Pakistan Journal of Engineering and Technology
ISSN (Print): 2664-2042; ISSN (Online): 2664-2050
Received: 2024-08-18
Accepted: 2024-09-21
https://doi.org/10.51846/vol7iss3pp120-128

PAKJET
Pakistan Journal of Engineering & Technology

Article

Texture Aware Deep Features for Precise Wrist Fracture Detection

Muhammad Hassan Butt *, Wakeel Ahmad, Syed M. Adnan Shah, and Usman Islam

Department of Computer Science, University of Engineering and Technology, Taxila 47080, Pakistan
* Correspondence: Muhammad Hassan Butt (muhammad.hassan2@students.uettaxila.edu.pk)

Abstract: Bone fractures are a critical medical condition requiring a thorough and accurate diagnosis to ensure proper treatment and healing. The physical examination of X-ray images by radiologists is time-consuming and subject to human error, indicating the need for an automated solution. This study proposes an advanced deep-learning approach to detect and classify bone fractures from X-ray images. This research aims to improve detection accuracy and streamline the diagnostic process using cutting-edge computer vision methods. Our approach starts with preprocessing the dataset, encompassing image resizing to 50x50 pixels, enhancement, and augmentation to improve diversity. Canny Edge detection is used to emphasize structural edges in the augmented dataset. The dataset is classified into training, testing, and validation subsets. A Convolutional Neural Network (CNN) automatically extracts deep learning features, obtaining complex patterns that show bone fractures. The Gray Level Co-occurrence Matrix (G.L.C.M) is utilized for texture features. These features are arranged to form a comprehensive feature set. The principal component analysis (PCA) is applied to this fused feature set to shrink dimensionality while preserving critical information, which is then used for fracture prediction. Experiments were conducted using a dataset of 14,718 X-ray images covering various bone fractures. The proposed method achieved an incredible classification accuracy of 98%, significantly outperforming traditional diagnostic methods and other contemporary models.

Keywords: CNN, GLCM, Fracture Detection, X-Ray Images, Bone Fracture, Medical Images

1. Introduction

1.1. Background and Motivation

Bone fractures are one of the most frequent types of injuries, which can manifest with a vast range of symptoms; the timely and precise diagnosis is essential. Manual review by Radiologist of the X-ray images is reliable but often became inaccurate because of a poor quality of images, human factors, and fractured patterns complexities [1, 2]. An oversight by the physician can result in delayed treatment process, the wrong treatment offered or can increase the patient's suffering. Computer Aided Diagnosis (CAD) offers a promising solution [3]. These methods can enhance radiologists' capacity and help them deliver a quicker and precise diagnosis to enhance the patient outcomes.

CAD systems have become more enhanced in the current world using Artificial Intelligence (AI) [4-6]. Artificial intelligence is a powerful approach for image analysis, in particular, deep learning. AI enabled CAD systems had proven their effectiveness better than medical experts in various healthcare diagnostics. The integration of machine learning in health care has resulted in today's solutions with a major focus on medical images. It is becoming essential to use computers in several areas and their use in healthcare makes an incredible opportunity in delivering patient care. Subsequently, deep learning has been considered one of the most promising techniques for the diagnosis of bone fracture from X-ray images, however, it has its limitations. The variations in fracture types, whether hairline, abraded, or mashed fractures complicate generalizability for models [7, 8]. In addition, there is a lack of high-quality annotated medical datasets that slow down the training process, and after that they hamper the performance of these models [9, 10]. Another issue is that deep learning model is usually a 'black box,' meaning that its output cannot be easily explained, which is important when making life-changing medical decisions [11, 12].

Though image analysis and using machine learning algorithms are in practice, each has their problem associated with it. Most of them rely on manually designed features, which cannot effectively represent some of the challenges pertaining to different types of fracture instances. This methodology also requires the input of skilled personnel hence it is not easily achievable. The conventional techniques as well also fail to address the variability of brightness and pixelation prevalent in X-rays and presence of multiple overlapping formations resulting in enhanced chances of errors which downplay the suitability of X-rays [13]. Both deep learning and other traditional approaches may therefore fail to achieve the required degree of accuracy, reliability, and especially interpretability when it comes to precise fracture diagnosis. The need to develop more comprehensive and effective solutions is therefore very evident. To overcome the above-mentioned limitations, we propose a novel approach that integrates concepts of both classical as well as deep learning algorithms. By integrating the handcrafted feature extraction of traditional methods with the pattern recognition capability of deep learning models this approach presents a more comprehensive diagnoses approach. This can be put on a well-

founded basis by emphasizing the significant fracture characteristics in classical machine learning, which subsequently deep learning can optimize and improve. Such combined approaches can enhance the model performance considerably, especially when data samples are scarce or noisy; classical methods help preprocess data samples and extract significant features that must otherwise be learned by deep learning models [14, 15]. Furthermore, when these two techniques are integrated, there is an improvement in the interpretability of the model, hence clinicians can be assured of the predictions made by the model. Classical factors provide a more direct link between deep learning and the end user as they explain the gap in a better way. Therefore, it makes more sense to integrate such solutions to obtain higher accuracy, reliability, and explainability of the bone fracture detection systems that can be effectively used in clinical practice.

The rest of the paper is organized as follows: Section 2 focuses on the literature review that provides the reader with the findings of state-of-the-art research in this domain. In Section 3 we describe the proposed methodology and discuss model training procedure. The results of the study are presented in section 4, along with a detailed discussion of the findings. Finally, section 5, presents the conclusion as well as future direction.

2. Literature Review

This study explores various advanced techniques for detecting fractures in long bones, ribs, and wrists. We have examined traditional techniques by highlighting their approaches, strengths, and reported outcomes. By identifying research gaps and evaluating commonly used metrics, this review aims to contribute to the development of improved fracture detection systems.

