An Comparative Analysis On Image Processing For Bone Fracture Detection System Using Machine Learning Techniques

[1] Dr. Vipul .A Shah
Instrumentation and Control Engineering
Dharmsinh Desai Univeristy
Nadiad,Gujarat,India
e-mail : vashah.ic@ddu.ac.in

[2] Hemantika P Chauhan
Ph.D scholar, Instruemntation and Control Engineering
Dharmsinh Desai Univeristy
Nadiad, Gujarat, India
e-mail:omacademy.infoeducation@gmail.com

Abstract— *When X-ray imaging is insufficient to identify a wrist fracture, other imaging modalities such as computed tomography and magnetic resonance imaging are often used. There are cases when radiologists fail to detect fractures due to their opacity. But some cracks are obvious and only slow down the radiologists due to reporting systems. This study compares the various machine and deep learning approaches for bone fracture detection on wrist X-ray image data from publicly available. Included in this collection were X-ray images of 82 healthy wrists and 111 fractured ones. They may put the model to the test by dividing the data set into two parts: training and testing. For fracture diagnosis, the RN-21CNN model—which is based on residual networks—was used. A confusion matrix was used to produce assessment metrics such as F1-score, recall, accuracy, and precision. The RN-21CNN model performed better than the others, with a 97% accuracy and 95% recall rate. These findings highlight the model's promise for practical applications in fracture diagnosis and treatment planning by demonstrating its ability to reliably identify wrist fractures from X-ray images. Future work needs to concentrate on reducing data set biases, increasing data set variety, and investigating cutting-edge deep learning methods to raise the effectiveness of fracture detection models.*

Keywords— *Bone Fracture Detection, Wrist, X-Ray Images, Machine Learning, Diagnosis, Deep Learning.*

I. INTRODUCTION

The development of more sophisticated hardware and software has made medical image processing an increasingly well-respected field of study in the healthcare sector. When a bone is subjected to a force that is greater than its structural resistance, a fracture may form. Modern society places a premium on early diagnosis and treatment of ailments, and one of them is bone fractures, which impact a large population. Developed nations are not immune to the widespread issue of bone fractures, and the rate of fractures is rising sharply. A bone fracture might be the result of a common injury or a medical condition. Fast and precise decision-making is essential for the success of any recommended therapy [1]. There are a lot of individuals who have fractures, and if they ignore one, it may cause serious harm or even death, according to new research from the WHO. Fractures of the wrist are fairly frequent, and the number of confirmed cases is rising every day [2]. It is common practice for radiologists and experts to use X-ray images to confirm the presence of a fracture and pinpoint its exact position. It takes a lot of time and effort to use manual assessment or the old-fashioned X-ray framework to find fractures.

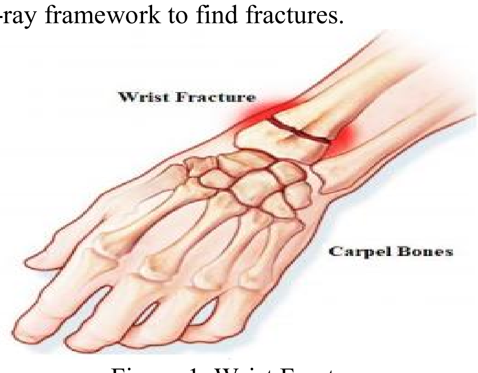

Figure 1. Wrist Fracture

Radiologists assess X-ray samples visually to ascertain if specific bone fractures are present. It requires a great deal of time and effort to manually check fractures as part of the radiological interpretation process. A high probability of false detection and poor fracture healing might also be the consequence of a shortage of specialized radiologists in overcrowded clinical settings, an absence of medical specialists in areas with insufficient medical resources, or simply weariness from excessive workloads [3]. It is common practice to first use traditional radiography or X-ray imaging when making a diagnosis of a fractured wrist [4]. Machine learning is a technique that uses algorithms and approaches to improve upon training data in order to make predictions autonomously [5]. A subset of supervised learning relies on receiving training data that has already been labeled. Among machine learning tools, deep learning has shown to be the most useful in computer vision and general imaging [6]. "Deep learning" refers to the latest developments in artificial neural networks that increase performance and abstraction by adding more network layers.

A. Contribution of the study

A main goal of this research to compare various machine and deep learning approaches for detection of fracture. The key contribution of this study as follow:

- To collect the wrist X-ray image data from Mendeley web page.
- To preprocess the data for normalization and augmentation.
- To assess how well machine learning algorithms identify wrist fractures in pictures.
- To compare the model's efficiency with performance measures like F1-score, recall, accuracy, and precision.

The remaining sections of the paper are structured as follows: Some prior research on bone fracture detection is detailed in **Section** II. **Section** III details the process by which the fracture detection was used, and **Section** IV presents and analyses the experimental data. Lastly, **Section** V presents the findings.

II. LITERATURE REVIEW

This research delves into the latest methods for fracture detection that make use of various deep learning and machine learning techniques.

