilitating precise diagnoses and prompt development of treatment strategies. Over the course of the last decade, there has been a notable progression in deep learning methods, demonstrating their effectiveness in a range of medical image analysis, such as fracture diagnosis and classification. This review article offers a full and extensive examination of recent advancements in using deep learning techniques for the purpose of detecting and identifying fractures from medical imaging. This study conducts a comprehensive examination of the prevailing methodology, commonly used datasets, inherent obstacles, and probable future trajectories within this rapidly developing field. The objective of this work is to provide significant insights into the transformational effects of deep learning on the interpretation of medical images linked to fractures. This research seeks to contribute to the improvement of patient care and diagnosis by examining the influence of deep learning in this field.

Keywords: CNN, Fracture, Deep learning, Binary classification, Multiclass classification.

1. Introduction:

The field of contemporary healthcare is significantly enhanced by the use of medical imaging, a remarkable technical advancement that reveals the complexities of the interior structure and operations of the human body [1]. The identification and detection of fractures play a fundamental role in diagnostic efforts within the broad range of medical imaging applications. The rapid and precise identification of fractures has considerable consequences, as it plays a fundamental role in clinical decision-making, development of suitable treatment approaches, and the ultimate objective of enhancing patient outcomes. In recent times, a significant transformation has occurred in this field, driven by the integration of deep learning algorithms with medical imaging tools [2]. This introductory section initiates a comprehensive review by delving into the significant importance of fracture detection, exploring the limitations of traditional methodologies, and highlighting the emergence of deep learning as a powerful tool in the field of medical image analysis.

\

1.1 Significance of Fracture Detection in Medical Imaging:

The human body, characterised by its complexity and beauty, often experiences fractures as a result of unexpected incidents, traumatic occurrences, or pre-existing medical disorders. In the present setting, the prompt and precise detection of fractures acquires great importance. In addition to serving as a significant diagnostic milestone, the discovery of a condition or disease has a substantial impact on the provision of patient care, influencing the course of therapies and prognostic outcomes [3]. The prompt highlights the need of promptly diagnosing fractures in order to tailor therapies to the unique characteristics and location of the fracture. This approach helps to alleviate pain, minimise possible consequences, and accelerate the healing process. However, the consequences of this extend beyond the immediate results mentioned, as they also involve areas such as surgery planning, rehabilitation measures, and the potential for promoting long-term well-being [4]. Therefore, while the field of medical imaging continues to advance, the skill of identifying fractures maintains its crucial position in the realm of diagnostic procedures and patient-centered healthcare.

1.2 Limitations of Traditional Fracture Detection Methods:

Historically, the responsibility of identifying fractures has been entrusted to radiologists, who possess the expertise and visual acuity necessary to interpret medical imaging. However, it should be noted that this typical technique is not impervious to restrictions. The act of human interpretation introduces a certain level of subjectivity, hence creating an opportunity for possible variations among various observers and interpreters [5]. The presence of diverse evaluation methods may significantly impact clinical decision-making, perhaps leading to notable effects on patient outcomes. Furthermore, the labour-intensive process of picture interpretation is inherently time-consuming, which may be a challenge in critical medical scenarios when prompt treatments are required [6]. The convergence of these variables highlights a clear and urgent need for innovation - a need for methods that go beyond personal biases, reduce inconsistencies, and accelerate the process of diagnosing.

1.3 Emergence of Deep Learning in Medical Image Analysis:

In recent times, there has been a notable rise in the popularity and use of deep learning techniques, leading to a significant transformation in the field of medical image analysis. Deep learning, which falls under the umbrella of artificial intelligence, demonstrates remarkable expertise in the areas of autonomous feature extraction and pattern identification from large datasets [7]. Amongst the many approaches used, Convolutional Neural Networks (CNNs) stand out conspicuously due to their highly regarded ability to recognise subtle elements inside medical pictures. By engaging in the thorough examination of vast datasets, these algorithms acquire the ability to differentiate with exceptional accuracy between typical anatomical structures and instances of fractures [8]. The integration of deep learning with medical image analysis presents a promising convergence, offering a wide range of potential applications and serving as a powerful alternative to conventional methodologies. During this significant period of change, the field of fracture detection is characterised by the potential for achieving consistency, immediacy, and accuracy.

2. Fundamentals of Deep Learning:

This section delves into the foundational aspects of deep learning, offering a comprehensive understanding of its architecture and its specific applications in medical imaging. By establishing these fundamentals, we can better appreciate how deep learning revolutionizes fracture detection methodologies. The Figure 1 shows the architecture of deep learning.

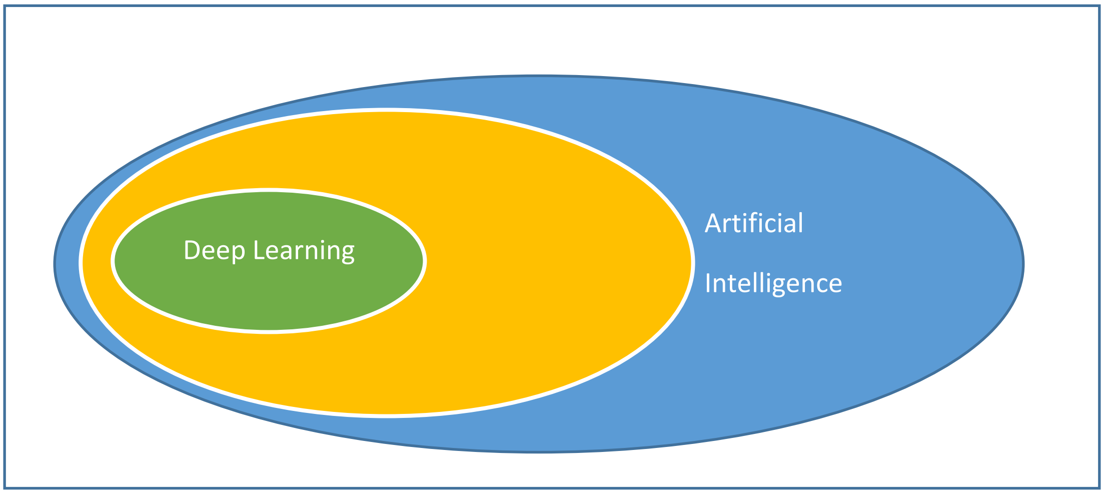

Fig 1. Architecture of Deep learning

2.1 Deep Learning Architecture Overview:

Deep learning, an essential component of artificial intelligence, mimics the complex neural networks seen in the human brain in order to uncover complicated patterns within convoluted datasets [9]. These networks, similar to the neurons seen in the human brain, function collectively to analyse and interpret fundamental patterns within the given data. A crucial aspect in grasping the transformational potential of deep learning models lies in the comprehension of their architectural structure.

