Detection of Bone Fracture Using X-Ray Images

DMML2: Project Report

Overall grade / 77.9% for CA

Sanjay Girish Dialani
National College of Ireland
Dublin, Ireland
X22102442@student.ncirl.ie

Akimuddin Aslam Shaikh
National College of Ireland
Dublin, Ireland
X22123245@student.ncirl.ie

need some refining in the abstract

***Abstract*— The study aims to address the task of diagnosing bone fractures through the analysis of X-ray images using machine learning techniques. The study focuses on employing Convolutional Neural Networks (CNNs), Random Forest combined with the Canny Edge Detection technique, and Transfer Learning techniques to enhance the accuracy of fracture detection. The study followed the CRISP-DM methodology which began by thoroughly exploring the dataset, evaluating class distributions, image dimensions, and class balancing.**

For the first model which is the CNN model built on the top of pre-trained EfficientNetB3, the study constructed a deep learning architecture with data augmentation which helps the model to learn intricate patterns and features from generated images. The results showcased promising performance, with an accuracy of 85.37% on the training data set. However, showed 63.33% of accuracy on the validation dataset.

Additionally, the Canny Edge Detection technique Combined with the Random Forest classifier showed an accuracy of 88.83%. Moreover, Transfer Learning using a pre-trained EfficientNetB3 model which used the Edge Detection yielded exceptional results, achieving an accuracy of 94.59%, recall of 97.15%, precision of 92.30%, and an AUC-ROC of 98.84%.

This project demonstrates the ability of deep learning models to accurately identify fractures in X-ray images. The results showed the importance of the selection of preprocessing steps and a suitable model. Furthermore, the usage of advanced technology improves the accuracy of medical diagnosis.

Future work includes model architectures refining and extending the analysis to multi-class scenarios. The methodology presented in this study serves as a valuable base for similar medical image analysis activities, which helps healthcare professionals in making more informed decisions through advanced techniques.

Keywords*—*Bone fracture detection, X-ray images, upper extremity joints, Convolutional Neural Network (CNN), Canny edge detection, image classification, Random Forest, deep learning

I. Introduction

references needed

In the diverse and active field of medical imaging, deep learning technology brought crucial changes, which changed the way of diagnosis and healthcare. The improvements in Convolutional Neural Networks (CNNs) made automating tasks such as image recognition and pattern extraction very easy. These improvements in the Machin learning technology led to the exploring the realm of automatic bone fracture detection using X-ray images. This effort's goal is to speed up the process of diagnosis.

The traditional way of diagnosing fractures with the help of X-ray images depends completely on the ability of the analysis of radiologist. However, manual assessments have their own limitations. Manual assessments are often time-consuming, highly influenced by the subjectivities of the radiologist, and hence prone to human errors and situations. The rapid rise in imaging data in the medical field leads to the need for well-designed diagnostic approaches. In this dynamic field, it is very helpful to incorporate machine learning and deep learning techniques.

The centre of the study's objective is the incredible ability of deep learning to transform the medical diagnosis field. This aims to use capabilities to create, construct, and assess models that can automate the task of bone fracture identification in X-ray images. This activity not only speeds up the process of the diagnostic procedure but also consistent and dependable approach that provides the evaluation of various scenarios precisely and consistently. Ultimately, effort looking forward to improving patient outcomes and the quality of clinical decision-making.

The core of the study is a fundamental research question that guides the exploration: Can deep learning models effectively differentiate in X-ray images to identify if the bone is fractured or not with reaching higher accuracy and operational efficiency? According to this research question, the study's prime objective is to design, implement, and rigorously evaluate a series of image classification models that can precisely differentiate between fractured and non-fractured X-ray images of bones.

In order to answer this research question, Study developed three unique models, each one using a different approach and methodologies. The first one utilized the CNN which is developed on the top of the base model of EfficientNetB3. For the latter two models, the images are prepared by applying the Canny Edge Detection technique. The second model applied after edge detection is Random Forest Model and the third one is CNN on edge-detected images.

The upcoming sections of the study report give thorough exploration of the complete step-by-step stages followed in this study, Starting with the literature review which discusses the previous research work conducted in the field, This section examines the gaps in these studies and compares them with the proposed solution. Followed by the methodologies section, This section details the stages followed while carrying out the project understanding business, understanding the dataset with exploratory data analysis, preprocessing and preparation of the data for a later stage, and selection and development of three different models. The evaluation of these models and finally the study provide the conclusion on the findings and future work which could improve the outcomes,

II. Literature Review

good

In the area of bone fracture detection classification, a number of different research studies have been conducted, Each of them made a different contribution with distinct ways of tackling the difficulties associated with the manual diagnostic approach.