2.1. Long Bone

In this study [16], a Dilated Convolutional Feature Pyramid Network (D.C.F.P.N) was utilized to detect fractures in 3,842 x-rays of the thigh. The network achieved an impressive AP score of 82.1%. In another study [17], researchers used 3,842 x-rays of thigh fractures to develop the Parallel Net method for identifying fractures. The average accuracy measured by the AP50 score was 87.8%. In this research [18], fracture detection algorithms were utilized on 2,333 x-ray images of the femur. The results showed a Mean score of 68.8% for nine types of femoral fractures. The study used a faster RCNN, an anchor-based model with FPNs of different resolutions and a ResNet50 base network. In this article [19], the researchers developed a CNN model that utilizes an ensemble technique to identify ankle fractures. They used Xception, InceptionV3 and ResNet to extract features. The algorithm based on the ensemble achieved an accuracy of 81% in distinguishing between fractures and healthy individuals. In this study [20], the InceptionV3 model was employed to classify X-ray images of the proximal femur. The method can accurately differentiate between images of healthy femur bones and those with fractures. In another research [21], Dense Net was utilized to categorize femur fractures, resulting in an accuracy rate of 89%. In this article [22], D. P. Yadav and Sandeep Rathor developed a deep neural network model to differentiate between healthy bones and fractures. They utilized flipping and shifting images to generate a new photo from the available dataset. CNN uses a convolution layer that automatically extracts features from the source images to classify features as cancerous or healthy bones. Tensor Flow and deep learning techniques have been introduced to diagnose long, short, and plain bone fractures. To ensure accurate performance, the system needs validation on a larger dataset. To avoid overfitting, the training dataset was expanded. A 5-fold cross-validation was conducted on the dataset. The activation function ReLU is implemented in each layer. The dense layer has been applied with activation functions Adam-ax and SoftMax one at a time. In this research [23], fracture identification in calcaneus CT images is an emerging study area; however, several attempts have been made to use these imaging modalities to detect fractures in other bones. Consider the research of Anu et al., who used GL-CMs to identify bone fractures in x-ray pictures. In this paper [24], Tian T. suggested a method of detecting fractures in femur bones by measuring the neck-shaft angle of the femur. Mahendran et al. [25], proposed a fusion classification approach for automated Tibia bone fracture detection in this research. The authors began the procedure by performing preprocessing operations including binary conversion, edge identification, noise reduction and segmentation. Using a straightforward majority voting method, they employed three widely used classifiers to categorize the outcomes: FFBP Neural Network, SVM, and Naive Bayes. This research classifies long bones such as the thighbone, femur and tibia. Fractures are then categorized based on their kind and the site of incidence [26]. Table 1 presents a summary of existing techniques on long bones.

Table 1. Summary of existing SOTA techniques on long-bone.

Ref	Types of Images	Technique Used	No. of Images	Acc (%)
[16]	X-Ray	DCFPN	3842	82.10
[17]	X-Ray	Parallel Net	3842	87.80
[18]	X-Ray	Faster R-CNN	2333	71.50
[19]	X-Ray	Inceptionv3, Resnet	596	81.00
[20]	X-Ray	Inceptionv3	2453	86.00
[21]	X-Ray	Dense Net	1347	89.00
[22]	X-Ray	DNN	4000	92.44
[23]	CT-Scan	GLCM	40	85.00
[24]	X-Ray	Gabor	420	91.67
[25]	X-Ray	SVM / BPN / NB	1000	94.85
[26]	X-Ray	Geometric	100	78.00

2.2. Rib Bone

In this study [27], the FracNet model was created to identify rib fractures. It consists of encoders and decoders, 3D convolution, batch normalization, nonlinearity, and maximum pooling. The model was trained using 420 images from the RibFrac dataset and tested on a subset of 120 prints. This approach's segmentation rate was 71.5%, and its detection sensitivity was 92.9% for the validation cohort. This paper [28], presents a faster R-CNN method to detect and categorize rib-bone fractures into groups automatically. The technique

cracks and three objectives: strengthening the model, identifying and classifying cracks, and establishing a reliable mechanism. The study shows that R-CNN outperforms YoloV3 in terms of detection speed and accuracy, with an impressive 91.1% accuracy and 86.3% sensitivity. In this article [29], researchers conducted a second study with the same objectives using cross-modal data to train a CNN model for automatically detecting and categorizing rib fractures in medical clinics. A faster R-CNN model was used to compile CT scans and clinical data, and a result-merging approach was used to generate 3D lesion outcomes from 2D inferences. By adding clinical data to CT scans, the CNN model could more accurately classify fresh, healing, and older fractures than when CT pictures were used alone. The CNN model improved sensitivity (from 0.77 to 0.95, from 0.61 to 0.89, and from 0.55 to 0.80) while maintaining accuracy (from 0.87 to 0.91, from 0.77 to 0.84, and from 0.70 to 0.95).

This article [30], used preprocessing to reduce the size of a rib fracture image to 128*128*333 pixels. Afterward, a method known as semantic segmentation was employed to pinpoint the specific areas where the ribs were cracked. Ultimately, the CT scan images were analyzed using the UNet model, which achieved an impressive classification accuracy of 88.54%. This research paper [31], accurately classified rib fractures with a remarkable 90.2% accuracy rate using a CNN model. In this study [32], CT images and deep convolutional networks (ConvNet) were used to automatically detect fractures in the posterior elements of the spine. It should be noted that these methods only see bone fractures in medical images of a specific bone. In [33], Uysal et al. conducted 26 deep learning-based classification methods to identify the type of shoulder bone fracture depicted in musculoskeletal radiograph (MURA) dataset x-ray images. They also created two ensemble learning models to enhance the accuracy of the classification outcomes.

In this study [34], Beyaz et al. conducted a fracture classification study on 1,341 femoral neck X-ray images using a CNN model and genetic algorithm (GA). The study achieved an accuracy of 79.3%. In this article [35], Tobler and their team utilized the ResNet18 model to classify fractures in a 15,775 frontal and lateral radiograph dataset. Their results showed an impressive accuracy rate of 94%. In this paper [36], Chen et al. utilized the ResNet model which accomplished an accuracy rate of 73.59% in detecting vertebral fractures among a dataset of 1,306 plain frontal radiographs. Table 2 presents a summary of existing techniques on rib bones.