Architecture Overview:

The structural composition of deep learning models may be likened to a series of linked layers, each possessing a distinct function in the process of converting input data into meaningful and interpretable outputs. These layers cover the following:

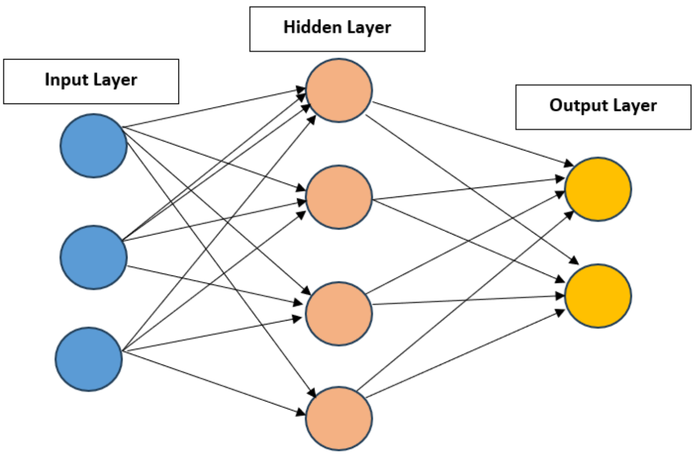

Fig 2, Structure of Deep learning [10]

Input Layer: The input layer is responsible for receiving the unprocessed data, such as photos, text, or other types of data, and transmitting it to the next layers for further processing.

Hidden Layer: The hidden layers play a crucial role in the extraction of more intricate characteristics from the input data. The process involves the acquisition and enhancement of data representations via the application of mathematical changes at each layer.

Output Layer: The output layer is responsible for generating the model's predictions or classifications, using the learnt characteristics. The structure of the job varies depending on its unique nature, such as whether it involves binary classification, multi-class classification, or regression [10].

Activation Functions:

Activation functions play a vital role in the functioning of individual neurons within a neural network. Non-linearity is included into the model, hence facilitating the representation of complex connections present in the data. Activation functions such as Rectified Linear Unit (ReLU) or sigmoid play a crucial role in the transformation of a neuron's output, hence defining its activation and subsequent effect on succeeding layers. The presence of non-linearity inside the model allows for the approximation of detailed patterns and complicated functions that may not be readily discernible in linear contexts.

Weight Initialization and Optimization:

The proper initialization of weights and the selection of optimisation methods play a crucial role in guaranteeing the efficient convergence of the model towards meaningful solutions. Weight initialization approaches such as Xavier/Glorot initialization or He initialization are used to assign appropriate starting weights, hence facilitating efficient information propagation across the network during the training process. Optimisation methods such as Adam or stochastic gradient descent repeatedly modify the weights of a model in order to minimise the loss function and improve overall performance. The learning process is enhanced by these algorithms by the modification of the model's parameters, which is determined by the gradients computed in each iteration.

2.2 Convolutional Neural Networks (CNNs) for Medical Imaging:

Convolutional Neural Networks (CNNs) have emerged as a fundamental cornerstone of deep learning in the field of medical imaging owing to its remarkable ability to identify and analyse spatial characteristics present in pictures. These networks effectively mimic the hierarchical processing of the human visual system, which has important implications for medical picture analysis, namely in the field of fracture identification.

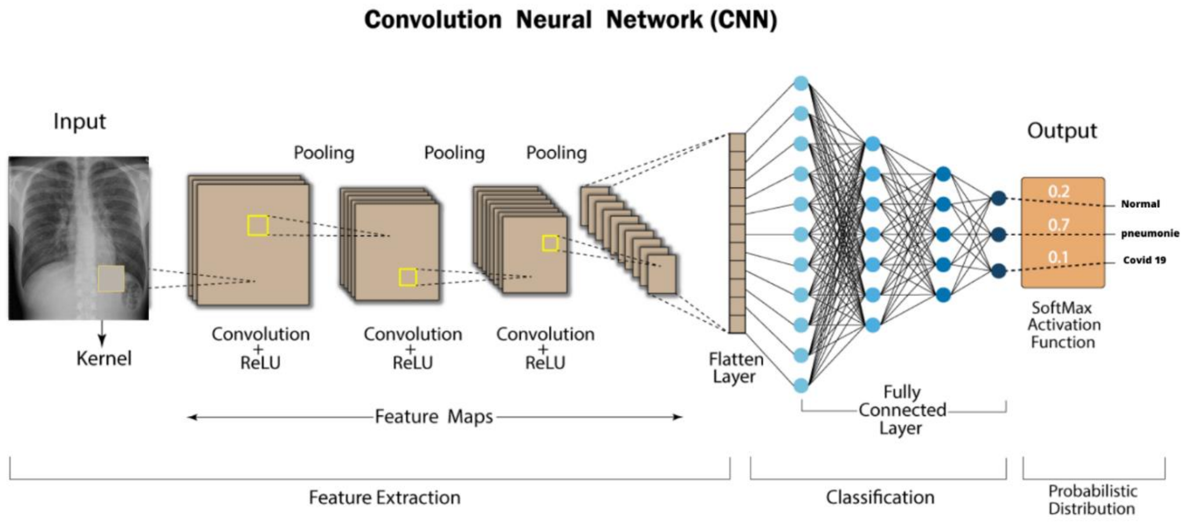

Fig. 3 CNNs for Medical Imaging [11]

Hierarchical Perception and Local Patterns:

Convolutional neural networks (CNNs) function by examining pictures in a way that resembles the visual processing of the human brain. They use the concept that pictures are composed of hierarchical layers of characteristics, ranging from basic edges to more complex patterns. Convolutional layers used in Convolutional Neural Networks (CNNs) make use of filters, also known as kernels, to perform scanning operations on pictures, hence facilitating the identification of localised patterns such as edges, textures, or forms [12]. The filters are applied to the input picture in a sliding manner, with the purpose of identifying and capturing certain elements that have a role in the overall visual context.

Pooling for Contextual Information:

Pooling layers are essential components of convolutional neural networks (CNNs) since they are responsible for gathering and incorporating wider contextual information. The aforementioned layers perform downsampling on the feature maps, so retaining crucial information while simultaneously decreasing the spatial dimensions. One example of a technique used in deep learning models is max pooling, which involves selecting the most significant feature inside a certain area. This process enhances the model's ability to handle fluctuations in spatial placement, hence improving its robustness [13]. The use of this pooling method enables Convolutional Neural Networks (CNNs) to effectively capture crucial information irrespective of their specific spatial locations within the picture.

Adeptness in Capturing Bone Structures and Nuances:

Convolutional Neural Networks (CNNs) demonstrate exceptional performance in the field of medical imaging, particularly in scenarios where the identification of fractures relies on the accurate analysis of complex bone structures and minor variations. The capability of individuals to acquire knowledge and discern diverse bone patterns, including both normal and broken conditions, demonstrates their aptitude for revealing essential diagnostic information [14]. Convolutional neural networks (CNNs) have the ability to detect even the most minute deviations that may indicate the presence of a fracture by examining the differences in pixel intensity, texture, and form.