The first article study researched is by Yadav and Rathor (2020) [1]. Their work emphasises the need for automated systems to address the time-consuming and error-prone aspects of manual fracture identification. The study's use of Deep Neural Networks (DNNs) illustrates the potential of deep learning in the area, achieving a notable accuracy of 92.44% in classification tasks. This is achieved through a 5-fold cross-validation process, which highlights the effectiveness of DNNs. However, this study addresses the concern of overfitting on limited datasets and introduces techniques for data augmentation to alleviate this obstacle.

Another study by Chai et al. (2011) [2] explores the domain of Computer-Aided Detection (CAD) to overcome the issues of subjectivity and time consumption associated with manual bone fracture analysis. Their approach, using Gray-Level Co-occurrence Matrix (GLCM) recognition, provides good results with an accuracy of 86.67%. The study shows the importance of the potential of image processing techniques to automate diagnosis. However, this study proposes that refining the feature extraction process could further enhance the system's accuracy, Which the current study at hand handles well.

The study by Yadav et al. (2022) [3] contributed a novel approach by clubbing machine learning and deep learning in their Hybrid SFNet model. This innovative approach integrates a convolution neural network (CNN) with a canny edge algorithm, improving fracture localization and computational efficiency. This achieves an impressive accuracy of 99.12%, F1-score of 99%, and recall of 100%, the study illustrates that advanced preprocessing methods can significantly impact model performance. By benchmarking their approach against state-of-the-art deep CNN models, the study highlights the comparative advantages of their proposed Hybrid SFNet model. The current study at hand makes use of some of the ideas from this particular study and improvises on top of this by adding additional layers into CNN.

The study [4] uses edge detection using Canny and Sobel techniques and feature detection done by using Hough line detection and Harris corner detection. These preprocessing stages helped in achieving the accuracy of 88.67% and 0.89AUC. [5] uses deep learning in detecting bone fractures in children. They achieved an accuracy of 95.7%. However, the detection was more promising in the bones of children who are older than 4 years.

Research work by Jones et al. (2020) [6] conducted a multi-site study demonstrating the efficacy of a deep-learning system in accurately identifying fractures across the adult musculoskeletal system. Their approach shows the potential to mitigate future diagnostic errors and enhance radiograph interpretation accuracy, especially for cases requiring prompt and accurate diagnosis. The results obtained are promising with a sensitivity of 95.2% and specificity of 81.3%. But, the drawback of the study is that Data is not equally diverse among different parts of the human body. Their data mostly contains more common regions like a foot.

A critical study by Raisuddin et al. (2021) [7] highlighted the importance of analysis of deep learning-based models before clinical deployment. This study mainly concentrates on wrist fracture detection and highlighted the differences between general test sets and challenging test sets asserted by CT imaging. While deep learning approaches achieved near-perfect performance on general test sets showing 0.99 precision but on challenging scenarios, it dropped to 0.64. This shows the need to thoroughly assess models' performance under different conditions to ensure reliability.

Zhang et al. (2021)[8] studied rib fracture detection using CT scans, exploring how deep learning software can enhance radiologists' detection accuracy and efficiency of reading. By utilizing deep learning as a concurrent reader, the study showed improvements in false positives per scan, sensitivity, and reading efficiency. However, This work is mainly concentrated on rib fractures.

Sharma et al. (2021) [9] proposed a novel approach employing machine learning and digital geometry for bone fracture detection. Their proposed method will identify the bone fractures by bone contours by discarding discontinuity using segmentation.

The study by Mutasa et al. (2020) [10] illustrated advanced deep-learning techniques in the area of femoral neck fracture detection and classification. By clubbing the data augmentation methods like generative adversarial networks and digitally reconstructed radiographs, the model achieved accurate classification results. Their study highlights the potential of deep learning to point to complex fracture classification tasks.

Guan et al.'s (2019) [11] research work stands out as they introduce the Dilated Convolutional Feature Pyramid Network (DCFPN) for thigh fracture detection. Their approach involves architectural innovation through the incorporation of dilated convolutions, which significantly enhances feature extraction, resulting in superior accuracy compared to existing methods. This showed the Average Precision of 78.2%. However, this study mainly concentrates on thigh fracture detection.

Krogue et al. (2020) [12] studied automated hip fracture identification, highlighting the collaborative objective of deep learning models and human expertise. This study used a Densely Connected Convolutional Neural Network (or DenseNet). This approach achieved good accuracy of 93.7% ad a sensitivity of 93.2% for binary classification. Multiclass classification accuracy of 90.8%. However, the greatest limitation of the model was the relatively low sensitivity to nondisplaced FN fractures which is 61% and only 51% correctly subclassified.

Joshi and Singh (2020) [13] provide a deep understanding of fracture detection techniques. The proposed project aims to contribute to aim by refining a specific deep-learning architecture for improved fracture detection in a particular type of bone. This focused approach aligns with the goal of maximizing deep learning's potential in fracture diagnosis.

Guan et al. (2020) [14] propose an improved deep convolutional neural network for the detection of arm fractures. Their study made use of 2-stage R-CNN. This study worked on a dataset of around 4000 arm X-ray images. This showed good results of accuracy. But, This model works mainly on arm fractures.