2.3. Wrist Bone

In this study [37], the Yolo 4 model was utilized to pinpoint the exact location of a bone fracture. Additionally, the researchers employed a data augmentation technique and evaluated the performance of their models using both the original and improved datasets. The method achieved an accuracy of 81.94%, effectively distinguishing between damaged and healthy sections within a single dataset.

Table 2. Summary of existing SOTA techniques on Rib bone.

Ref	Types of Images	Technique Used	No. of Images	Acc (%)
[27]	X-Ray	FracNet	7,473	71.50
[28]	CT Image	Faster R-CNN	25054	91.10
[29]	CT Image	CNN, Faster R-CNN	894	88.00 78.00
[30]	CT Image	U-Net CNN	720	88.54
[31]	CT Image	CNN	511	90.20
[32]	CT Image	Convnet	Random Subset	85.70
[33]	X-Ray	Densenet169	8942	88.03
[34]	X-Ray	Genetic Algorithm	1341	79.30
[35]	Radiograph	Resnet18	15775	94.00
[36]	Radiograph	Resnet Model	1306	73.59

This research paper [38], presents a deep CNN ensemble model for wrist fracture identification. The method successfully distinguishes between healthy and fractured wrist bones with an accuracy of 86.39%. In this article [39], the researchers utilized 11,112 images of wrist radiographs to train the model, and an additional 100 pictures were incorporated, comprising 50 fracture images and 50 typical images. The diagnostic sensitivity, selectivity, and AUC values were determined to be 88%, 90%, and 0.954%, respectively.

In this research [40], a convolutional neural model was developed using 135,409 radiographs to detect wrist fractures. The doctor's ability to interpret images improved from 88% to 94% with the aid of the model, resulting in a 53% reduction in misinterpretations. In this study [41], they examined radiographic images of the hand and wrist to assess a model for distal radius fractures. The model was compared to the diagnostic accuracy of experienced orthopedic physicians, and its efficiency was evaluated. The results showed that the model attained a sensitivity of 90% and a specificity of 88%. This research [42] aims to develop a system that can automatically detect fractures in hand bones using advanced filtering techniques to remove noise, edge detection methods to identify edges, and Wavelet and Curvelet transforms to extract significant features. Additionally, classification algorithms like decision trees will determine the fracture type. The system will work specifically with X-ray images. Table 3 presents a summary of existing techniques on wrist bones.

Table 3. Summary of existing SOTA techniques on wrist bone.

Ref	Types of Images	Technique Used	No. of Images	Acc (%)
[37]	X-Ray	Yolo4, Faster-RCNN	40,561	81.94
[38]	X-Ray	Deep CNN, WFD-C	542	86.39
[39]	Radiograph	Inceptionv3	11,112	95.40
[40]	Radiograph	CNN	135,409	94.00
[41]	X-Ray	VGG16	256,458	88.00
[42]	X-Ray	Wavelet / Curvelet	98	91.80

3. Proposed Methodology

This section outlines the methodologies proposed for automatic bone fraction detection from X-ray images. Fig. 1 depicts the proposed system's step-by-step procedure.

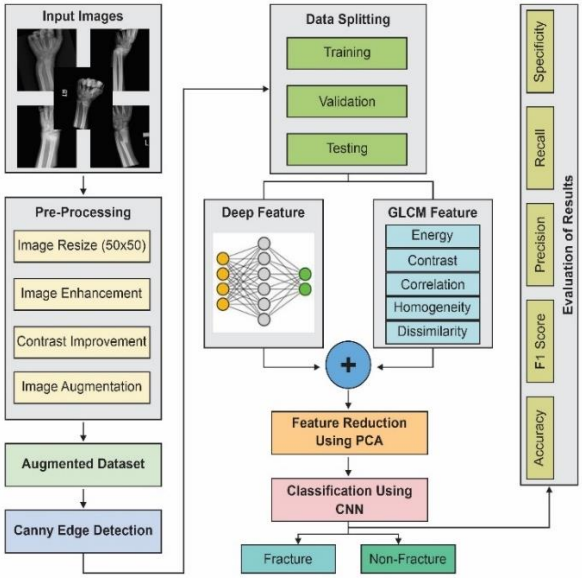

Figure 1. The workflow diagram of the proposed system.

3.1. Dataset

The dataset for automatic bone fracture detection using X-ray images collected from Kaggle [43], comprises 4,906 images categorized into two classes: fracture and non-fracture. To enhance the dataset diversity data augmentation techniques were applied which also improved the model performance. This dataset has been divided into three subsets for training, testing, and validation, with the training set containing 12,297 images, the testing set containing 1,203 images, and the validation set containing 1,218 images. Detailed statistics are provided in Table 4.

3.2. Pre-Processing

Due to the inherent noise, inconsistencies, and incompleteness of real-world medical images, preprocessing techniques are essential for enhancing data quality. By effectively removing artifacts and improving image clarity, preprocessing significantly improves the performance of subsequent analysis stages, such as feature extraction and classification. Initially, images were resized to 50x50 pixels to maintain uniformity and improve model training efficiency. For image enhancement, we used histogram equalization, making both fracture and non-fracture details more distinct. These preprocessing steps collectively enhance the accuracy of fracture detection and contribute to more reliable diagnostic outcomes. Sample images are shown in Fig. 2.

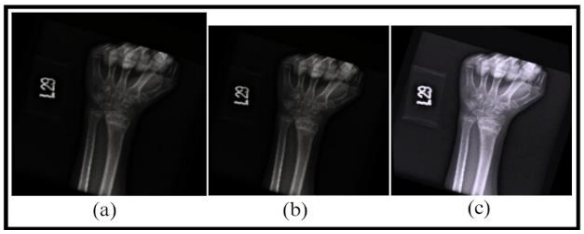

Figure 2. Sample images after preprocessing (a) original image (b) resize (c) enhanced image.

3.3. Image Augmentation

Starting with 4.906 images across two bone categories (fracture and non-fracture), we have applied various augmentation techniques it helps to enhance the model performance [44]. These techniques involved generating new, slightly modified versions of the existing images, expanding the dataset to a total of 14,718 images. Methods such as 20% shearing, 30% zooming, and horizontal flipping were used [45]. Detailed statistics are provided in Table 4, with sample images shown in Fig. 3. This augmentation process introduces a wider range of image variations, simulating real-world conditions where the appearance of fractures may vary. As a result, the dataset becomes more robust, and the model's ability to generalize to unseen data is significantly improved.