In M, Imrankhan (2023) [7] suggested a ResNet 50-based deep neural network model for bone injury detection. However, our DL model has a tendency to overfit when presented with sparse data. In order to get around this issue, data augmentation techniques are employed to expand the dataset. Our model achieves an object discrimination accuracy of 92.44% using this approach, which employs 5-fold cross-validation.

In Vasker, Nishat and Hasan (2023) [8] presents an innovative strategy to fracture diagnosis that addresses these issues by leveraging AI-assisted methods, including deep learning. To provide real-time broken bone identification and classification, a deep neural network (DNN) technique is employed. The suggested model achieves a remarkable 92.44% accuracy rate in differentiating between healthy and broken bones by employing 5-fold cross-validation. The accuracy is also higher than 93% when tested on 20% of the data and 95% when tested on 10% of the data. The results show that the model that was created is better than the ones that are already out there.

In Moon, Gwiseong and Kim (2022)[9] present a CA-FBFD method that uses computers to help diagnose facial bone fractures in order to overcome these difficulties. The object identification model YoloX-S is utilized by this system. YoloX-S was trained utilizing CT image Mixup data augmentation and solely IoU Loss for box prediction. While testing included a variety of face fracture data, training relied solely on data pertaining to nasal bone fractures. With an average accuracy of 69.8 percent for facial fractures, the CA-FBFD system significantly surpassed the baseline YoloX-S model by 10.2 percent across the board. Furthermore, the CA-FBFD system outperformed the baseline YoloX-S model by 66.7% with a sensitivity/person of 100% for facial fractures. As a result, the CA-FBFD method can significantly reduce the amount of work that clinicians have to do when using facial CT to diagnose bone fractures.

In Joshi, Deepa and Singh, (2022) [10]suggested design is an adaptation of the mask-RCNN architecture that transfers a weights by a surface crack dataset to a wrist X-ray dataset to improve model convergence. The results of the modifications performed at the sub-architecture level (levels 1 and 2) are examined. By including the level 1 and 2 modifications, they were able to improve the results compared to the traditional mask-RCNN model on the wrist fracture dataset. They were able to attain an average accuracy of 92.278% for fracture identification and 79.003% for fracture segmentation on a 50 0 scale, and 52.156% on a 75 0 scale, etc. In H. P. Nguyen (2021) [11] offer an innovative method for detecting fractures in X-ray images of the arm's bones using DL. They begin applying a method that combines YOLACT++ for picture segmentation and Contrast Limited Adaptive Histogram Equalization for picture contrast enhancement to the X-ray image to better prepare it for further processing. Their proposed strategy yielded an 81.91% maximum result. Even on the tiny dataset, our strategy beats the Faster-RCNN based solution, according to the experimental findings.

The following table 1 shows the comparative analysis of Bone fracture detection using various methods and techniques.

Table 1: Comparative analysis of bone fracture detection during various methods and techniques

Author	Methodology	Data set	Findings	Limitations/Challenges
Olczak et al. [12]	Deep learning networks	Wrist, hand, and ankle radiographs	Fracture accuracy: 83%	the low picture quality in orthopaedic radiographs did not prevent deep learning networks from accurately identifying important image features.
Dimililer [13]	Intelligent bone fracture detection system (IBFDS)	30 training images, 70 testing	94.3%	Small dataset size, lack of external validation, performance may vary on larger datasets.
Ebsim et al. [14]	CNNs with random forest regression	Lateral and posterioanterior radiographs	AUC: 96	Radiographs of the wrist have not yet been extensively used for fracture identification.
Thian et al. [15]	Inception-ResNet Faster R-CNN	7356 wrist images	91.8%	finding effective ways to train deep learning models with sparse data from a particular class is a hot topic in computer vision research right now (24).
Kim and MacKinnon [16]	CNNs, pre-training	Wrist images	95.4%	Limited to wrist fractures, potential bias in pre-training on non-medical images.
Jiménez-Sánchez et al. [17]	CAD tool, DL algorithms	Musculoskeletal X-ray images	87% (organizing), 94% (recognition)	Lack of external validation, potential bias in CAD tool implementation.
Chada [18]	Deep transfer learning, novel architectures	MURA dataset	83%-92%	Limited to upper limb abnormalities, potential bias in using public dataset.
Cheng et al. [19]	Deep convolutional networks (DCN), Grad-CAM	Hip fracture dataset	91%	Lack of interpretability in Grad-CAM, need for further validation on diverse datasets.

A. Research gaps

There are still several problems to overcome in a field of bone fracture detection, even though DL and machine learning have made significant advances. Problems occur when datasets are too small or lack diversity to adequately represent the whole range of fracture forms, sites, and patient demographics observed in clinical practice. Over fitting and low real-world applicability could result from established models' limited generalization and resilience. The absence of external validation just makes matters worse, since the results of the model could not be applicable to new data or patients. Patients may arrive with fractures in a variety of locations, and many current techniques are only applicable to certain kinds of fractures or parts of the body. This limits their use in complete clinical contexts. Although deep learning models are very accurate, they are still not easily Interpretable, which makes it hard to incorporate them into clinical procedures.