Another work [15] introduces the detection of wrist fracture from posteroanterior as lateral radiographs using CNN, this demonstrated the promising result of 96% of the Area Under the Receiver Operating Characteristic Curve.

The study [16] proposes a new type of classification network which they call CrackNet (Crack Sensitive Convolution Neural Network). They propose 2 steps solution, first one is Faster R-CNN, and the result of this is sent to the second one which is CrackNet. This work achieved an accuracy of 90.11% and an f1-score of 90.14%. The limitation of this work is time-consuming as the images go through 2 stages. [17] uses the modified Ada-ResNeSt backbone network and AC-BiFPN method. This work uses data from the MURA dataset, achieving higher detection AP of 68.4%. The study in [18] helped in understanding bone disorders like the genetic bone disorder Fibrous Dysplasia (FD). In this study, 3 different ways for identifying bone fracture studied machine learning based is the last and the best one.

The studies [19] and [20] worked on the detection of rib fracture and wrist fracture respectively. Where the former uses deep learning and later uses the R-CNN for localization wrist radiographs. The approach of ResNet Faster R-CNN showed a good result with 98.1% of sensitivity, and 89.5 % of AUC. However, the limitation is this method more specific to ulna fractures on the wrist.

III. METHODOLOGIES

For achieving the success of this study a systematic approach has been followed. These methodologies combine the different stages of performing machine learning and deep learning task achievement. The study made use of a sophisticated framework like CRISP-DM to carry out the systematic analysis of the project.

Methodologies address the important steps like data collection, Business Understanding, Data Understanding, preparation of the data for later stages, and Applying 3 different models to the data. Later Evaluation of the model will be conducted.

The subsequent sections deep dive into the detailed descriptions of each methodology, comprehending on their technical aspects and potraying their contributions to advancing fracture detection in the medical field.

A. Data Collection:

The dataset used in this research study was sourced from Kaggle, a platform for data science and machine learning resources. This dataset is a collection of X-ray images, it provides a diverse and relevant set of data which comprises different types of bones and fractures. This dataset from Kaggle serves as a foundation for the study, allowing us to train and evaluate the deep-learning model. In this stage of the study finding a suitable dataset, downloading and analysing the compatibility of the dataset has been performed.

B. Data Understanding:

A comprehensive exploration of the obtained dataset sourced from Kaggle has been done. The dataset consists of a collection of X-ray images capturing different types of fracture scenarios.

For the initial assessment of image quality, a subset of images examines as shown in Figure 1. Figure shows the sample images of X-rays of bones belonging to fractured and not Fractured class.

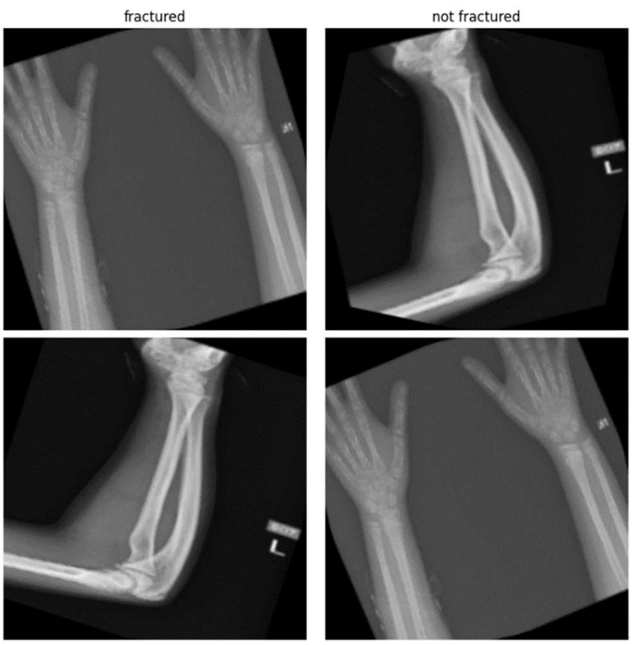

Fig. 1. Sample X-ray images belong to fracture and not fracture classes

The distribution of images belonging to fracture categories and non-fracture instances has been assessed. This will address whether class balancing is needed so that model should not be biased towards any category.