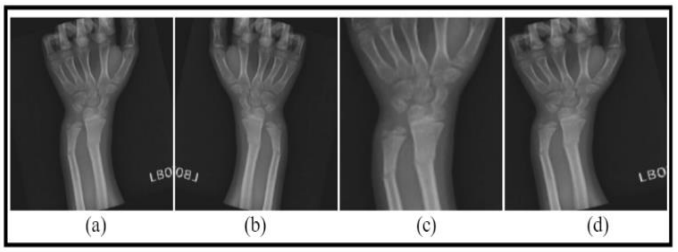

Figure 3. Sample images after augmentation (a) original image (b) horizontal flip (c) zoom 30% (d) shear 20%.

Table 4. Detail of dataset before and after augmentation.

Sr. No.	Classes	Original Dataset	Augmented Dataset
1.	Training	4099	12,297
2.	Validation	406	1218
3.	Testing	401	1203
Total		**4906**	**14718**

3.4. Edge Detection Using Canny Edge

Bone edges are detected from X-ray images using Canny Edge [46] which is widely recognized for its accuracy and robustness in identifying edges in images. The Canny Edge detection algorithm operates in several steps to ensure precise edge identification. At first step, to reduce the noise from X-ray images a gaussian filter applied. This is crucial for minimizing false edges that could result from noise artifacts. After that the brightness intensity of each pixel is determined and the Sobel operators are used to measure any changes in the gradient of intensity thus indicating edges. These gradient values represent the areas where intensity changes from one value to the other with large differences which is possible to indicate edges. The next step that the algorithm does is non-maximum suppression in which case it only preserves the local maxima in the gradient direction of the edges.

This thins the edges to about one pixel wide, which is critical if bones edges are to be highlighted and there is also a check on thick or blurred edges that may be occasioned by over lapping structures or noise. Thirdly, double thresholding is conducted in order to identify the strong edges and the weak edges. Essentially, strong edge maps which reveal the greatest variation of pixel's intensity are effectively and promptly detected as the edges of bone. Only candidates that are linked

to strong edges are preserved, thus guaranteeing the produced edges are contiguous and robust. The method used here results in the precise definition of edges of the bone which is crucial for the feature extraction part of the proposed model. Some of the sample images when the Canny edge algorithm has been applied are shown in Fig. 4.

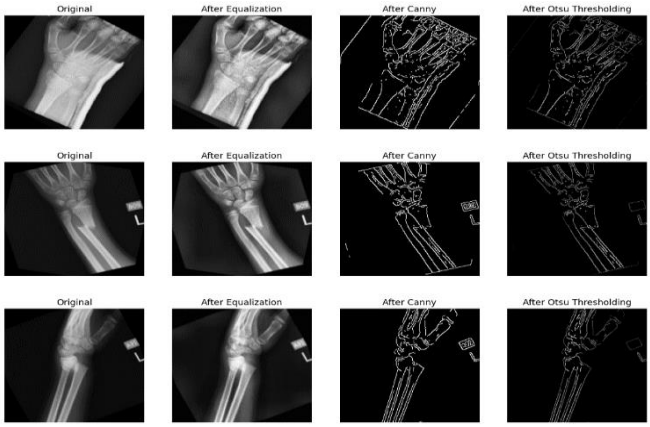

Figure 4. Sample images of canny edge detection results.

3.5. Feature Extraction

A hybrid feature extraction technique helps in extracting key features from the obtained bone edges. Somewhat more complex, patterns within the images are learned in the form of deep features that are inherent within a Convolutional Neural Network (CNN). However, handcrafted features are extracted using Gray-Level Co-Occurrence Matrices (GLCM), which describe the texture and structure of the bones. So, by adding these two sets of features, a better understanding of the image is achieved: improving the efficiency and specificity of the classification model.

3.5.1. Deep Feature Extraction

Convolutional Neural Network (CNN) architectural design starts with the input layer which takes X-ray images as the input then follow by three convolution layers which include Conv2D-1, Conv2D-2 and Conv2D-3. Every convolutional layer uses filters to extract the important aspects of the images including edges and textures. In order to speed up the training process and make the models more stable, to each convolutional layer, Group Normalization layer is added which helps to normalize the output. MaxPooling2D layers are then used to decrease the size of feature maps and therefore decrease the dimension of image. This reduction in dimensionality reduces the number of parameters making it less likely to overfit in the process.

3.5.2. Gray-Level Co-Occurrence Matrix (G.L.C.M)

GLCM [47] is a statistical technique applied for texture analysis to define discriminate features from X-ray images for bone fracture detection. GLCM gives a true picture of textural description of the image by giving a measure of the spatial dependency of gray levels. Some of the selected GLCM parameters that describe the textural attributes include energy, contrast, correlation, homogeneity, and dissimilarity. Energy expresses image uniformity where higher its value corresponds to domination of specific gray levels and can be calculated by Eq. (1). Contrast enhances intensity difference within the image and boosts the edges and intricate features and can be determined from Eq. (2). Correlation measures the local changes of pixel pairs, which is used to understand the relative homogeneity of the image and is obtained from Eq. (3). Homogeneity measures how close the gray-level co-occurrences are to the diagonal of the GLCM matrix that depicts image smoothness defined by the Eq. (4). Last of all, dissimilarity computes for mean intensity difference between two pixels and yields the degree of contrast of the images computed by the Eq. (5). Thus, the GLCM features directing the textural properties of the bone tissue plays a very important role in the accurate discrimination of normal and fractured bone regions.

$$\text{Energy} = \sqrt{\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} P(i,j)^2} \quad (1)$$

$$\text{Contrast} = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} P(i,j) * (i-j)^2 \quad (2)$$

$$\text{Correlation} = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} \frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i)(\sigma_j)}} \quad (3)$$

$$\text{Homogeneity} = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} \frac{P(i,j)}{1+|i-j|} \quad (4)$$

Here $P_{i,j}$ = Element (i, j) of the normalized symmetrical GLCM, and N = number of gray levels in the X-Ray image, μ = GLCM mean, calculated as $\mu = \sqrt{\sum_{i,j=0}^{N-1} iP_{ij}}$ and σ^2 =The variance of the intensities of all concerned pixels in the relationships that added to the GLCM, calculated as $\sigma^2 = \sum_{i,j=0}^{N-1} P_{ij}(i-\mu)^2$.