III. METHODOLOGY

A comprehensive assessment of each of a DL model that were utilized is provided in this section. The technique begins with a description of the data set and flowchart utilized in the experiments in Section.

For the purpose of identifying bone fractures, the research technique that was utilized in this study included a number of important phases. To begin, X-ray images of the wrist were taken by a publicly accessible Mendeley data set. This data set included 211 fractured instances and 82 normal cases. To maintain equality across the data set, the preprocessing step entailed scaling all images to a standard resolution of 512 × 256 × 3 pixels. This was accomplished by the utilization of bi-cubic interpolation. Subsequently, the data was normalized to ensure consistency. After enhance an overall quality and diversity of the data set, several data augmentation techniques, such as affine transformations, were utilized. The next step in getting the model ready for deployment was to split the data set into two parts: training and testing. For the classification method, a machine and deep learning model was utilized. This model was constructed on a residual network architecture, and it included particular requirements for layer configurations and activation functions. The model was evaluated using a variety of performance measures, including as F1-score, recall, accuracy, and precision. Presented below is a flow diagram depicting the data set:

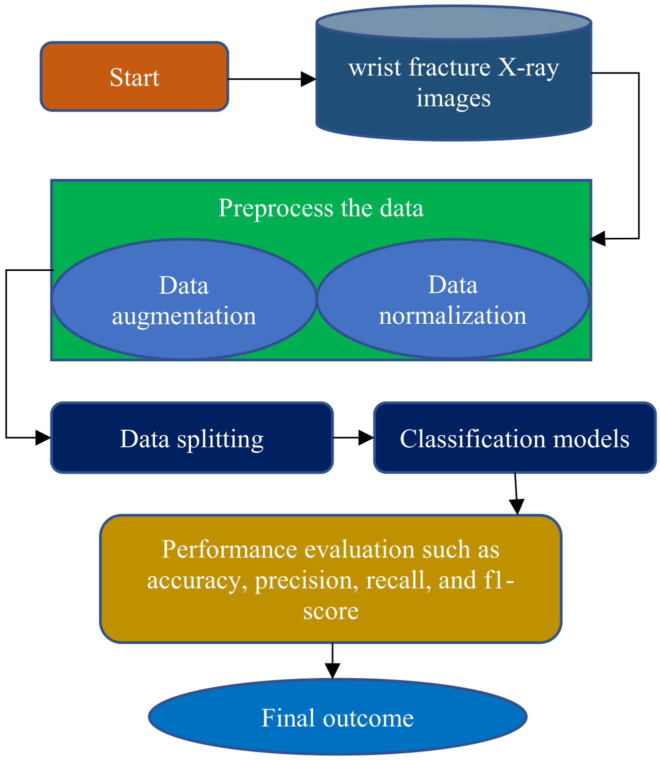

Figure 2: Proposed flowchart

The proposed flowchart shows the following step like data collection, preprocessing, splitting. Classification models, and evaluation parameters are described in below.

A. Data collection and preprocessing

The publicly accessible Mendeley data set was used to get the wrist fracture X-ray pictures of patients that were utilized in our study. There are a total of 111 X-ray pictures of broken wrists and 82 normal wrists in the collection. The JPG formatted images taken in real time come in a range of sizes. Prior to any further processing, the photos underwent preprocessing. During the preprocessing step, they enhanced and resized the images. By using the bi-cubic interpolation approach, all of the photos were downsized to 512 × 256 × 3 pixels' resolution. To calculate the weighted average in the bi-cubic approach, 4 × 4 adjacent pixels were used.

1. Data Normalization

Data normalization has long been a major part of preprocessing, which aims to remove inconsistent or redundant data from the database, control the network's complexity, and provide reliable results. The real-time visuals are blurry and pixelated for a variety of reasons. Resizing the image is necessary so it fits the dimensions of the model. factors. Resizing the image is necessary so it fits the dimensions of the model. The X-ray images utilized for this research were in JPG format and were in a range of sizes. When doing calculations, the closest pixels in both the vertical and horizontal dimensions were given more weight. Given that they wanted to investigate the new pixel from the point with coordinates (x, y), they used the following equation to get its value:

$$v(x,y) = \sum_{i=0}^{3}\sum_{i=0}^{3} a_{ijx^i y^i} \ldots\ldots.(1)$$

2. Data augmentation

An efficient and successful method to lessen the "over fitting" of the models brought on by insufficient training photos is data augmentation, which calculates the data probability space by adjusting input images by random cropping, rotation, scaling, and noise disruption. Improving the dataset's volume, quality, and diversity may generally boost the model's performance. In this work, the affine transformation approach was used to enhance photographs.

B. *Data splitting*

For the purpose of model implementation, split the data into training and testing. Whenever data is split, an algorithm is trained on a test set of data, and a training set of data is set apart. The model train on the training data set.