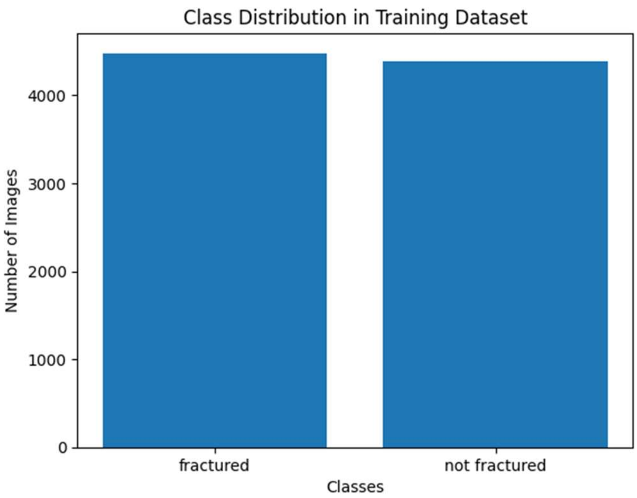

Fig. 2. Distribution of fracture and not fractured class images

The count plot displayed in Figure 2 shows the distribution of two classes in the dataset. The graph suggests that the distribution of these images is balanced. And there is no need of balancing.

C. Data Preparation

For the purpose of effective model training and evaluation having good quality of dataset is very important. As the dataset collected is raw and may contain different shapes, the Data preparation stage is very crucial to carry out.

1) Image Transformation:

To make sure consistency across different platforms, the colour space of images is converted from BGR to RGB using the cv2 library. This is an important stage as deep learning frameworks typically work on RGB images.

2) Image Resizing:

As the dataset may contain images which have different dimensions this may hamper the performance of the model. To overcome this issue, all the images are resized to a standard dimension. This is done with the help of resize function of the CV2 library.

3) Data Augmentation:

To diversify the dataset, augmentation techniques have been utilized. This technique will create the images out of the original images. These generated images will be used for the training. This stage involves introducing controlled variations to the preexisting dataset, This technique is very helpful especially when there is a limited dataset.

To achieve this study used the Keras framework's ImageDataGeneratore module. This module enables the application of different image transformations while model training.

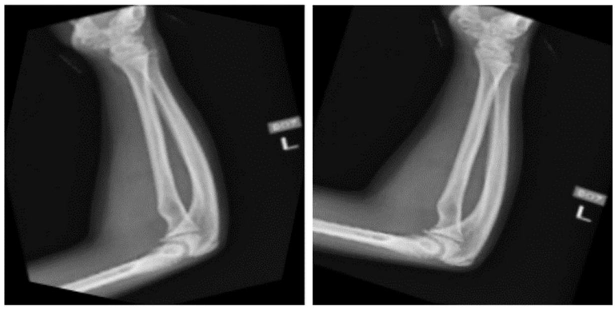
Fig. 3. Distribution of fracture and not fractured class images

The strategies used for the augmentation include Rotation with 15 degrees, shearing (which induces non-uniform deformation) of range 0.2, and zoom of randomly up to 20%. Horizontal Flip with 50% probability, Fill mode to fill the gap caused while transforming and finally width and height shift of up to 10%.

Figure 3 shows the instance where an alternative image is generated by using the image augmentation technique. This demonstrates one or more criteria for transforming the images.

By carrying out these data preparation steps, we set the stage for robust model training, where the dataset is well-preprocessed and augmented, allowing our deep learning models to learn effectively from the available data while minimizing the risk of overfitting.

4) Edge Detection:

In an effort to improvise the dataset and facilitate more robust fracture detection, the study integrated the Canny edge detection technique. The outcome of this technique is images with edges plotted, these images are used for the 2nd (Random Forest) and 3rd (CNN with EfficientNet B3) models.

This approach starts by converting the input RGB image to grayscale, representing it in a single channel. next, the Canny edge detection algorithm is applied, using specified threshold values. The resulting edge map is later converted back to an RGB format to enable integration with deep-learning models.

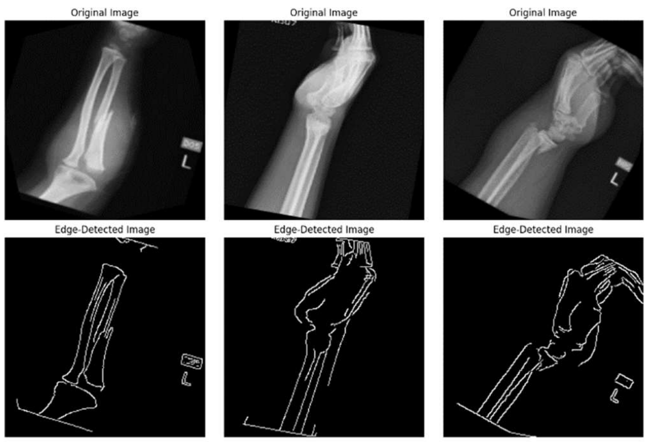

Fig. 4. Sample images before and after applying canny edge detection

Please use references to justify your choices in the methodology

The technique of Canny edge detection includes an additional dimension to our dataset. For each original image, the study generates an edge-enhanced version as shown in Figure 4. This can provide models with better-highlighted contours and fracture-related features, which allows them to better discriminate fractures from non-fracture samples. As can be observed from figure 4 the first image belong to the fractured class and with edge detection it is very easy to recognize the difference.