3.6. Feature Fusion

The feature fusion phase combines the extracted deep and handcrafted features into a single feature vector. This is achieved through early fusion, where features are concatenated before classification. Deep features, extracted from Convolutional Neural Networks (CNNs), capture complex patterns, while handcrafted features, such as those derived from Gray Level Co-occurrence Matrices (GLCM), offer domain-specific information. To harmonize these features, normalization is applied before concatenating them into a unified feature vector. This integrated representation is subsequently fed into PCA for feature reduction, which is then fed into a classifier. By combining deep and handcrafted features, we aim to leverage the strengths of both approaches and improve the overall performance of the fracture detection system.

3.7. Principal Component Analysis (PCA)

PCA [48, 49] is employed to reduce the dimensionality of the fused feature vector derived from deep and handcrafted feature extraction. By projecting the data onto a new subspace defined by the principal components, PCA effectively captures the most significant variations within the data while discarding less informative components. This process not only streamlines computational efficiency but also aids in mitigating the impact of noise and redundancy present in the original feature space. Through eigenvalue decomposition of the covariance matrix, PCA identifies orthogonal principal components that maximize variance, enabling the projection of data onto a lower-dimensional subspace while preserving essential information.

3.8. Classification using CNN

Following feature fusion and potential dimensionality reduction, a classification layer is employed to detect the presence or absence of a fracture in X-Ray image. This is achieved using fully connected (dense) layers within a neural network architecture. The extracted features are flattened into a one-dimensional vector suitable for input to dense layers. Using activation function (ReLU) these layers introduce non-linearity and enabling the model to learn complex relationships between features and the target variable. In order to minimize overfitting, the dropout layers are placed between the dense layers, where a certain amount of neurons is removed from the computation during the training process. The final dense layer employs sigmoid activation function to give a probability between 0 and 1 as the chance of a fracture. A decision threshold is then applied on this probability to come up with binary classification result.

4. Result And Discussion

This section gives a detailed discussion on the experiment of the proposed hybrid model. Subsequently, a detailed explanation of the findings endeavors to further elaborate on the consequences that can be derived from them.

4.1. Performance Evaluation Measure

Standard performance evaluation measures have been employed when measuring the performance of the proposed model. Accuracy is a basic assessment of the ability of models and is determined by the Eq. (6). Precision, it acknowledges how many times the positive class was correctly predicted by the model, as it is formulating in Eq. (7). Sensitivity or true positive rate (TPR) which is equal to recall can be determined using the Eq. (8). Finally, the F1- score, which measures the weighted average between precision and recall is computed using Eq. (9). Specificity calculates the ability of the proposed model to categorically exclude true negatives of each type of image using Eq. (10).

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{6}$$

$$\text{Precision Rate} = \frac{\text{TP}}{\text{TP+FP}} \tag{7}$$

$$\text{Recall Rate} = \frac{\text{Tp}}{\text{TP} + \text{FN}} \tag{8}$$

$$\text{F1} - \text{Score} = 2 * \frac{(\text{Precision} * \text{Recall})}{\text{Precision} + \text{Recall}} \tag{9}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN+FP}} \tag{10}$$

The CNN model was trained for 10 epochs, carefully balancing the risks of under-fitting and overfitting and allowing the loss function to converge properly. The trained model achieved a 98% accuracy in identifying fracture and non-fracture bone, as illustrated in Fig. 5, with the model's loss depicted in Fig. 6. The model also demonstrated a precision of 98% with an F1 score of 98%, reflecting its high accuracy in positive classification.

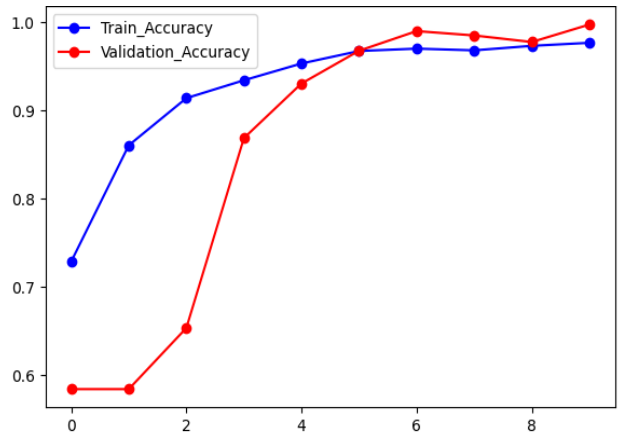

Figure 5. Training and validation accuracy.

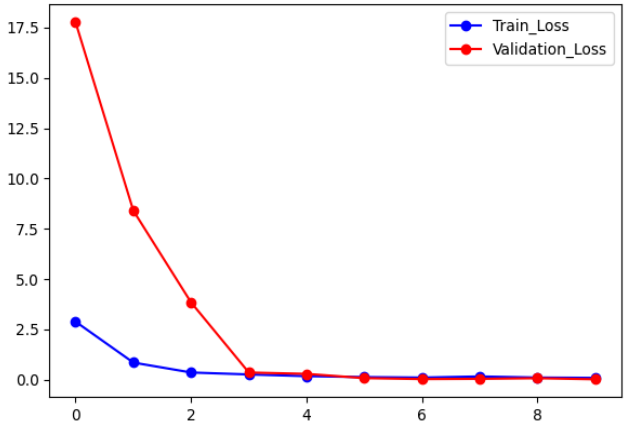

Figure 6. Training and validation loss.

The model achieved a recall of 98%, effectively identifying true positive cases, while its specificity of 100% demonstrated its ability to accurately detect fracture and non-fracture bone or confirm their absence. The results are presented in Fig. 7.