C. *Classification techniques*

Select the most appropriate classification model to build a reliable system for fracture detection. The models that follow are:

1) *RN-21CNN Model*

The RN-21CNN model is based on a kind of network known as a residual network. Within the realm of mathematical statistics, the residual network encompasses the dispersion of expected and actual values. The model may put an emphasis on learning these little substitutions by showing the residual idea to delete the same piece and highlighting the slight variances. Researchers have shown that this method fixes the problem where fitting effects degrade with more neural network layers. Consequently, the network's input is changed to a 150 × 150 picture block after preprocessing and is then complicated by a sequence of 21 layers. Next, in the convolution layer, they choose a 1 × 1 and 3 × 3 filter size. To get greater non-linear activation, they are now using the ReLU function with a convolution part that is less than 5 × 5 and 7 × 7. This allows us to simply guarantee that each neuron is small enough to be trained to the input's accessible field in order to capture local texture characteristics associated with the desired output. In this case, the whole network was unable to provide accurate results, even if the previous convolution network performance was satisfactory. Former convolution network flatly disregarded a lot of data that is strongly related to the goal because of their little size and excessive dimensionality reduction. Consequently, they have provided a CNN model based on residual networks to guarantee that the purpose of the original network is preserved while the aim of the image's size is reduced. Figure 3 shows the RN-21CNN model's design.

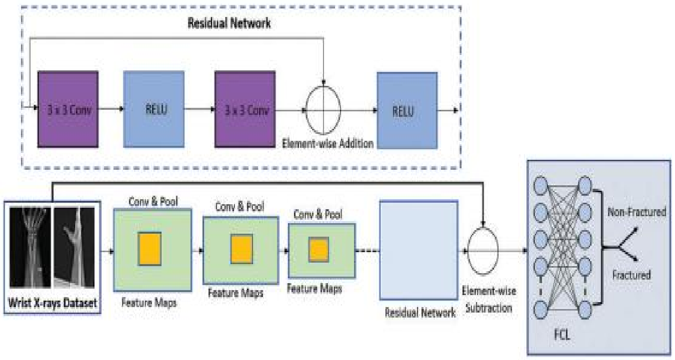

Figure 3: RN-21CNN architecture

As shown in Figure 4, there are standards for raising the convolution. Therefore, protecting the feature map's resolution from significant damage is achieved by raising the convolution supports, which in turn increases the receptive area of the convolution section. The sensitive input region undergoes non-linear input value adjustment after each convolution layer with the addition of a batch normalization layer. This prevents gradient vanishing, a problem that becomes more prevalent as training and architectural complexity increase. The global average pooling layer takes the output of the previous convolution layer and uses it to determine the mean of all feature maps.

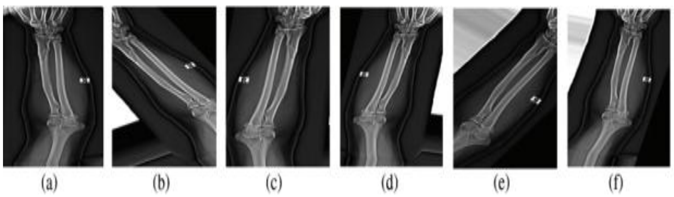

Figure 4: Sample images after data augmentation

Two linked layers may be produced, with the number of feature maps in the final product being equal to that of the previous layer. The first layer has 128 nodes, while the second is called the classification layer. They have included a dropout layer before a global pooling layer to prevent the model from being overfit. The value has been changed to 0.5. Unavoidable units that depend on detailed inputs are trained using a modified procedure that randomly terminates half of the neurons [20].

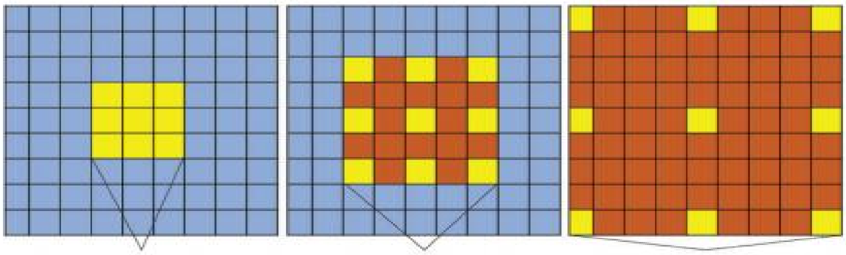

Figure 5: Three various dilated rates

An activation task is a crucial part of a neural network model for learning and interpreting a very complicated and non-linear workload. The activation function is responsible for presenting the non-linear properties to the neurons. This allows the neural network to estimate any non-linear function at random and has a significant impact on the pace of convergence. It was shown that using ReLU prevented gradient vanishing and reduced processing costs compared to standard sigmoid options [21][22]. Consequently, lowering the over-fitting problem, increasing system sparsity, and decreasing parameter dependency are all outcomes of having many neurons' ReLU values set to zero, which is a really amazing conclusion. They have used the ReLU function to activate every of the convolutional layers in our research study. In order to activate the standard ReLU for the fully connected network layout, the first dense layer is used.