2.3 Transfer Learning and Pretrained Models:

Transfer learning has emerged as a significant method in the field of medical imaging, particularly in situations where data availability is a limitation. Transfer learning is a technique that utilises prior knowledge from a relevant activity or domain to augment the process of learning in a novel task. This segment provides an in-depth exploration of the notion of transfer learning and its significant ramifications in the field of medical image analysis[15]. Pretrained models, which are often trained on large datasets such as ImageNet, provide transferable characteristics that may be refined for targeted medical imaging applications, such as fracture identification. This methodology not only expedites the convergence of the model but also improves its performance, especially in cases when there is a scarcity of labelled medical picture datasets. Medical practitioners may effectively implement fracture detection systems with lower data requirements by using pretrained models. This approach allows for the use of robust and precise algorithms.

3. Datasets and Annotations:

This section explores the pivotal significance of datasets and annotations in the use of deep learning techniques for fracture identification. This paper presents a comprehensive examination of the medical picture datasets used in this particular context, with a specific focus on the difficulties and intricacies involved in data pre-processing and annotations.

3.1 Overview of Medical Image Datasets for Fracture Detection:

The presence of well selected and organised medical picture datasets plays a crucial role in the advancement and education of precise deep learning models for the identification of fractures. This article offers a comprehensive review of the pivotal datasets that have significantly contributed to the advancement of this field. The datasets under consideration include MURA, NIH Chest X-rays, and RSNA Bone Age, which provide significant contributions to the understanding of fracture diagnosis in medical pictures.

Fig 4. Fracture Images

Dataset Characteristics and Modalities:

The MURA dataset, developed by the National Institutes of Health (NIH), is a widely used dataset in the field of medical. The datasets of Chest X-rays and RSNA Bone Age consist of a wide variety of imaging modalities, mostly X-rays and sometimes CT scans. These modalities function as a medium for recording bone architecture, fractures, and other abnormalities, facilitating the creation of models with diagnostic capabilities.

Diversity of Fracture Cases:

These datasets are characterized by their diversity, encapsulating a spectrum of fracture cases that span various anatomical regions, age groups, and clinical contexts. The inclusion of different fracture patterns, such as wrist fractures, femur fractures, and vertebral fractures, enriches the dataset's complexity [16]. The datasets also encompass various age groups, which is vital given that fracture patterns can vary considerably across different stages of life.

Size and Impact:

The size of the datasets vary, but, when combined, they provide a significant amount of data for both training and evaluating models. The influence of their work in this domain is significant, since it establishes a fundamental basis for evaluating the effectiveness of deep learning models in the context of fracture diagnosis [17]. Scholars and professionals use these datasets in order to evaluate the applicability, responsiveness, and precision of their models, hence enhancing the overall precision and clinical significance of the area.

3.2 Data Pre-processing and Annotation Challenges:

The process of preparing medical image data for fracture identification necessitates the implementation of extensive pre-processing procedures, owing to the inherent diversity and complexity associated with medical pictures. This part examines the difficulties that arise in the process of data preparation. The need for using picture normalisation, scaling, and noise reduction methods is discussed in order to establish a consistent format for input photographs and enhance the performance of the model [18]. Moreover, the procedure of annotation, which requires accurate labelling of fractures, poses many problems. The intrinsically subjective characteristic of fracture annotation necessitates specialised expertise, which may result in heterogeneity in labelling. The limited availability of annotated data and the labour-intensive process of annotation provide further challenges.

4. Methodologies for Fracture Detection:

This section explores the strategies used in fracture diagnosis via the application of deep learning algorithms. This paper examines many methodologies used in medical image analysis, ranging from binary classification to more complex localization and segmentation methods [19]. It aims to provide a comprehensive understanding of the subtle distinctions and benefits associated with each strategy within this field.

4.1 Binary classification models:

Binary classification models serve as the fundamental components for the identification of fractures via the use of deep learning methodologies. This part provides a comprehensive analysis of these models, elucidating their architectural design, functioning, and significant contribution to the field of medical imaging [20].

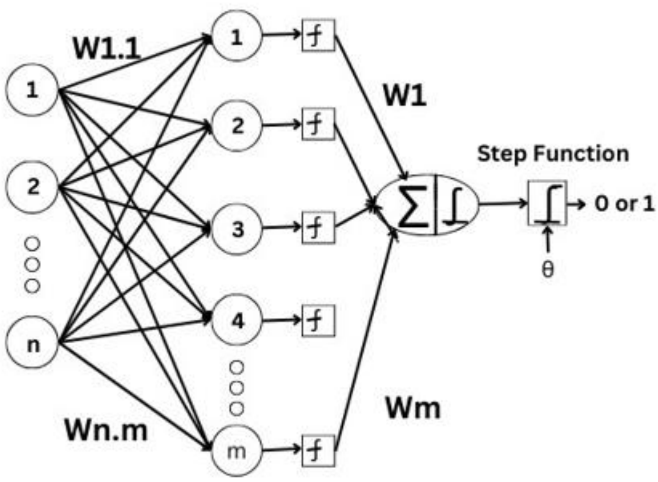

Fig 5. Architecture of Binary Classification [20]

Architecture and Operation:

Fundamentally, binary classification models are specifically developed to classify medical pictures into two distinct categories: fractured and non-fractured. The architectural design of these models often involves the use of convolutional neural networks (CNNs), which demonstrate exceptional proficiency in extracting detailed characteristics from pictures. The use of convolutional and pooling layers in CNNs facilitates the model's ability to identify distinctive patterns and forms that differentiate intact bone structures from those that have had fractures. The successive layers of the model gradually capture more advanced variables, allowing for a deeper understanding of the distinct attributes related to fractures.

Tailoring CNNs for Fracture Detection:

Convolutional neural networks (CNNs) have a high level of suitability for the task of fracture identification, mostly owing to their inherent ability to extract hierarchical features [21]. The ability to identify minor variations in skeletal patterns, irrespective of their spatial position within the picture, is present in these individuals. Convolutional Neural Networks (CNNs) acquire the ability to differentiate between normal and broken bone structures via the process of training on labelled data, which provides information on whether an image has a fracture or not. Consequently, individuals may extrapolate this acquired knowledge to novel and unfamiliar pictures, so making them efficacious instruments for diagnostic purposes.

Role of Labelled Data:

The efficacy of binary classification models is contingent upon the availability of accurately labelled data of high quality. The use of annotated photos, in which each occurrence is categorised as either fractured or non-fractured, allows the model to acquire knowledge about the distinctive characteristics that are associated with fractures [22]. The presence of a wide range of data that is appropriately categorised and labelled is of utmost importance in ensuring the precision and applicability of the model.

4.2 Multiclass Classification Models for Fracture Identification:

Progressing beyond binary classification, this subsection plunges into the nuanced realm of multiclass classification models tailored for fracture identification. The intricacies of fracture identification involve categorizing fractures into specific types, such as wrist fractures, femur fractures, or vertebral fractures [23]. The section unveils how deep learning models are adapted to tackle these complexities and facilitate the recognition of multiple fracture categories.