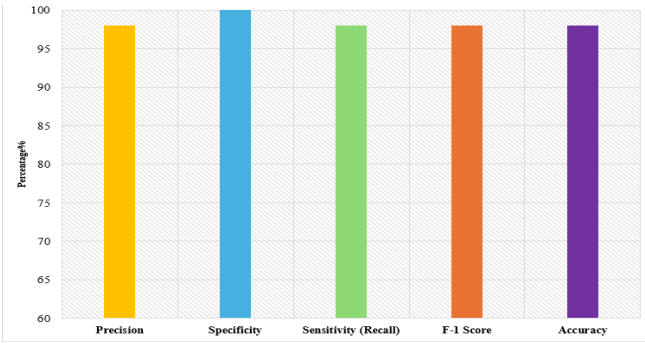

Figure 7. Results of the classification report.

Fig. 8 displays the results of our proposed model compared to state-of-the-art studies that utilized different learning models. In our comparison, we specifically focused on studies involving datasets related to bone fractures. Our analysis showed that our model addressed a broader range of bone fracture classes, and our fine-tuned, pre-trained model achieved an impressive accuracy of 98%. Same comparison is presented in tabular format using Table 5.

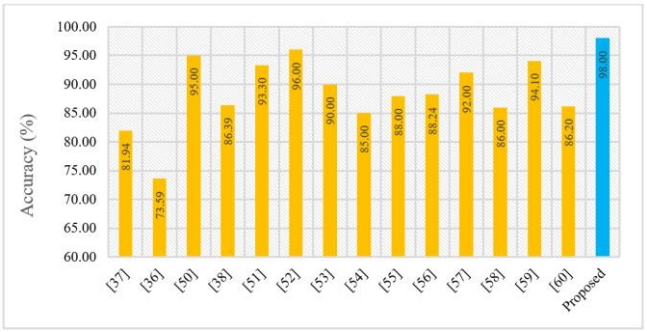

Figure 8. Accuracy comparison with existing SOTA techniques.

Table 5. Accuracy Comparison with existing SOTA Techniques.

Ref	Year	Technique	Acc %
[36]	2021	DCNN	73.59
[37]	2021	YOLO 4, FASTER-RCNN	81.94
[38]	2022	DEEP CNN, WFD-C	86.39
[50]	2021	DCNN	95.00
[51]	2022	CNN WITH TL	93.30
[52]	2022	U-NET NN	96.00
[53]	2022	CNN WITH GRAD-CAM	90.00
[54]	2023	LSNET	85.00
[55]	2023	FASTER R-CNN	88.00
[56]	2023	DCNN-2, DCNN-LSTM 2	86.54 88.24
[57]	2023	AI ALGORITHM	92.00
[58]	2024	YOLOV4	86.00
[59]	2024	RESNET-18	94.10
[60]	2024	RESNET152, VGG19	86.20 73.20
PRO	--	**CNN+GLCM**	**98.00**

The proposed CNN+GLCM model shows a significant improvement in accuracy compared to existing state-of-the-art techniques. This is because, in medical images like X-rays, texture variations are crucial for detecting fractures. GLCM effectively captures these texture details, while CNN focuses on recognizing patterns and structures. By combining both methods, the model becomes more robust and precise, overcoming the limitations of CNN alone and leading to better fracture detection.

5. Conclusion and Future Work

Accurate and timely bone fracture detection is essential for optimal patient care. Usually, traditional procedures that rely on human expertise are slow and error prone. To overcome these challenges, this study recommends a combined approach that uses both handcrafted and deep features for better fracture identification. A dataset comprising of 14,718 X-ray images was selected deliberately to mimic real-life conditions thereby enhancing the model's performance under different settings. Preprocessing steps were essential in improving image quality by performing noise reduction and normalization. Bone structures were outlined by performing edge detection using Canny algorithm. Similarly, GLCM analysis was used to obtain associated bone microstructure through texture features extraction. In order to exploit deep learning capabilities, this enhanced dataset was used for fine-tuning a pre-trained CNN which allows the model to learn complex feature representations and adapt itself for fracture detection task as such. A hyperparameter search has been performed to improve the accuracy of our models. A significantly high classification rate of 98% is reported with the proposed method for fracture prediction. The efficiency of the model was thoroughly evaluated using precision, recall, specificity, F1-score and support measures. Future studies will consider performing more experiments on a variety of fractures and their complexities will create deeper and stronger fracture detection systems.

Declarations

Funding: This research did not receive any external funding.

Conflicts of Interest: This version is straightforward and clearly communicates the intended message.

Data Availability Statement: The data used in this study is free available on internet, the source can be found in the dataset section.

Acknowledge: The authors acknowledge the dataset owners for providing the datasets on the Kaggle platform.

References

[1] A. P. Brady, "Error and discrepancy in radiology: inevitable or avoidable?," *Insights into Imaging,* vol. 8, no. 1, pp. 171-182, 2017/02/01 2017.

[2] J. N. Itri, R. R. Tappouni, R. O. McEachern, A. J. Pesch, and S. H. Patel, "Fundamentals of Diagnostic Error in Imaging," vol. 38, no. 6, pp. 1845-1865, 2018.

[3] H. P. Chan, L. M. Hadjiiski, and R. K. J. M. p. Samala, "Computer-aided diagnosis in the era of deep learning," vol. 47, no. 5, pp. e218-e227, 2020.

[4] R. Masud, M. Al-Rei, and C. Lokker, "Computer-Aided Detection for Breast Cancer Screening in Clinical Settings: Scoping Review," (in English), Review %J JMIR Med Inform vol. 7, no. 3, p. e12660, 2019.

[5] W. Shafik, A. F. Hidayatullah, K. Kalinaki, and M. M. Aslam, "Artificial Intelligence (AI)-Assisted Computer Vision (CV) in Healthcare Systems," in *Computer Vision and AI-Integrated IoT Technologies in the Medical Ecosystem*: CRC Press, 2024, pp. 17-36.

[6] J. Gong, J.-y. Liu, L.-j. Wang, B. Zheng, and S.-d. Nie, "Computer-aided detection of pulmonary nodules using dynamic self-adaptive template matching and a FLDA classifier," *Physica Medica,* vol. 32, no. 12, pp. 1502-1509, 2016/12/01/ 2016.