D. Model Evaluation

A wide variety of performance metrics can be employed to assess the efficacy of algorithms built for fracture detection. Model evaluation includes an use of accuracy, F1-score, recall, and precision.

IV. COMPARATIVE RESULTS & DISCUSSIONS

Provide details on the datasets retrieved from various sources and the results of our analyses in this section. Results for the evaluations using the models are shown to be f1-sore, recall, precision, and accuracy.

A. Data set Description

The X-ray pictures of patients' wrist fractures used in the research were obtained from the publicly accessible Mendeley data set. Mendeley collects X-ray images from the Al-Huda Digital X-ray Laboratory in Multan, Pakistan, which is located on Nishtar Road. The data set contains X-ray images of 111 fractured wrists and 82 uninjured ones. Shot live in Multan, Pakistan, the film was captured at the Huda Digital X-ray Laboratory on Nishtar Road. The sample has a total of 111 fractured wrists and 82 healthy ones.

B. Performance matrix

Metrics for assessment provide light on how well a model performs. An important quality of assessment metrics is their capacity to distinguish between various model outputs.

a) Confusion Matrix

An approach to demonstrating the efficacy of a classification system is the confusion matrix. For models with more than two classes or when a number of instances in every class is uneven, relying just on accuracy might be misleading. A confusion matrix has the potential benefit of making it easy to determine whether the system often mislabels one class as another, or if it confuses between two classes.

Table 2: Confusion Matrix

	Actually Positive (1)	Actually Negative (0)
Predicted Positive (1)	True Positives (TPs)	False Positives (FPs)
Predicted Negative (0)	False Negatives (FNs)	True Negatives (TNs)

b) Accuracy

This metric measures how well a model predicts outcomes relative to the number of samples used as inputs. It is provided as-

$$Accuracy = (\mathrm{TP} + \mathrm{TN})/\mathrm{TP} + \mathrm{Fp} + \mathrm{TN} + \mathrm{FN} \ldots .(2)$$

c) Precision

A good predictor of future positive outcomes is the ratio of the number of actual positive results to the number of projected positive results. The way it is stated is-

$$Precision = (\mathrm{TP})/(\mathrm{TP} + \mathrm{FP}) \ldots. (3)$$

d) Recall

It is calculated by dividing the total number of relevant samples by the number of accurate positive outcomes. It may be expressed mathematically as -

$$Recall = (\mathrm{TP})/(TP + FN) \ldots\ldots. (4)$$

e) F1 score

Accuracy in testing is assessed using it. A score of F1 is obtained by summing the two metrics of precision and recall. "F1 Score" may take on values between zero and one. Mathematically, it is given as-

$$F1 - Score = \frac{2(Precision * Recall)}{\mathrm{Precision} + \mathrm{Recall}} \ldots. (5)$$

The receiver operating characteristic (ROC) curve shows how changing the discriminating threshold affects the ratio of true positives (Sensitivity) to erroneous positives (100 - Specificity). Each point on the ROC curve represents a decision threshold and its associated sensitivity and specificity.

C. Experimental results

In this section offers a result of the models of a models used in this research for wrist fracture detection.

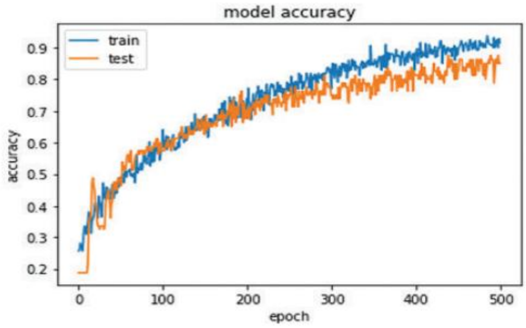

Figure 6: Accuracy graph of RN-21CNN model

The above figure 6 shows the accuracy graph of RN-21CNN model. The CNN model ran through all 500 epochs. Training achieved a maximum accuracy of 0.99 while validation achieved a maximum accuracy of 0.95.

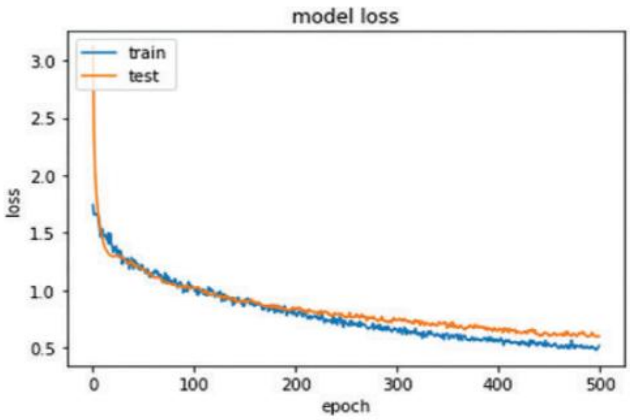

Figure 7: loss graph of RN-21CNN model

A loss graph of RN-21CNN model present in figure 7. A total of 0.021 and 0.026 were lost during training and validation. The CNN model used 500 epochs of execution.