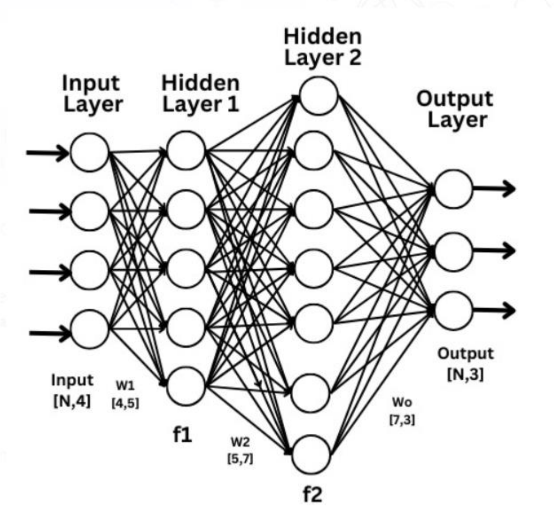

Fig 6. Architecture of Multi Class model [23]

Categorizing Diverse Fracture Types:

The process of fracture diagnosis offers an additional degree of complexity as it necessitates models to classify fractures into discrete categories. Every form of fracture has distinct features, which highlights the importance of the model accurately capturing the numerous details that distinguish them. In contrast to the binary classification approach that primarily distinguishes between normal and fractured instances, multiclass classification requires a broader learning framework to incorporate the many kinds of fractures.

Adaptations for Multiclass Classification:

Deep learning models demonstrate a high level of proficiency in effectively adjusting to the complexities associated with multiclass classification tasks. A pivotal approach is the use of softmax activation in the output layer. The Softmax function is used to transform the initial outputs of a model into probabilities for each class. This transformation allows the model to not only detect fractures but also generate a probability distribution across various fracture kinds. Furthermore, the use of one-hot encoding enables the conversion of category labels into binary vectors, hence simplifying the representation of several classes.

4.3 Localization and Segmentation Approaches:

Reaching the pinnacle of sophistication in fracture detection, this subsection ventures into the intricate realm of localizing fractures within medical images and achieving pixel-level segmentation of fractured regions [24]. These techniques represent a profound advancement, enabling not only fracture identification but also precise localization and mapping of the fractured areas.

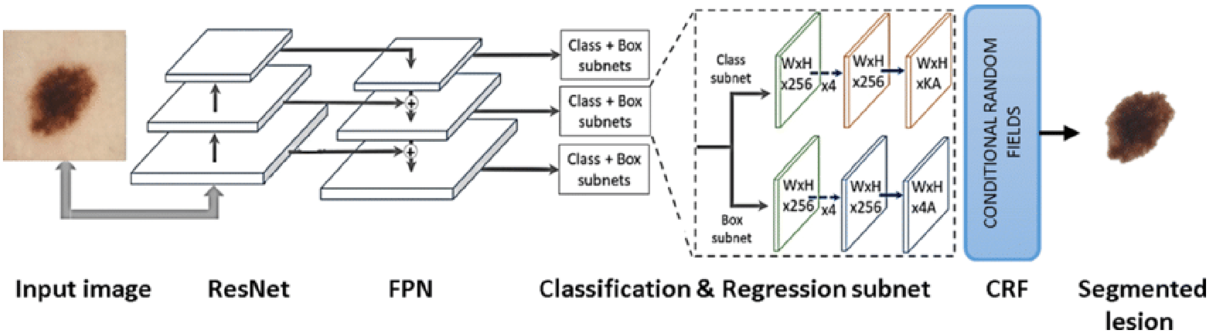

Fig 7. Architecture of Localization and Segmentation [24]

Fracture Localization:

Localization techniques are designed to pinpoint the exact location of fractures within images. Utilizing approaches like bounding boxes or heatmaps, these techniques provide a clear spatial indication of the fracture's position. Bounding boxes encapsulate the area surrounding the fracture, offering a region of interest for clinicians. Heatmaps visualize the intensity of fracture presence, effectively highlighting the fractured region within the image.

Fracture Segmentation:

Taking sophistication a step further, segmentation techniques go beyond localization to achieve pixel-level accuracy in isolating fractured regions. These techniques meticulously delineate the boundaries of the fractured area, allowing clinicians to precisely understand the extent of the fracture. By segmenting fractured areas at the pixel level, these techniques provide invaluable insights for treatment planning and assessment.

Architectures like U-Net:

Deep learning architectures like U-Net are exemplary in the domain of image segmentation. U-Net employs an encoder-decoder architecture that learns to map input images to pixel-level segmentations. It excels at capturing fine details and intricate patterns within images, making it particularly potent for identifying and segmenting fractured regions.

Deep learning in radiology/orthopaedic traumatology

The deep learning's application in radiology underscore its potential advantages for both fracture detection and characterization tasks, as demonstrated in the aforementioned instances (Figure 3).

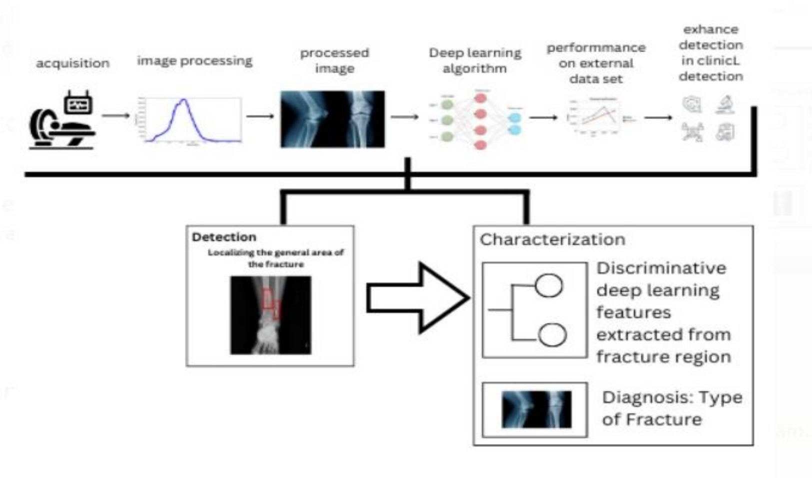

Fig 8. Deep learning aided workflow in fracture detection [25]

The integration of deep learning algorithms into radiology practise shows potential in enhancing diagnostic accuracy and efficiency, as well as reducing the burden on radiologists by automating labour-intensive responsibilities. Nevertheless, it is crucial to acknowledge that deep learning models are not impervious to the intrinsic variability seen in human-based diagnoses, which encompasses inconsistencies between different observers as well as within the same observer. In research-oriented settings, deep learning has the potential to achieve comparable or superior performance to humans in tasks related to fracture identification and classification in X-rays and CT scans. This implies that the use of deep learning has the potential to improve the diagnostic capabilities of radiologists and orthopaedic traumatologists. Although deep learning has significant promise, it is important to acknowledge the importance of acknowledging its limits and the possibility of mistakes when using it. The optimal use of deep learning in fracture-related activities within the fields of radiology and orthopaedic traumatology relies heavily on striking a balance between AI help and human competence.