[7] L. Alzubaidi *et al.*, "Comprehensive review of deep learning in orthopaedics: Applications, challenges, trustworthiness, and fusion," *Artificial Intelligence in Medicine,* vol. 155, p. 102935, 2024/09/01/ 2024.

[8] A. Bhatnagar, A. L. Kekatpure, V. R. Velagala, and A. Kekatpure, "A Review on the Use of Artificial Intelligence in Fracture Detection," (in eng), *Cureus,* vol. 16, no. 4, p. e58364, Apr 2024.

[9] S. Wang *et al.*, "Annotation-efficient deep learning for automatic medical image segmentation," *Nature Communications,* vol. 12, no. 1, p. 5915, 2021/10/08 2021.

[10] W. Ahmad, S. Shah, A. J. K. T. o. I. Irtaza, and I. Systems, "Plants disease phenotyping using quinary patterns as texture descriptor," vol. 14, no. 8, pp. 3312-3327, 2020.

[11] V. Hassija *et al.*, "Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence," *Cognitive Computation,* vol. 16, no. 1, pp. 45-74, 2024/01/01 2024.

[12] A. Rahman *et al.*, "Machine learning and deep learning-based approach in smart healthcare: Recent advances, applications, challenges and opportunities," (in eng), *AIMS Public Health,* vol. 11, no. 1, pp. 58-109, 2024.

[13] X. Liu and Z. J. a. p. a. Wang, "Deep learning in medical image classification from mri-based brain tumor images," 2024.

[14] M. Suganthi and J. G. R. Sathiaseelan, "An Exploratory of Hybrid Techniques on Deep Learning for Image Classification," in *2020 4th International Conference on Computer, Communication and Signal Processing (ICCCSP)*, 2020, pp. 1-4.

[15] M. Luo and K. Zhang, "A hybrid approach combining extreme learning machine and sparse representation for image classification," *Engineering Applications of Artificial Intelligence,* vol. 27, pp. 228-235, 2014/01/01/ 2014.

[16] B. Guan, J. Yao, G. Zhang, and X. J. P. R. L. Wang, "Thigh fracture detection using deep learning method based on new dilated convolutional feature pyramid network," vol. 125, pp. 521-526, 2019.

[17] M. Wang, J. Yao, G. Zhang, B. Guan, X. Wang, and Y. J. M. S. Zhang, "ParallelNet: Multiple backbone network for detection tasks on thigh bone fracture," pp. 1-10, 2021.

[18] Y. Qi *et al.*, "Ground truth annotated femoral X-ray image dataset and object detection based method for fracture types classification," vol. 8, pp. 189436-189444, 2020.

[19] G. Kitamura, C. Y. Chung, and B. E. J. J. o. d. i. Moore, "Ankle fracture detection utilizing a convolutional neural network ensemble implemented with a small sample, de novo training, and multiview incorporation," vol. 32, pp. 672-677, 2019.

[20] L. Tanzi, E. Vezzetti, R. Moreno, A. Aprato, A. Audisio, and A. J. E. j. o. r. Massè, "Hierarchical fracture classification of proximal femur X-Ray images using a multistage Deep Learning approach," vol. 133, p. 109373, 2020.

[21] M. Lotfy, R. M. Shubair, N. Navab, and S. Albarqouni, "Investigation of focal loss in deep learning models for femur fractures classification," in *2019 International Conference on Electrical and Computing Technologies and Applications (ICECTA)*, 2019, pp. 1-4: IEEE.

[22] D. Yadav and S. Rathor, "Bone fracture detection and classification using deep learning approach," in *2020 International Conference on Power Electronics & IoT Applications in Renewable Energy and its Control (PARC)*, 2020, pp. 282-285: IEEE.

[23] T. Anu and R. J. I. J. C. A. Raman, "Detection of bone fracture using image processing methods," vol. 975, p. 8887, 2015.

[24] J. C. He, W. K. Leow, and T. S. Howe, "Hierarchical classifiers for detection of fractures in X-ray images," in *Computer Analysis of Images and Patterns: 12th International Conference, CAIP 2007, Vienna, Austria, August 27-29, 2007. Proceedings 12*, 2007, pp. 962-969: Springer.

[25] S. Mahendran, S. S. J. G. J. o. C. S. Baboo, and Technology, "An enhanced tibia fracture detection tool using image processing and classification fusion techniques in X-ray images," vol. 11, no. 14, pp. 22-28, 2011.

[26] O. Bandyopadhyay, A. Biswas, B. J. P. R. Bhattacharya, and I. Analysis, "Classification of long-bone fractures based on digital-geometric analysis of X-ray images," vol. 26, pp. 742-757, 2016.

[27] L. Jin *et al.*, "Deep-learning-assisted detection and segmentation of rib fractures from CT scans: Development and validation of FracNet," vol. 62, 2020.

[28] Q.-Q. Zhou *et al.*, "Automatic detection and classification of rib fractures on thoracic CT using convolutional neural network: accuracy and feasibility," vol. 21, no. 7, p. 869, 2020.

[29] Q.-Q. Zhou *et al.*, "Automatic detection and classification of rib fractures based on patients' CT images and clinical information via convolutional neural network," vol. 31, pp. 3815-3825, 2021.

[30] Z. N. Haitaamar and N. Abdulaziz, "Detection and semantic segmentation of rib fractures using a convolutional neural network approach," in *2021 IEEE Region 10 Symposium (TENSYMP)*, 2021, pp. 1-4: IEEE.

[31] T. Weikert *et al.*, "Assessment of a deep learning algorithm for the detection of rib fractures on whole-body trauma computed tomography," vol. 21, no. 7, p. 891, 2020.

[32] H. R. Roth, Y. Wang, J. Yao, L. Lu, J. E. Burns, and R. M. Summers, "Deep convolutional networks for automated detection of posterior-element fractures on spine CT," in *Medical imaging 2016: computer-aided diagnosis*, 2016, vol. 9785, pp. 165-171: SPIE.