In the subject of data preparation, the integration of Canny edge detection aligns with the goal of providing models with a diversified dataset. By including edge-enhanced images in the dataset, the study aims to contribute to the model's feature extraction ability, improving its overall fracture detection accuracy.

D. Modelling

The modelling phase is a crucial phase in the project's attempt to automate bone fracture detection using machine learning and deep learning methodologies. Building on the top of the refined dataset which is prepared during the phase of data preparation. 3 different models have been developed including Ransom Forest and the two CNN models which operate on different types of preprocessed data. However, The concentration is on leveraging CNN and other pertinent architectures, capitalizing on their feature extraction capabilities and hierarchical learning representations.

1) Convolution Neural Network (Without Edge Detection):

To address our research question, the study opted to begin with a Convolutional Neural Network (CNN) approach. For the implementation of the CNN model, the study employed the TensorFlow library and made use of a pre-trained architecture, specifically the EfficientNetB3 model, initialized with weights of ImageNet. The designed model architecture is a sequential composition of layers. The pre-trained base model was used, and its layers were set as non-trainable for the purpose of preserving their learned features.

On top of this later, a Global Average Pooling layer followed by densely connected layers and dropout layers to allow the learning of higher-level features and prevent overfitting.

In addition, the performance of the model is monitored including accuracy, precision, recall and AUC. For optimization, the Study utilized the Adam Optimizer along with the binary cross-entropy loss function, which is best for binary classification tasks. Additionally, training and validation accuracy as well as training and validation loss has been monitored and plotted.

The training stage started, with the model being trained over multiple epochs. During every epoch, training and validation data were fed through the model, and the parameters were updated iteratively to minimize the loss function. The model went through 15 epochs, with each epoch consisting of 50 steps per epoch. The training history was monitored and stored with the help of the custom callback, recording metrics such as accuracy, loss, as well as validation loss.

In conclusion, this initial model employed a CNN-based approach to develop fracture detection. This model acted as the initial stage of modelling.

2) Edge Detection and Random Forest Approach for Bone Fracture Detection (With Edge Detection)

For the second model study made use of the images generated after complete preprocessing stages including Canny Edge detection. The motivation behind the development of the second model is to compare the ability of a machine learning model like Radom Forest with the deep learning model.

For the machine learning context, a Random Forest classifier is developed. This model is trained on the features extracted from the resized and flattened edge-detected X-ray images. This model learns to distinguish between fractured and non-fractured bone radiographs based on the edge patterns of the images. After the training process, the model's performance is evaluated against the testing subset using evaluation metrics such as accuracy, precision, recall, and F1-score.

The edge detection and Random Forest approach provide a unique perspective on bone fracture detection in X-ray images. This methodology not only uses image processing techniques for feature extraction but also integrates machine learning for classification. This development was mainly for comparative study and assessment of machine learning models in analysing image data.

3) Convolution Neural Network (with Edge Detection)

The 3rd model in the study incorporates an innovative methodology that clubs edge detection, transfer learning using the EfficientNetB3 architecture, and a custom data generator to enhance the accuracy of bone fracture identification in X-ray images.

The core of the model development involves harnessing transfer learning with the EfficientNetB3 architecture on the top on-edge detected images. This pre-trained model, assigned with weights from the ImageNet dataset, is loaded and augmented with extra layers for bone fracture detection. Similar to the first model, the pre-trained layers are set to preserve their learned features, while the newly added layers are trained and fine-tuned for the fracture detection task. This approach makes use of the features learned by the EfficientNetB3 base model while adapting it to the task at hand which is fracture detection. In addition, training and validation accuracy, as well as training and validation loss, are plotted for the understanding of accuracy and loss trends over the course of training.

In summary, the third model clubs edge detection, transfer learning utilizing EffivientNetB2, and a custom data generator for evaluation. This gives a promising way of image detection.

IV. Evaluation and Result Interpretation

The application of different methodologies to the bone fracture detection task showed insightful results, showcasing the efficacy and potential impact of each of the above-proposed approaches in improvising the accuracy. Each methodology was evaluated using different performance metrics, which gives knowledge of their strengths and limitations while providing valuable insights into their implications for medical diagnostics.

1) Convolution Neural Network (Without Edge Detection):

The CNN methodology was employed to address the bone fracture detection challenge, This model is developed as an initial model. Evaluation is done through a comprehensive set of metrics, including accuracy, precision, recall, and area under the curve (AUC).

The trained CNN model achieved an accuracy of approximately 85.75%, a precision of 85.17%, a recall of 86.20%, and an AUC of 93.49% for the training dataset. These metrics indicate that the model achieved a balanced trade-off for correctly identifying fractured cases (recall) and minimizing the instances of misclassifying non-fractured cases as fractures (precision) for the training set evaluation. The high AUC value further validates the model's ability to distinguish between the two classes.

However, It achieved an accuracy of around 63.33%, a precision of 54.39%, a recall of around 51% and finally AUC of 69.5% on the validation dataset. This indicates model doesn't perform as well as it does on the training dataset.