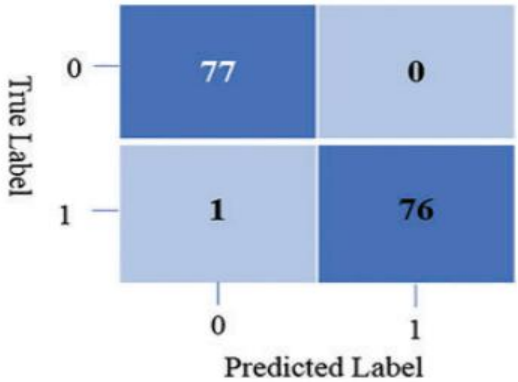

Figure 8: Confusion matrix of RN-21CNN model

Figure 8 displays the RN-21CNN model's confusion matrix. In this figure, x-axis shows a forecasted labels and y-axis present a true label. Class 0 represent the false negative (77) and true negative (0). Class 1 represent the true positive (1) and false positive (76).

D. Comparative analysis

This section showcases the outcomes of our suggested approach applied to a data set. An accuracy of our model's predictions is assessed here. Bar graphs, tables, and figures are the formats used to display the results.

Table 3: Comparison of model's performance on different parameter

Models	**Accuracy (%)**	**Precision**	**Recall (%)**	**F1-score**
SVM[23]	92	93	97	-
DCNN-LSTM 2 [24]	88.24	87.93	92	0.90
YOLO 512 [25]	95	95	95	0.95
Inception-v3 [16]	93	90	88	0.89
RN-21CNN	97	0	95	0.97

The table provides performance metrics that were assessed on a particular task for a range of machine learning models. RN-21CNN achieved the highest accuracy of 97%, with recall at 95%. The YOLO 512 model achieve a 95% accuracy with precision, recall and f1-score is same 95%. Overall, RN-21CNN exhibited the highest accuracy and recall among the models evaluated, while YOLO 512 demonstrated uniform performance across F1-score, recall, and precision.

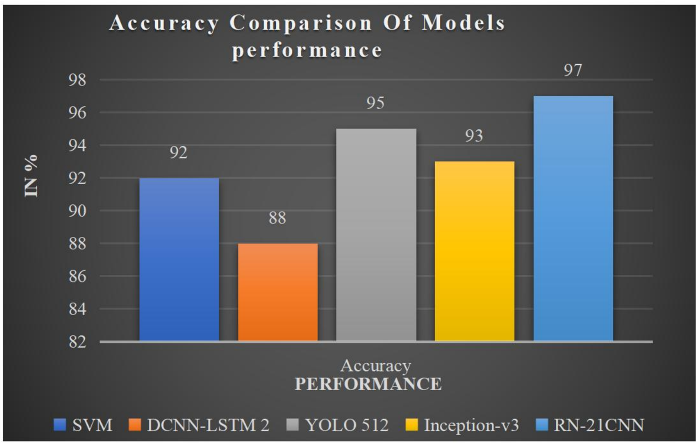

Figure 9: Comparison of Model's Accuracy

In comparing of accuracy of different models shows in figure 9. The RN-21CNN model achieving a remarkable 97% accuracy and YOLO 512 closely with an accuracy 95 accuracy. The DCNN model perform a lowest accuracy.

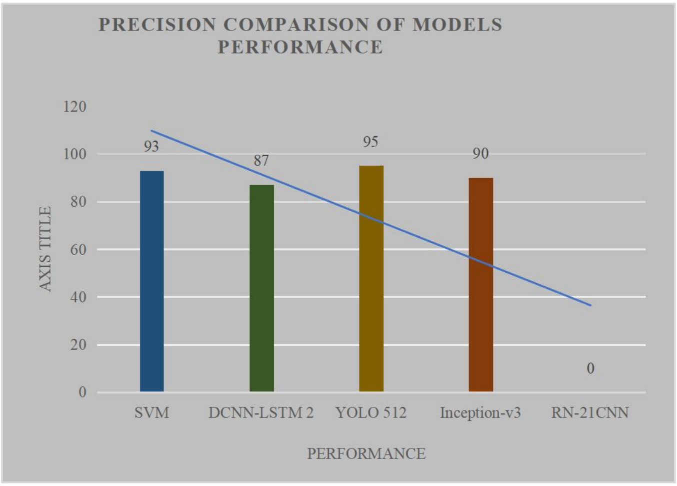

Figure 10: Precision comparison of model's performance

The above figure 10 shows the Precision comparison of model's performance. RN-21CNN achieves a precision 0 and SVM model attain a precision is 93%. The inception-V3 and DCNN-LSTM 2 achieve precision is 90% and 87%, respectively.