Author/year	Modality	Methods	Performance Metrics	Key features	Limitations
Tanushree Meena et al. 2023 [31]	X-ray	InceptionNet (CNN-based)	Accuracy, Sensitivity	High accuracy in X-ray fracture detection, Automation	Limited investigation of various disorders; outside variables affecting the quality of X-ray images
Dhirendra. et. al 2022 [32]	X-ray	SFNet + Canny Edge Algorithm	Accuracy: 99.12%, F1-Score: 99%, Recall: 100%	Multi-scale feature fusion, Sequential DL, Improved Canny	Insufficient variety of real-world data; algorithm susceptible to picture noise
Firat hardalac et.al 2022 [33]	Wrist X-ray	Object Detection Models	AP Score, LRP	Clinical dataset, 10 object detection models	Sensitivity to changes in imaging circumstances and restricted applicability to a variety of datasets
Yu Jin Seol et al. 2022 [34]	CT scan	3D-ResNet34 and ResNet50 deep learning	AUC: 94.5% for 3D-ResNet34, 93.4% for ResNet50	Automatic diagnosis of nasal fractures, High AUC scores	Deep learning models' limited interpretability and resource-intensive computing
Barhoom et al 2020 [35]	X-ray	VGG16-based deep learning	Precision: 85.96%	Utilization of Mura-v1.1 dataset with 42,000 X-rays	Problems with generalization with a variety of patient demographics; possible bias in the dataset used
Tomita et al 2018 [36]	CT	Deep convolutional neural network	Accuracy: 89.2%, F1 score: 90.8%	High precision, Deep learning for CT image analysis	supervised learning's reliance on labeled data and its susceptibility to changes in imaging methods
Sahin et.al (2023) [37]	X-ray	Image pre-processing, Canny, Sobel, Hough	Accuracy, Training Time, Testing Time, AUC (LDA)	LDA achieved 88.67% accuracy, 0.89 AUC for CAD	Limited resistance to noise and artifacts; reliance on methods for pre-processing images
Firtz et al (2023) [38]	MRI and CT	Radiomics, Machine Learning, Deep Learning	Accuracy range 0.72	Diagnosis of various musculoskeletal conditions	Feature selection dependence; possible overfitting in intricate models
Ahmed et al (2023) [39]	CT	Naïve Bayes, Decision Tree, Nearest Neighbours, Random Forest, SVM	Accuracy range 0.92	Application in Medicine, Seismology, Remote Sensing	Sensitivity to dataset imbalances and certain techniques' computational complexity
Paul et al., (2023) [40]	CT scan	MobileNetV2 with augmentation	Accuracy: 99.75%	Smartphone Android application for decision-making	Limited research on a range of patient demographics; possible prejudice in smartphone-assisted diagnosis

Lin et al (2023) [41]	CT scan	DCNN	Accuracy: 93.8%	Utilization of modified transfer and deep learning methods	Deep learning models' limited interpretability and possible transferability issues
Kassem et al (2023) [42]	X ray	Explainable artificial intelligence (XAI)	Accuracy: 98.5%	Assisting doctors in disease diagnosis, Trustworthiness	complicated XAI models' interpretability trade-offs and possible shortcomings in capturing complicated medical decision-making
Mutasa et al., 2020	X-ray (AP hip radiographs)	CNN, GAN, DRR	AUC, Accuracy, Sensitivity, Specificity, PPV, NPV	Garden fracture classification, Two-class and Three-class prediction	Limited to femoral neck fractures, may not generalize to other fractures.
Krogue et al., 2020	Hip and pelvic radiographs	Deep learning-based object detection, DenseNet	Binary accuracy, Sensitivity, Specificity, Multiclass classification accuracy	Automated placement of bounding boxes, Detection and classification of hip fractures	Limited to hip fractures, potential challenges in real-world scenarios.
Raisuddin et al., 2021	Wrist radiographs	Convolutional Neural Networks	Average precision, AUC	Deep Wrist pipeline, General population test set, Challenging test set	Reduced performance on challenging cases, emphasizes the need for careful model analysis.
Lindsey et al., 2018	Radiographs	Deep neural network	Sensitivity, Specificity, Reduction in misinterpretation rate	Fracture detection, Comparison with human observers, Improved diagnostic accuracy	Emphasis on aiding human performance, potential for improved patient care.
Kim et al., 2018	Lateral wrist radiographs	Inception v3 network, Transfer learning	AUC, Sensitivity, Specificity	Transfer learning from non-medical images, Fracture classification	Proof of concept for transfer learning in fracture detection, potential for clinical applications.
Chung et al., 2018	Anteroposterior shoulder radiographs	Deep convolutional neural network	Top-1 accuracy, AUC, Sensitivity, Specificity, Youden index	Proximal humerus fracture detection and classification	Superior performance compared to general physicians and orthopedists, potential clinical applicability.
Tomita et al., 2018	Chest, abdomen, and pelvis CT scans	Deep convolutional neural network	Accuracy, F1 score	Osteoporotic vertebral fracture detection	Comparable performance to practicing radiologists, potential for clinical assistance.
Urakawa et al., 2019	Proximal femoral radiographs	Convolutional neural network	Accuracy, Sensitivity, Specificity	Intertrochanteric hip fracture detection	Exceeded orthopedic surgeons' performance under limited conditions, potential for screening in emergency settings.

Cheng et al., 2019	Pelvic Radiographs (PXR)	DCNN, Grad-CAM	Accuracy, Sensitivity, False-Negative Rate, AUC	Accuracy: 91%, Sensitivity: 98%, False-Negative Rate: 2%, AUC: 0.98	Limited to the specific dataset used for training and evaluation. Generalizability to diverse populations and imaging conditions needs further validation.
Arif et al., 2018	X-ray	Deep Fully Convolutional Neural Network (Spinal Region Localization)	Dice Similarity Coefficient: 0.84	Automatic localization of the spinal region in X-ray images using deep learning.	High complexity
Meena et al., 2022	Various bone imaging modalities	DL in traumatology and orthopaedics	-	Demonstrated potential of DL in diagnosing fractures and diseases	Specific studies in traumatology and orthopaedics not mentioned, general over vi
Thian et al., 2019	Wrist Radiographs	Object Detection CNN (Inception-ResNet Faster R-CNN)	Per-image Sensitivity: 95.7%, Specificity: 82.5%, AUC: 0.918 (Frontal View)	Demonstrated high sensitivity and specificity in detecting radius fractures	Limited to wrist radiographs,
Burns et al., 2017	CT Images	Automated Machine Learning System	Accuracy (Genant type): 95.0% (95% CI: 0.89, 0.98), Weighted κ: 0.90 (95% CI: 0.81, 0.99)	Accurate classification by Genant type with high weighted κ,	Limited accuracy, Potential variability in bone attenuation measurement.