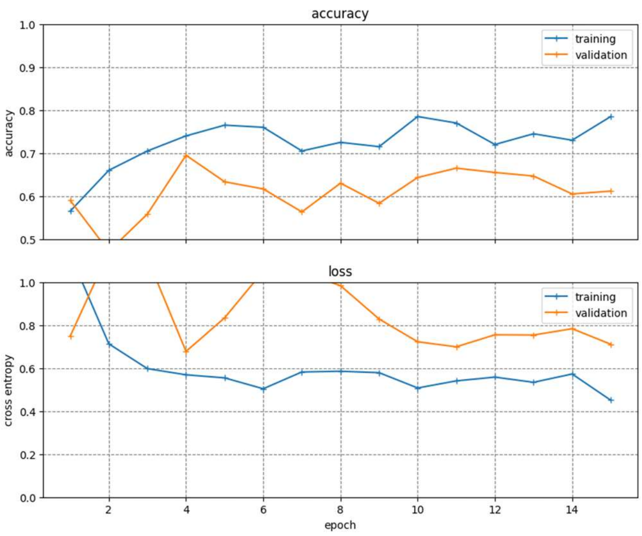

Fig. 5. Training curve showing accuracy and loss per epoch

The graph displaying accuracy and loss per epoch in Figure 5 presents valuable insights into the training dynamics of our deep learning model. It shows the evolving performance of the model as it learns from the training data over a number of iterations.

Starting with the accuracy trend, the study observes a significant progression in the ability to make the right predictions. Initially, during the initial epoch, the accuracy stands at approximately 56.5%, indicating a mid-level of correct predictions. As training proceeds, the accuracy consistently increases, reaching around 78.5% at its highest point. This upward trajectory signifies the capacity of the model to gradually identify patterns including complex ones resulting in better accuracy while making predictions.

On the other hand, the loss trend shows a compelling narrative of the model's improvement process as shown in Figure 5. At the second epochs loss function with the value of 1.13 which suggests a significant gap between prediction and actual value. However, as each epoch progresses, the loss decreases gradually. Which signifies the model's ability to reduce error.

While the initial stages show consistent improvement, the accuracy trend does show few fluctuations in the later epochs, ranging from 70% to 78.5%. This behaviour might show that the model's learning rate has slowed down and might lead to overfitting. Additionally, the loss graph varies within the range of 0.45 to 0.58 corroborating the model's slow convergence to a minimum error point.

In short, CNN exhibited promise in fracture detection in terms of the training data like accuracy, precision, recall and AUC. However, it is proven not great model for validation data. This is also proved with the training curve which is displayed in Figure 5.

2) Edge Detection and Random Forest Approach for Bone Fracture Detection (With Edge Detection)

The combination of Edge Detection and Random Forest methodology offered a different approach to bone fracture detection, leveraging edge information extracted from X-ray images and utilizing a Random Forest classifier. The evaluation of this methodology produces good results, highlighting the model's potential in accurately categorizing X-ray images as fractured or non-fractured classes.

The Random Forest model showed an accuracy of 87.6%, a precision of 88.07%, a recall of 87.29%, and an F1 score of 87.68%. These performance metrics demonstrate strong overall performance and a well-balanced trade-off between precision and recall. The model's accuracy demonstrates its ability to make accurate predictions, while a good F1-score indicated the model's ability to handle imbalanced data.

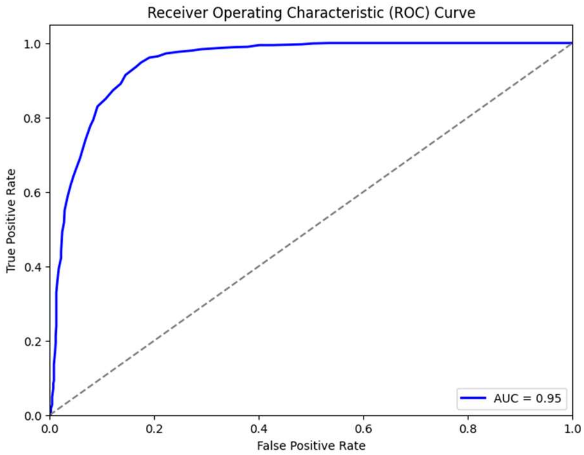

Fig. 6. Receiver Operating Characteristic Curve for Random Forest Model

The AUC curve for the Random Forest model shown in Figure 6, demonstrates a very good result of 0.95. This result indicates that the Random Forest model is able to accurately distinguish between the positive and negative cases (fractured and non-fractured classes). This promising result of AUC shows the reliability of the model in the context of medical data. However, the Radom Forest model is not a great choice for image analysis in the general context.

In summary, Edge Detection combined with Random Forest methodology illustrated impressive performance in bone fracture detection. The balanced precision-recall metrics and F1 score emphasize its usability in clinical settings, improvising diagnostic accuracy.

