Chapter Title - A COMPARATIVE ANALYSIS OF DEEP LEARNING MODELS AND CONVENTIONAL APPROACHES FOR OSTEOPOROSIS DETECTION IN HIP X-RAY IMAGES

Subtopic - Machine Learning and Deep Learning Applications

1) Introduction

Osteoporosis is a common bone disorder characterized by low bone mass and structural deterioration of bone tissue, leading to increased bone fragility and a higher risk of fractures. It primarily affects older individuals, particularly women after menopause, but can also occur in men and younger individuals.

Prevention and management of osteoporosis involve a multifaceted approach. Adequate intake of calcium and vitamin D, regular weight-bearing and muscle-strengthening exercises, avoidance of tobacco and excessive alcohol, and fall prevention strategies are essential for maintaining bone health. In some cases, medications may be prescribed to slow bone loss or promote bone formation.

Early detection of osteoporosis is crucial to initiate appropriate interventions. Dual-energy X-ray absorptiometry (DXA or DEXA) is the most commonly used test to measure bone mineral density and assess fracture risk. Screening guidelines vary, but generally, women over 65 and men over 70 should undergo DXA testing.

2) Dataset employed

Data cleaning is an essential step in the data preparation process, ensuring that datasets are accurate, consistent, and reliable for analysis. It involves identifying and rectifying errors, inconsistencies, and inaccuracies in data, ranging from missing values and duplicate entries to formatting issues and outliers. The process often starts with data profiling and exploratory analysis to gain insights into the quality and structure of the data. Then, various techniques are applied to handle different types of issues, such as imputation methods for missing values, deduplication algorithms for removing duplicates, and data transformation approaches for standardizing formats. Validation checks and quality assurance measures are also performed to ensure that the cleaned data meets the predefined criteria or constraints. Through effective data cleaning, organizations can enhance data integrity, reduce biases, and improve the accuracy and reliability of subsequent data analysis and modelling tasks.

The data used in the model was collected from various sources like Kaggle, Roboflow, etc. The data consisted of 49 hip X-Ray images of males and females. Later the data was cleaned and annotated with the help of Roboflow under the supervision of Orthopedic surgeons from Lata Mangeshkar Hospital, Nagpur. Next, the data was analyzed. The following properties were kept as default for all images:

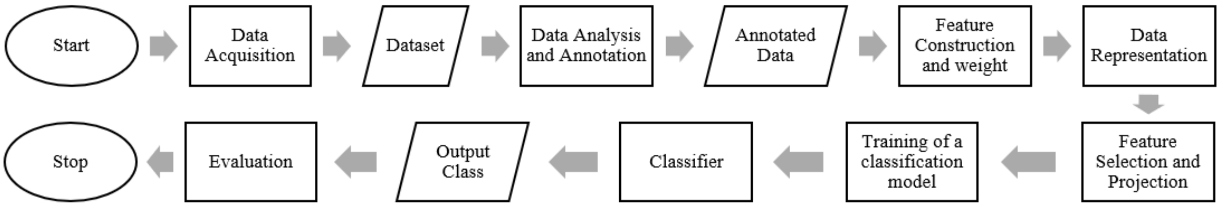
1) Auto-Orient: Applied and
2) Resize: Stretch to 640X640.

Finally, the data was applied with the following list of augmentations after which the data increased to 117 images:

1) Flip: Horizontal,
2) Minimum Zoom: 40%
3) Rotation: Between -15° and +15°

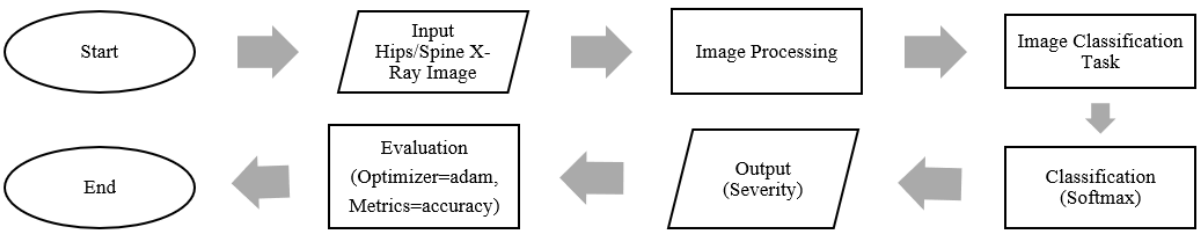
3) Flow Chart & Block Diagram

Flowchart



Flow chart for the working of Osteoporosis Detection using Deep Learning Models

Block Diagram



Block diagram depicting the workflow of the program for Osteoporosis Detection using Deep Learning Model

5) Results

ResNet's "skip connections" feature offers a novel approach to the vanishing gradient problem. Initial inactive convolutional layers (many identity mappings; ResNet) are stacked, skipped, and activated once more from the previous layer. By reducing the number of layers in the network, skipping accelerates early training. The hip X-Ray dataset was observed to have an accuracy of 80% when trained on the ResNet model. the confusion matrix obtained during the evaluation of ResNet-50 on Hips X-Ray Dataset and Table I shows the accuracy, precision, recall and F1-score values obtained on the testing dataset.