5. Performance Evaluation Metrics:

This section delves into the essential metrics used to evaluate the performance of deep learning models in fracture detection. It offers a comprehensive understanding of metrics ranging from basic accuracy to more nuanced measures that provide deeper insights into a model's performance [26].

5.1 Accuracy, Sensitivity, Specificity, and Precision:

These fundamental metrics serve as the cornerstone of model evaluation. The subsection begins by elucidating accuracy, the ratio of correctly predicted fractures to total predictions. Sensitivity (recall) quantifies the proportion of actual fractures correctly identified by the model. Specificity gauges the ability of the model to accurately identify non-fractured cases.

Precision denotes the ratio of true positive predictions to all positive predictions, providing a measure of the model's reliability in labelling fractures. The significance of these metrics lies in their ability to provide a holistic view of model performance, including its ability to detect both fractures and non-fractures.

5.2 Receiver Operating Characteristic (ROC) Curves and AUC:

The Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) offer insights into a model's performance across different thresholds. This subsection navigates through ROC curves, illustrating the trade-off between sensitivity and specificity as the threshold for positive predictions varies. AUC quantifies the curve's performance, with higher values indicating better discrimination ability. The section elaborates on the significance of ROC and AUC in evaluating fracture detection models' overall effectiveness, even in cases where class imbalance is prevalent.

5.3 Intersection over Union (IoU) for Segmentation:

In the context of segmentation tasks, such as fracture area delineation, the Intersection over Union (IoU) metric becomes pivotal. Also known as the Jaccard index, IoU measures the overlap between the predicted segmentation mask and the ground truth mask. This subsection dissects IoU's calculation, emphasizing its significance in assessing the accuracy of fracture region delineation. It discusses how IoU's intuitive interpretation, ranging from 0 to 1, reflects the degree of overlap between predicted and actual fractured regions.

6. Challenges in Deep Learning for Fracture Detection:

This section delves into the challenges and hurdles encountered when deploying deep learning techniques for fracture detection within medical images. It navigates through the intricacies of limited annotated data, class imbalance, domain adaptation, and the imperative for model interpretability [27].

6.1 Limited Annotated Data and Data Augmentation:

The scarcity of meticulously annotated medical image data is a recurring challenge in training accurate deep learning models. This subsection explores the complexities of this challenge and the solutions it has spawned. Data augmentation emerges as a strategy to amplify the available dataset by applying transformations such as rotations, flips, and scaling. The section delves into the efficacy of data augmentation in enhancing model generalization and robustness, bridging the gap between the limited available data and the complex patterns deep learning models need to comprehend.

6.2 Class Imbalance and Domain Adaptation:

Class imbalance, where one class (e.g., fractured) is significantly underrepresented compared to the other (e.g., non-fractured), poses a unique set of challenges in fracture detection. This subsection dissects the repercussions of class imbalance on model performance, including biased predictions towards the majority class. Strategies like oversampling, under sampling, and the introduction of class weights are discussed to counteract this imbalance. Moreover, the challenge of domain adaptation arises when models trained on one dataset struggle to perform optimally on another dataset due to variations in imaging conditions and characteristics. Techniques like domain adaptation and transfer learning are explored as potential solutions to address this challenge.

6.3 Interpretability and Explain ability of Deep Learning Models:

As deep learning models become more complex, their decisions can appear as "black-box" outcomes, raising questions about their trustworthiness and clinical adoption. This subsection delves into the challenge of model interpretability and explain ability. It elucidates techniques such as Grad-CAM (Gradient-weighted Class Activation Mapping) that provide insights into the regions of interest influencing the model's predictions. These approaches aim to bridge the gap between sophisticated models and their comprehensibility by clinicians, fostering trust and adoption in real-world medical scenarios.

7. Clinical Implementation and Impact:

This section delves into the real-world clinical implications and transformative potential of integrating deep learning models for fracture detection into the healthcare landscape. It explores the integration process, impact on radiologist workload, ethical considerations, and patient privacy concerns [28].

7.1 Integration of Deep Learning Models into Clinical Workflow:

The integration of deep learning models into clinical practice is a pivotal step towards harnessing their potential. This subsection dissects the process of seamlessly incorporating fracture detection models into existing clinical workflows. It explores the interoperability of deep learning systems with Picture Archiving and Communication Systems (PACS) and Radiology Information Systems (RIS). Furthermore, the importance of integrating these models as decision support tools, rather than replacements for radiologists, is highlighted. The section also discusses the training required for healthcare practitioners to effectively utilize these tools.

7.2 Reduced Radiologist Workload and Improved Diagnosis Speed:

Deep learning's ability to swiftly and accurately detect fractures bears profound implications for radiologists' workload and patient care. This subsection delves into how deep learning systems can expedite the diagnosis process, providing rapid insights into fracture presence or absence. The potential for reducing radiologist fatigue and improving overall efficiency is explored. Additionally, the augmentation of radiologists' diagnostic prowess through AI-assisted decision-making is discussed, highlighting the collaborative potential between human expertise and AI capabilities.

7.3 Ethical Considerations and Patient Privacy:

Introducing deep learning systems into clinical practice necessitates addressing ethical and privacy concerns. This subsection delves into the ethical considerations surrounding the integration of AI into healthcare. It discusses issues related to algorithm transparency, accountability, and potential biases. Patient privacy concerns, particularly related to medical image data sharing and storage, are explored in the context of deep learning systems. The importance of complying with regulatory standards like HIPAA (Health Insurance Portability and Accountability Act) and GDPR (General Data Protection Regulation) is emphasized.

8. Future Directions and Emerging Trends:

This section peers into the horizon of future possibilities and emerging trends in the domain of fracture detection using deep learning [29]. It explores the integration of multimodal data, explainable AI, and collaborative approaches as key drivers of innovation and advancement.

8.1 Incorporating Multimodal Data for Enhanced Fracture Detection:

The future of fracture detection is marked by the integration of diverse data modalities. This subsection delves into the fusion of multiple imaging techniques, such as combining X-ray, CT, and MRI data, to provide a comprehensive view of fractures. It discusses the potential synergies arising from these multimodal inputs, offering deeper insights into fracture characteristics, location, and severity. The section further explores challenges such as data preprocessing, alignment, and the development of models capable of comprehending and exploiting this rich information landscape.

8.2 Explainable AI for Enhanced Clinical Acceptance:

As deep learning models become more sophisticated, ensuring their interpretability and explain ability becomes imperative. This subsection delves into the emerging trend of explainable AI (XAI) within fracture detection. Techniques like saliency maps, attention mechanisms, and feature visualization are explored, enabling clinicians to comprehend and trust the decisions made by AI models. The significance of XAI in fostering clinical acceptance, reducing skepticism, and augmenting collaboration between AI and healthcare practitioners is highlighted [30].