3) Convolution Neural Network (with Edge Detection)

The Convolution Neural Network with Canny Edge Detection and Transfer Learning methodology showed a comprehensive approach to bone fracture detection, combining CNN with Canny edge detection preprocessing and improvising a pre-trained model for feature extraction. The evaluation of this methodology showed very good results, showing the model's potential to deliver accurate and reliable fracture detection results.

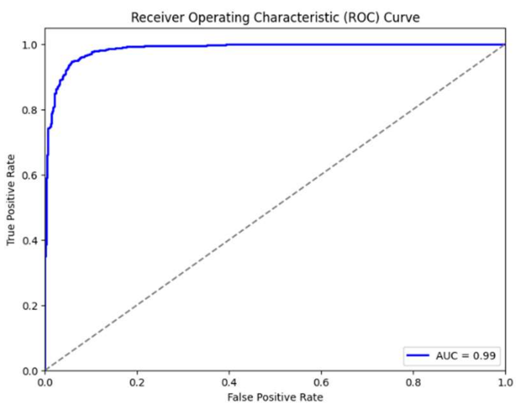

Fig. 7. Receiver Operating Characteristic Curve for Random Forest Model

The model achieved an accuracy of approximately 93.91%, a precision of 91.56% and a recall of 96.58%. These performance metrics showed the effectiveness of the methodology in classifying bone fracture cases accurately. In addition, the model showed a really good score for the area under the curve of 0.98 as shown in Figure 7. The high recall and AUC values show the ability of the model to identify a significant portion of true positive cases and its strong ability to distinguish between the fracture and non-fractured images.

The validation loss and accuracy obtained further validate the methodology's robustness, with a validation loss of 0.1555 and a validation accuracy of nearly 94.59%. These results emphasize the model's capacity to generalize well to unseen data and make accurate predictions on unseen X-ray image data.

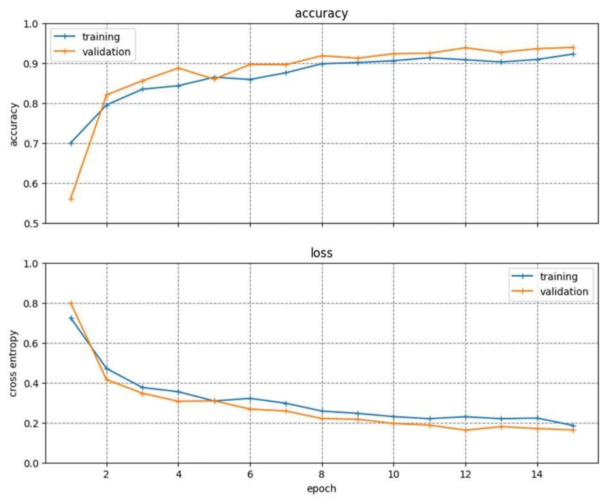

Fig. 8. Training curve showing accuracy and loss per epoch

In addition, the training progress of the final model across 15 epochs shows a compelling evolution of performance metrics as shown in Figure 8. The model starts with a loss value of 0.7248 and eventually refines its predictions, leading to a reduction in both losses. The accuracy increases from around 0.70 to as high as 0.92, signifying the model's enhanced ability to make accurate predictions.

Precision and recall are crucial indicators of a model's performance consistently improving over epochs. The precision, demonstrating the ratio of correctly predicted positive cases, progresses from 0.69 to 0.92. Similarly, recall, which shows the proportion of actual positives correctly predicted by the model, increases from 0.70 to 0.92. These increasing trends show the capacity of the model to recognize fractures more accurately while minimizing false positives and negatives.

Table 1 displays the comparison of the different evaluation matric values obtained after the evaluation of the models. The result shows the CNN model trained on the top of the canny edge detected images, performs the best among all the other models. And the results are very promising.

TABLE 1. COMPARISON OF THE EVALUATION RESULTS OF 3 MODELS

	CNN	Random Forest (With Edge Detection)	CNN (With Edge Detection)
Accuracy	0.85	0.87	0.93
Precision	0.85	0.88	0.92
Recall	0.86	0.87	0.97
AUC	0.93	0.95	0.98

Overall, the CNN with Canny Edge Detection and Transfer Learning methodology demonstrated outstanding performance in the primary task. The combination of edge information, transfer learning, and deep neural networks led to a model with high accuracy, precision, and recall values. The strong results have meaningful decision-making processes and contribute to more accurate fracture diagnoses, thereby potentially making patient outcomes better.

V. CONCLUSION AND FUTURE WORK

In conclusion, The study deep dives into the field of bone fracture detection with the help of advanced deep learning methodologies. Through thorough exploration of distinct models, the study demonstrated the potential of Convolutional Neural Networks (CNNs) and Random Forest in identifying fractures from X-ray images accurately. The integration of edge detection techniques and transfer learning further improved the accuracy of the model, with impressive AUC values demonstrating their efficacy.