[33] F. Uysal, F. Hardalaç, O. Peker, T. Tolunay, and N. J. A. S. Tokgöz, "Classification of shoulder x-ray images with deep learning ensemble models," vol. 11, no. 6, p. 2723, 2021.

[34] S. Beyaz, K. Açıcı, E. J. J. d. Sümer, and r. surgery, "Femoral neck fracture detection in X-ray images using deep learning and genetic algorithm approaches," vol. 31, no. 2, p. 175, 2020.

[35] P. Tobler *et al.*, "AI-based detection and classification of distal radius fractures using low-effort data labeling: evaluation of applicability and effect of training set size," vol. 31, pp. 6816-6824, 2021.

[36] H.-Y. Chen *et al.*, "Application of deep learning algorithm to detect and visualize vertebral fractures on plain frontal radiographs," vol. 16, no. 1, p. e0245992, 2021.

[37] H. P. Nguyen, T. P. Hoang, and H. H. Nguyen, "A deep learning based fracture detection in arm bone X-ray images," in *2021 international conference on multimedia analysis and pattern recognition (MAPR)*, 2021, pp. 1-6: IEEE.

[38] F. Hardalaç *et al.*, "Fracture detection in wrist X-ray images using deep learning-based object detection models," vol. 22, no. 3, p. 1285, 2022.

[39] D. Kim and T. J. C. r. MacKinnon, "Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks," vol. 73, no. 5, pp. 439-445, 2018.

[40] R. Lindsey *et al.*, "Deep neural network improves fracture detection by clinicians," vol. 115, no. 45, pp. 11591-11596, 2018.

[41] J. Olczak *et al.*, "Artificial intelligence for analyzing orthopedic trauma radiographs: deep learning algorithms—are they on par with humans for diagnosing fractures?," vol. 88, no. 6, pp. 581-586, 2017.

[42] M. Al-Ayyoub, I. Hmeidi, and H. J. J. M. P. T. Rababah, "Detecting Hand Bone Fractures in X-Ray Images," vol. 4, no. 3, pp. 155-168, 2013.

[43] (August 8, 2024). *Bone Fracture*. Available: kaggle.com/datasets/ahmedashrafahmed/bone-fracture

[44] K. Ghosh, C. Bellinger, R. Corizzo, P. Branco, B. Krawczyk, and N. Japkowicz, "The class imbalance problem in deep learning," *Machine Learning,* vol. 113, no. 7, pp. 4845-4901, 2024/07/01 2024.

[45] M. Xu, S. Yoon, A. Fuentes, and D. S. Park, "A Comprehensive Survey of Image Augmentation Techniques for Deep Learning," *Pattern Recognition,* vol. 137, p. 109347, 2023/05/01/ 2023.

[46] J. Canny, "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. PAMI-8, no. 6, pp. 679-698, 1986.

[47] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification," *IEEE Transactions on Systems, Man, and Cybernetics,* vol. SMC-3, no. 6, pp. 610-621, 1973.
[48] A. Maćkiewicz and W. Ratajczak, "Principal components analysis (PCA)," *Computers & Geosciences,* vol. 19, no. 3, pp. 303-342, 1993/03/01/ 1993.
[49] B. M. Salih Hasan and A. M. Abdulazeez, "A Review of Principal Component Analysis Algorithm for Dimensionality Reduction," *Journal of Soft Computing and Data Mining,* vol. 2, no. 1, pp. 20-30, 04/15 2021.
[50] A. P. Yoon, Y. L. Lee, R. L. Kane, C. F. Kuo, C. Lin, and K. C. Chung, "Development and Validation of a Deep Learning Model Using Convolutional Neural Networks to Identify Scaphoid Fractures in Radiographs," (in eng), *JAMA Netw Open,* vol. 4, no. 5, p. e216096, May 3 2021.
[51] K. Üreten, H. F. Sevinç, U. İğdeli, A. Onay, and Y. Maraş, "Use of deep learning methods for hand fracture detection from plain hand radiographs," (in eng), *Ulus Travma Acil Cerrahi Derg,* vol. 28, no. 2, pp. 196-201, Jan 2022. Düz el radyografilerinden el kırıklarının tespiti için derin öğrenme yöntemlerinin kullanılması.
[52] F. Hržić, S. Tschauner, E. Sorantin, and I. Štajduhar, "Fracture Recognition in Paediatric Wrist Radiographs: An Object Detection Approach," vol. 10, no. 16, p. 2939, 2022.
[53] A. Singh *et al.*, "Automated detection of scaphoid fractures using deep neural networks in radiographs," *Engineering Applications of Artificial Intelligence,* vol. 122, p. 106165, 06/01 2023.
[54] T. Anwar and H. J. I. J. o. I. T. Anwar, "Lsnet: a novel cnn architecture to identify wrist fracture from a small x-ray dataset," vol. 15, no. 5, pp. 2469-2477, 2023.
[55] J. R. Zech *et al.*, "Detecting pediatric wrist fractures using deep-learning-based object detection," vol. 53, no. 6, pp. 1125-1134, 2023.
[56] T. Rashid, M. S. Zia, Najam-ur-Rehman, T. Meraj, H. T. Rauf, and S. Kadry, "A Minority Class Balanced Approach Using the DCNN-LSTM Method to Detect Human Wrist Fracture," vol. 13, no. 1, p. 133, 2023.
[57] N. Hendrix *et al.*, "Musculoskeletal radiologist-level performance by using deep learning for detection of scaphoid fractures on conventional multi-view radiographs of hand and wrist," vol. 33, no. 3, pp. 1575-1588, 2023.
[58] L.-W. Cheng *et al.*, "Automated detection of vertebral fractures from X-ray images: A novel machine learning model and survey of the field," vol. 566, p. 126946, 2024.
[59] S. R. Thorat, D. G. Jha, A. K. Sharma, D. V. J. J. o. M. S. Katkar, and Research, "Wrist fracture detection using self-supervised learning methodology," vol. 8, no. 2, pp. 133-141, 2024.
[60] J. Li *et al.*, "Detection of hidden pediatric elbow fractures in X-ray images based on deep learning," vol. 17, no. 2, p. 100893, 2024.