8.3 Federated Learning and Collaborative Approaches:

The future envisions collaborative paradigms where deep learning models are trained across distributed institutions without centralizing patient data. Federated learning, a burgeoning approach, enables models to be trained collaboratively while preserving data privacy. This subsection explores the concept of federated learning, detailing its advantages in healthcare settings where data privacy and security are paramount. It discusses the challenges of aggregating model updates, federated optimization techniques, and the potential for creating robust and generalizable fracture detection models through collaboration.

9. Conclusion:

This concluding section encapsulates the key takeaways and overarching insights garnered throughout the comprehensive review of fracture detection using deep learning methods. The section revisits the pivotal significance of accurate fracture detection within medical imaging, juxtaposing it against the limitations of traditional methods. It underscores the transformation ushered in by deep learning, illuminating how architectures like CNNs have become pivotal instruments in deciphering intricate patterns within medical images. The exploration of diverse methodologies, performance metrics, and challenges highlights the multifaceted nature of fracture detection using deep learning. The intersection of clinical implementation, ethical considerations, and the fusion of AI and human expertise is also reflected upon. Throughout the journey, the review underscores the potential for deep learning to reshape clinical workflows, expedite diagnoses, and foster better patient outcomes. As the review culminates, it casts a gaze towards the future horizons of fracture detection using deep learning. The prospects ahead are promising, as evidenced by emerging trends such as the fusion of multimodal data, the imperative of explainable AI, and the paradigm-shifting potential of federated learning. These trends are poised to augment the accuracy, efficiency, and clinical acceptance of deep learning models in fracture detection. The integration of AI into clinical workflows offers the potential to reduce radiologist workload and enhance diagnosis speed while adhering to ethical considerations and safeguarding patient privacy.

References:

1. Meena, T., & Roy, S. (2022). Bone fracture detection using deep supervised learning from radiological images: A paradigm shift. Diagnostics, 12(10), 2420.
2. Kim, D. H., & MacKinnon, T. (2018). Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks. Clinical radiology, 73(5), 439-445.
3. Beyaz, S., Açıcı, K., & Sümer, E. (2020). Femoral neck fracture detection in X-ray images using deep learning and genetic algorithm approaches. Joint diseases and related surgery, 31(2), 175.
4. Ma, Y., & Luo, Y. (2021). Bone fracture detection through the two-stage system of crack-sensitive convolutional neural network. Informatics in Medicine Unlocked, 22, 100452.
5. Mutasa, S., Varada, S., Goel, A., Wong, T. T., & Rasiej, M. J. (2020). Advanced deep learning techniques applied to automated femoral neck fracture detection and classification. Journal of Digital Imaging, 33, 1209-1217.
6. Liu, X., Faes, L., Kale, A. U., Wagner, S. K., Fu, D. J., Bruynseels, A., ... & Denniston, A. K. (2019). A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. The lancet digital health, 1(6), e271-e297.
7. Basha, C. Z., Padmaja, T. M., & Balaji, G. N. (2019). An Effective and Reliable Computer Automated Technique for Bone Fracture Detection. EAI Endorsed Transactions on Pervasive Health and Technology, 5(18), e2-e2.
8. Joshi, D., & Singh, T. P. (2020). A survey of fracture detection techniques in bone X-ray images. Artificial Intelligence Review, 53(6), 4475-4517.
9. Santosh, K. C., Antani, S., Guru, D. S., & Dey, N. (Eds.). (2019). Medical imaging: artificial intelligence, image recognition, and machine learning techniques. CRC Press.
10. Morid, M. A., Borjali, A., & Del Fiol, G. (2021). A scoping review of transfer learning research on medical image analysis using ImageNet. Computers in biology and medicine, 128, 104115.
11. Rguibi, Z., Hajami, A., Zitouni, D., Elqaraoui, A., & Bedraoui, A. (2022). Cxai: Explaining convolutional neural networks for medical imaging diagnostic. Electronics, 11(11), 1775.
12. Storey, O., Wei, B., Zhang, L., & Fotso Mtope, F. R. (2020). Adaptive bone abnormality detection in medical imagery using deep neural networks. In Developments of Artificial Intelligence Technologies in Computation and Robotics: Proceedings of the 14th International FLINS Conference (FLINS 2020) (pp. 915-922).
13. Gunz, S., Erne, S., Rawdon, E. J., Ampanozi, G., Sieberth, T., Affolter, R., ... & Dobay, A. (2019). Automated rib fracture detection of postmortem computed tomography images using machine learning techniques. arXiv preprint arXiv:1908.05467.
14. Mutasa, S., Varada, S., Goel, A., Wong, T. T., & Rasiej, M. J. (2020). Advanced deep learning techniques applied to automated femoral neck fracture detection and classification. Journal of Digital Imaging, 33, 1209-1217.

15. Kitamura, G., Chung, C. Y., & Moore, B. E. (2019). Ankle fracture detection utilizing a convolutional neural network ensemble implemented with a small sample, de novo training, and multiview incorporation. Journal of digital imaging, 32, 672-677.
16. Joshi, D., & Singh, T. P. (2020). A survey of fracture detection techniques in bone X-ray images. Artificial Intelligence Review, 53(6), 4475-4517.
17. Jin, L., Yang, J., Kuang, K., Ni, B., Gao, Y., Sun, Y., ... & Li, M. (2020). Deep-learning-assisted detection and segmentation of rib fractures from CT scans: Development and validation of FracNet. EBioMedicine, 62.
18. Koska, O. I., Çilengir, A. H., Uluç, M. E., Yücel, A., & Tosun, Ö. (2023). All-star approach to a small medical imaging dataset: combined deep, transfer, and classical machine learning approaches for the determination of radial head fractures. Acta Radiologica, 64(4), 1476-1483.
19. Yilmaz, E. B., Buerger, C., Fricke, T., Sagar, M. M. R., Peña, J., Lorenz, C., ... & Meyer, C. (2021). Automated deep learning-based detection of osteoporotic fractures in CT images. In Machine Learning in Medical Imaging: 12th International Workshop, MLMI 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, September 27, 2021, Proceedings 12 (pp. 376-385). Springer International Publishing.
20. Alzubaidi, F., Makuluni, P., Clark, S. R., Lie, J. E., Mostaghimi, P., & Armstrong, R. T. (2022). Automatic fracture detection and characterization from unwrapped drill-core images using mask R–CNN. Journal of Petroleum Science and Engineering, 208, 109471.
21. Chilamkurthy, S., Ghosh, R., Tanamala, S., Biviji, M., Campeau, N. G., Venugopal, V. K., ... & Warier, P. (2018). Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. The Lancet, 392(10162), 2388-2396.
22. Warin, K., Limprasert, W., Suebnukarn, S., Inglam, S., Jantana, P., & Vicharueang, S. (2022). Assessment of deep convolutional neural network models for mandibular fracture detection in panoramic radiographs. International Journal of Oral and Maxillofacial Surgery, 51(11), 1488-1494.
23. Kalmet, P. H., Sanduleanu, S., Primakov, S., Wu, G., Jochems, A., Refaee, T., ... & Poeze, M. (2020). Deep learning in fracture detection: a narrative review. *Acta orthopaedica*, *91*(2), 215-220.
24. Chai, Z., Lin, H., Luo, L., Heng, P. A., & Chen, H. (2022, September). ORF-Net: Deep Omni-Supervised Rib Fracture Detection from Chest CT Scans. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 238-248). Cham: Springer Nature Switzerland.
25. Malik, G. K., Nainwal, A., Amrish, Bijalwan, V., & Semwal, V. B. (2022, November). Arm Fracture Detection Using Deep Convolution Neural Network. In International conference on Advances in Materials and Manufacturing (pp. 203-212). Singapore: Springer Nature Singapore.
26. Peña-Solórzano, C. A., Albrecht, D. W., Bassed, R. B., Burke, M. D., & Dimmock, M. R. (2020). Findings from machine learning in clinical medical imaging applications–Lessons for translation to the forensic setting. Forensic Science International, 316, 110538.
27. Twinprai, N., Boonrod, A., Boonrod, A., Chindaprasirt, J., Sirithanaphol, W., Chindaprasirt, P., & Twinprai, P. (2022). Artificial intelligence (AI) vs. human in hip fracture detection. Heliyon, 8(11).