To discuss the limitation. Even though the data collected for this study has a good number of images, all the images are from the same dataset. This is the limitation of the study. In the real world, there could be many varieties of medical challenges and hence different kinds of X-ray images.

Given more time, several routes could be explored to improvise and extend our study. Fine-tuning hyperparameters of the model, exploring more advanced techniques for data augmentation, and including additional clinical metadata could further enhance model performance. Including more different and diverse datasets which contain different possible cases of fractures could lead to models that are better suited for real-world applications.

In summary, the study advances the comprehension of using deep learning for bone fracture detection. While offering valuable insights, it also highlights the need for continued research and validation to ensure the practical viability of the model in real clinical contexts. The study's contributions lie in the fusion of technology and medicine, looking forward to making a positive impact on patient care through accurate and automated fracture detection.

REFERENCES

[1] Yadav, D. P., and Sandeep Rathor. "Bone fracture detection and classification using deep learning approach." 2020 International Conference on Power Electronics & IoT Applications in Renewable Energy and its Control (PARC). IEEE, 2020.

[2] H.Y Chai, K.L Wee, T.T Swee, S.H Salleh, A.K Ariff and Kamarulafizam, "Gray-Level Co-occurrence Matrix Bone Fracture Detection", American Journal of Applied Sciences, vol. 8, pp. 26-32, 2011.

[3] Yadav, Dhirendra Prasad, et al. "Hybrid SFNet model for bone fracture detection and classification using ML/DL." Sensors 22.15 (2022): 5823.

[4] Sahin, Muhammet Emin. "Image processing and machine learning‐based bone fracture detection and classification using X‐ray images." International Journal of Imaging Systems and Technology 33.3 (2023): 853-865.

[5] Dupuis, Michel, et al. "External validation of a commercially available deep learning algorithm for fracture detection in children." Diagnostic and Interventional Imaging 103.3 (2022): 151-159.

[6] Jones, Rebecca M., et al. "Assessment of a deep-learning system for fracture detection in musculoskeletal radiographs." NPJ digital medicine 3.1 (2020): 144.

[7] Raisuddin, Abu Mohammed, et al. "Critical evaluation of deep neural networks for wrist fracture detection." Scientific reports 11.1 (2021): 6006.

[8] Zhang, Bin, et al. "Improving rib fracture detection accuracy and reading efficiency with deep learning-based detection software: a clinical evaluation." The British Journal of Radiology 94.1118 (2021): 20200870.

[9] Sharma, Ashish, et al. "Bone fractured detection using machine learning and digital geometry." Mobile Radio Communications and 5G Networks: Proceedings of MRCN 2020. Springer Singapore, 2021.

[10] Mutasa, Simukayi, et al. "Advanced deep learning techniques applied to automated femoral neck fracture detection and classification." Journal of Digital Imaging 33 (2020): 1209-1217.

[11] Guan, Bin, et al. "Thigh fracture detection using deep learning method based on new dilated convolutional feature pyramid network." Pattern Recognition Letters 125 (2019): 521-526.

[12] Krogue, Justin D., et al. "Automatic hip fracture identification and functional subclassification with deep learning." Radiology: Artificial Intelligence 2.2 (2020): e190023.

[13] Joshi, Deepa, and Thipendra P. Singh. "A survey of fracture detection techniques in bone X-ray images." Artificial Intelligence Review 53.6 (2020): 4475-4517.

[14] Guan, Bin, et al. "Arm fracture detection in X-rays based on improved deep convolutional neural network." Computers & Electrical Engineering 81 (2020): 106530.

[15] Kitamura, Gene. "Deep learning evaluation of pelvic radiographs for position, hardware presence, and fracture detection." European journal of radiology 130 (2020): 109139.

[16] Ma, Yangling, and Yixin Luo. "Bone fracture detection through the two-stage system of crack-sensitive convolutional neural network." Informatics in Medicine Unlocked 22 (2021): 100452.

[17] Lu, Shuzhen, Shengsheng Wang, and Guangyao Wang. "Automated universal fractures detection in X-ray images based on deep learning approach." Multimedia Tools and Applications 81.30 (2022): 44487-44503.

[18] Saranya, A., and K. Kottilingam. "A survey on bone fracture identification techniques using quantitative and learning based algorithms." 2021 international conference on artificial intelligence and smart systems (ICAIS). IEEE, 2021.

[19] Wu, Mingxiang, et al. "Development and evaluation of a deep learning algorithm for rib segmentation and fracture detection from multicenter chest CT images." Radiology: Artificial Intelligence 3.5 (2021): e200248.

[20] Thian, Yee Liang, et al. "Convolutional neural networks for automated fracture detection and localization on wrist radiographs." Radiology: Artificial Intelligence 1.1 (2019): e180001.