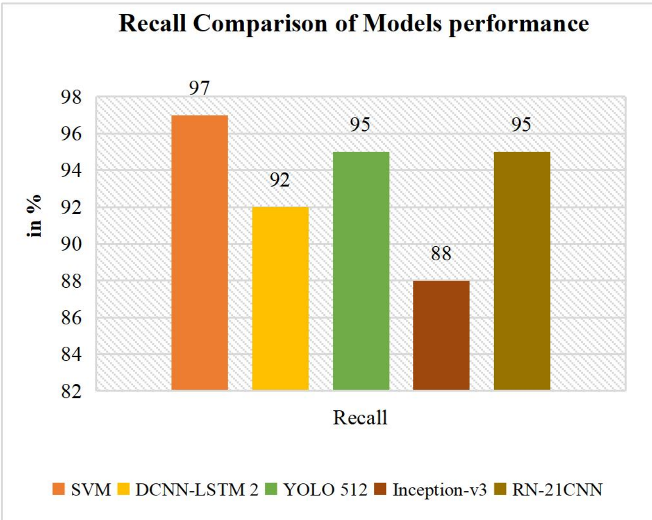

Figure 11: Recall comparison of model's

The above figure 11 shows the Recall comparison of models. SVM stands out with a recall of 97%. The YOLO512 and RN-21CNN achieve the same 95% recall and inception-V3 recall rate is lower.

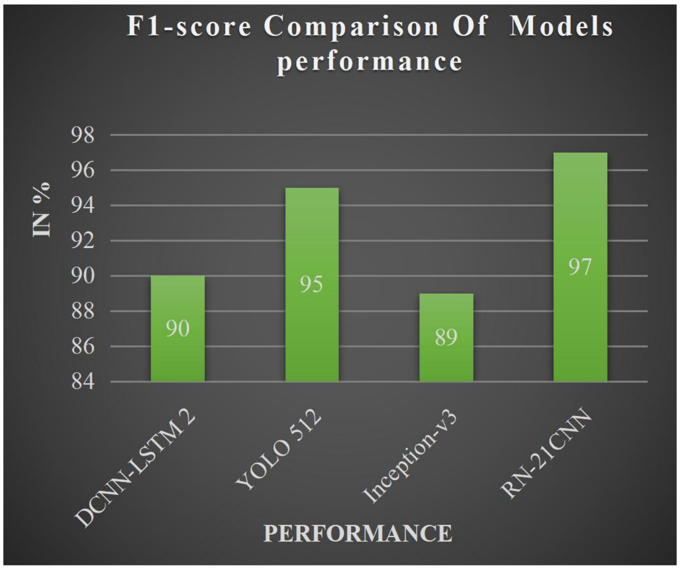

Figure 12: F1-score comparison of model's

The above figure 12 shows the comparison of model's performance on F1-score. RN-21CNN shoes a robust F1-score is 0.97 and YOLO 512 achieve 0.95 f1-score. The Inception-v3 achieves an F1-score of 0.89 respectively.

V. CONCLUSION AND FUTURE WORK

Machine learning plays a vital role in medical imaging when it comes to diagnosing wrist fractures. Medical practitioners may use it to improve the precision and efficiency of fracture diagnosis. Optimal treatment, pain reduction, and a speedier, more complete recovery may all be achieved with early fracture identification. In order to identify wrist fractures from X-ray pictures, this study contrasts machine learning and deep learning techniques. A data set of 111 wrist X-ray pictures with fractures and 82 with normal wrists was taken from Mendeley for the proposed study. Using X-ray images, the RN-21CNN model successfully attained an impressive 97% accuracy and 95% recall. A finding of a research indicated that a suggested model is superior to other models that were compared in the study in terms of its ability to effectively detect bone fractures. The study showcased the effectiveness of the RN-21CNN architecture, which leverages residual networks and convolution neural networks to achieve superior performance in fracture detection tasks. The study needs to overcome a few problems. Initial biases in the training and evaluation data set, such as unequal class distribution or low diversity, may have an impact on a performance and generalizability of a model. To improve an efficacy of bone fracture detection models and solve their limitations, several areas for future study might be investigated.

References

[1] T. Nahid Sultana and A. Parveen, "Efficient Bone Fracture Detection And Classification Using Machine Learning Approaches," vol. 19, no. 2, pp. 2022–1046, 2022, [Online]. Available: http://www.webology.org.

[2] M. Brink *et al.*, "Single-shot CT after wrist trauma: impact on detection accuracy and treatment of fractures," *Skeletal Radiol.*, 2019, doi: 10.1007/s00256-018-3097-z.

[3] D. Joshi and T. P. Singh, "A survey of fracture detection techniques in bone X-ray images," *Artif. Intell. Rev.*, 2020, doi: 10.1007/s10462-019-09799-0.

[4] M. A. A. Basha, A. A. A. Ismail, and A. H. F. Imam, "Does radiography still have a significant diagnostic role in evaluation of acute traumatic wrist injuries? A prospective comparative study," *Emerg. Radiol.*, 2018, doi: 10.1007/s10140-017-1559-6.