28. Warin, K., Limprasert, W., Suebnukarn, S., Paipongna, T., Jantana, P., & Vicharueang, S. (2023). Maxillofacial fracture detection and classification in computed tomography images using convolutional neural network-based models. Scientific Reports, 13(1), 3434.
29. Yadav, D. P., Sharma, A., Athithan, S., Bhola, A., Sharma, B., & Dhaou, I. B. (2022). Hybrid SFNet model for bone fracture detection and classification using ML/DL. Sensors, 22(15), 5823.
30. Hardalaç, F., Uysal, F., Peker, O., Çiçeklidağ, M., Tolunay, T., Tokgöz, N., ... & Mert, F. (2022). Fracture detection in wrist X-ray images using deep learning-based object detection models. Sensors, 22(3), 1285.
31. Seol, Y. J., Kim, Y. J., Kim, Y. S., Cheon, Y. W., & Kim, K. G. (2022). A study on 3D deep learning-based automatic diagnosis of nasal fractures. Sensors, 22(2), 506.
32. Barhoom, A. M., Al-Hiealy, M. R. J., & Abu-Naser, S. S. (2022). Bone Abnormalities Detection and Classification Using Deep Learning-Vgg16 Algorithm. Journal of Theoretical and Applied Information Technology, 100(20), 6173-6184.
33. Tomita, N.; Cheung, Y.Y.; Hassanpour, S. Deep neural networks for automatic detection of osteoporotic vertebral fractures on CT scans. Comput. Biol. Med. 2018, 98, 8–15.
34. Sahin, M. E. (2023). Image processing and machine learning-based bone fracture detection and classification using X-ray images. International Journal of Imaging Systems and Technology, 33(3), 853-865.
35. Fritz, B., Paul, H. Y., Kijowski, R., & Fritz, J. (2023). Radiomics and deep learning for disease detection in musculoskeletal radiology: an overview of novel MRI-and CT-based approaches. Investigative radiology, 58(1), 3-13.
36. Ahmed, K. D., & Hawezi, R. (2023). Detection of bone fracture based on machine learning techniques. Measurement: Sensors, 27, 100723.
37. Paul, S. G., Saha, A., & Assaduzzaman, M. (2023). A real-time deep learning approach for classifying cervical spine fractures. Healthcare Analytics, 4, 100265.
38. Lin, Z. W., Dai, W. L., Lai, Q. Q., & Wu, H. (2023). Deep learning-based computed tomography applied to the diagnosis of rib fractures. Journal of Radiation Research and Applied Sciences, 16(2), 100558.
39. Kassem, M. A., Naguib, S. M., Hamza, H. M., Fouda, M. M., Saleh, M. K., & Hosny, K. M. (2023). Explainable Transfer Learning-Based Deep Learning Model for Pelvis Fracture Detection. International Journal of Intelligent Systems, 2023.
40. Krogue, J.D., Cheng, K.V., Hwang, K.M., Toogood, P., Meinberg, E.G., Geiger, E.J., Zaid, M., McGill, K.C., Patel, R., Sohn, J.H. and Wright, A., 2020. Automatic hip fracture identification and functional subclassification with deep learning. *Radiology: Artificial Intelligence*, *2*(2), p.e190023.
41. Raisuddin, A.M., Vaattovaara, E., Nevalainen, M., Nikki, M., Järvenpää, E., Makkonen, K., Pinola, P., Palsio, T., Niemensivu, A., Tervonen, O. and Tiulpin, A., 2021. Critical evaluation of deep neural networks for wrist fracture detection. *Scientific reports*, *11*(1), p.6006.

42. Lindsey, R., Daluiski, A., Chopra, S., Lachapelle, A., Mozer, M., Sicular, S., Hanel, D., Gardner, M., Gupta, A., Hotchkiss, R. and Potter, H., 2018. Deep neural network improves fracture detection by clinicians. *Proceedings of the National Academy of Sciences*, *115*(45), pp.11591-11596.
43. Kim, D.H. and MacKinnon, T., 2018. Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks. *Clinical radiology*, *73*(5), pp.439-445.
44. Chung, S.W., Han, S.S., Lee, J.W., Oh, K.S., Kim, N.R., Yoon, J.P., Kim, J.Y., Moon, S.H., Kwon, J., Lee, H.J. and Noh, Y.M., 2018. Automated detection and classification of the proximal humerus fracture by using deep learning algorithm. *Acta orthopaedica*, *89*(4), pp.468-473.
45. Tomita, N., Cheung, Y.Y. and Hassanpour, S., 2018. Deep neural networks for automatic detection of osteoporotic vertebral fractures on CT scans. *Computers in biology and medicine*, *98*, pp.8-15.
46. Urakawa, T., Tanaka, Y., Goto, S., Matsuzawa, H., Watanabe, K. and Endo, N., 2019. Detecting intertrochanteric hip fractures with orthopedist-level accuracy using a deep convolutional neural network. *Skeletal radiology*, *48*, pp.239-244.
47. Badgeley, M.A., Zech, J.R., Oakden-Rayner, L., Glicksberg, B.S., Liu, M., Gale, W., McConnell, M.V., Percha, B., Snyder, T.M. and Dudley, J.T., 2019. Deep learning predicts hip fracture using confounding patient and healthcare variables. *NPJ digital medicine*, *2*(1), p.31.
48. Cheng, C.T., Ho, T.Y., Lee, T.Y., Chang, C.C., Chou, C.C., Chen, C.C., Chung, I. and Liao, C.H., 2019. Application of a deep learning algorithm for detection and visualization of hip fractures on plain pelvic radiographs. *European radiology*, *29*(10), pp.5469-5477.
49. Al Arif, S.M.R., Knapp, K. and Slabaugh, G., 2018. Fully automatic cervical vertebrae segmentation framework for X-ray images. *Computer methods and programs in biomedicine*, *157*, pp.95-111.