[5] M. Kohli, L. M. Prevedello, R. W. Filice, and J. R. Geis, "Implementing machine learning in radiology practice and research," *American Journal of Roentgenology*. 2017, doi: 10.2214/AJR.16.17224.

[6] H. Greenspan, B. Van Ginneken, and R. M. Summers, "Guest Editorial Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique," *IEEE Transactions on Medical Imaging*. 2016, doi: 10.1109/TMI.2016.2553401.

[7] I. M, V. I, A. J, P. P, and R. J, "Deep Learning Model to Detect and Classify Bone Fracture in X-Ray Images," in *2023 International Conference on System, Computation, Automation and Networking (ICSCAN)*, 2023, pp. 1–6, doi: 10.1109/ICSCAN58655.2023.10394986.

[8] N. Vasker, M. Hasan, M. B. R. Nuha, S. Jahan, M. Tahsin, and M. Y. A. Emon, "Real-time Classification of Bone Fractures Utilizing Different Convolutional Neural Network Approaches," in *2023 26th International Conference on Computer and Information Technology (ICCIT)*, 2023, pp. 1–6, doi: 10.1109/ICCIT60459.2023.10441387.

[9] G. Moon, S. Kim, W. Kim, Y. Kim, Y. Jeong, and H. S. Choi, "Computer Aided Facial Bone Fracture Diagnosis (CA-FBFD) System Based on Object Detection Model," *IEEE Access*, 2022, doi: 10.1109/ACCESS.2022.3192389.

[10] D. Joshi and T. P. Singh, "Novel Use of Deep Convolution Architecture Pre-Trained on Surface Crack Dataset to Localize and Segment Wrist Bone Fractures," 2022, doi: 10.1109/SMART55829.2022.10047826.

[11] H. P. Nguyen, T. P. Hoang, and H. H. Nguyen, "A deep learning based fracture detection in arm bone X-ray images," 2021, doi: 10.1109/MAPR53640.2021.9585292.

[12] J. Olczak *et al.*, "Artificial intelligence for analyzing orthopedic trauma radiographs," *Acta Orthop.*, 2017, doi: 10.1080/17453674.2017.1344459.

[13] K. Dimililer, "IBFDS: Intelligent bone fracture detection system," 2017, doi: 10.1016/j.procs.2017.11.237.

[14] R. Ebsim, J. Naqvi, and T. F. Cootes, "Automatic detection of wrist fractures from posteroanterior and lateral radiographs: A deep learning-based approach," 2019, doi: 10.1007/978-3-030-11166-3_10.

[15] Y. L. Thian, Y. Li, P. Jagmohan, D. Sia, V. E. Yao Chan, and R. T. Tan, "Convolutional neural networks for automated fracture detection and localization on wrist radiographs," *Radiol. Artif. Intell.*, 2019, doi: 10.1148/ryai.2019180001.

[16] D. H. Kim and T. MacKinnon, "Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks," *Clin. Radiol.*, 2018, doi: 10.1016/j.crad.2017.11.015.

[17] A. Jiménez-Sánchez *et al.*, "Precise proximal femur fracture classification for interactive training and surgical planning," *Int. J. Comput. Assist. Radiol. Surg.*, 2020, doi: 10.1007/s11548-020-02150-x.

[18] G. Chada, "Machine Learning Models for Abnormality Detection in Musculoskeletal Radiographs," *Reports*, 2019, doi: 10.3390/reports2040026.

[19] C. T. Cheng *et al.*, "Application of a deep learning algorithm for detection and visualization of hip fractures on plain pelvic radiographs," *Eur. Radiol.*, 2019, doi: 10.1007/s00330-019-06167-y.

[20] V. Nair and G. E. Hinton, "Rectified linear units improve Restricted Boltzmann machines," 2010.

[21] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," 2015.

[22] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," *Comput. Electron. Agric.*, 2019, doi: 10.1016/j.compag.2018.03.032.

[23] K. Dlshad Ahmed and R. Hawezi, "Detection of bone fracture based on machine learning techniques," *Meas. Sensors*, 2023, doi: 10.1016/j.measen.2023.100723.

[24] 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," *Life*, 2023, doi: 10.3390/life13010133.

[25] F. Hržić, S. Tschauner, E. Sorantin, and I. Štajduhar, "Fracture Recognition in Paediatric Wrist Radiographs: An Object Detection Approach," *Mathematics*, 2022, doi: 10.3390/math10162939. \

[26] V.A.Shah, H.P.Chauhan, H.R.Patel, "Enhancing Diagnostic Accuracy in Wrist X-ray Analysis Using Deep Transfer Learning Based Hybrid Model",, Journal of Technology, Page No:-959-968 Vol:12,Issue:10,October-2024. DOI:18.15001/JOT.2024/V12I10.24.850

[27] V.A.Shah, H.P.Chauhan, H.R.Patel, "Image Processing for bone Fracture Detection System", Journal of Technology, Page No:-998-1005 Vol:12,Issue:10,October-2024. DOI:18.15001/JOT.2024/V12I10.24.854