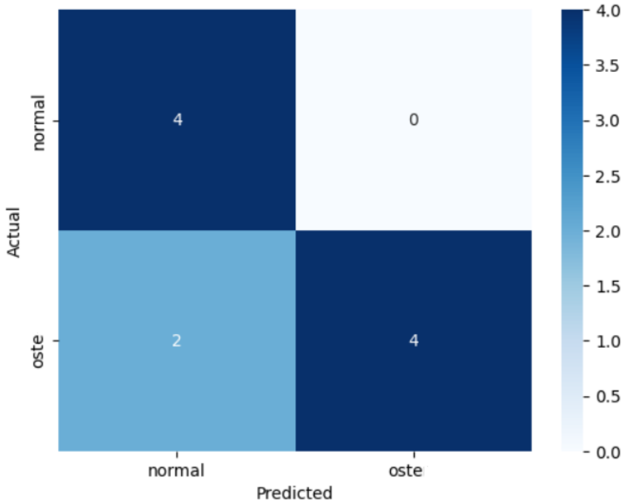


Fig. 1. Confusion matrix for ResNet model on X-Ray data of hips consisting of 10 images belonging to two classes: a) Normal, b) Osteoporosis.

Classes	**Evaluation Metrics**			
	Accuracy	***Precision***	***Recall***	***F1-Score***
All Classes	0.80	1.00	0.67	0.80
Normal	0.67	0.67	1.00	0.80
Osteoporotic	0.80	1.00	0.67	0.85

Table 1. Precision, recall, and mean average precision values of ResNet-50 on hip X-Ray dataset

4.1.2. Inception Net

Large networks are more susceptible to overfitting, and when multiple convolutional processes are connected together, the network's computing cost increases. In order to realize the Inception network and module, researchers constructed straightforward Conv net designs. An inception network is a type of deep neural network with repeating architectural structures called Inception modules. The Inception module consists of a 1x1 convolution layer for feature reduction, 3x3 and 5x5 convolutional layers for learning different spatial features at different scales, and a max pooling layer.

85.2% accuracy was attained once CNN's Inception Net architecture was implemented for the prediction of osteoporosis through the Hips X-Ray dataset. The results obtained on Inception Net model are depicted in the form of a confusion matrix and the corresponding accuracy, precision, recall and F1-score values are depicted in Table II.

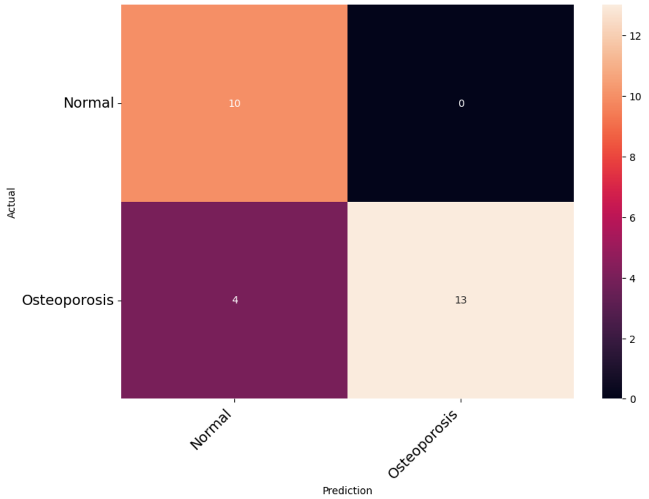


Fig. 2. Confusion matrix for X-Ray data of hips consisting of 27 images belonging to two classes: a) Normal, and b) Osteoporosis, trained on the InceptionNet model.

Classes	**Evaluation Metrics**			
	Accuracy	***Precision***	***Recall***	***F1-Score***
All Classes	0.851	1.000	0.714	0.833
Normal	0.851	0.714	1.000	0.851

Classes	Evaluation Metrics			
	Accuracy	***Precision***	***Recall***	***F1-Score***
Osteoporotic	0.851	1.000	0.714	0.750

TABLE 2. PRECISION, RECALL, AND MEAN AVERAGE PRECISION VALUES OF INCEPTION NET ON HIP X-RAY DATASET

4.1.3. YOLO v7

YOLO (You Only Look Once) first divides the image into N grids before doing object detection in a single step. These grids are of the same 5x5 size. It is utilized to find and locate any objects that might be present in any of these areas. Bounding box coordinates, B, for any prospective objects are predicted for each grid along with their object labels and a likelihood rating for their presence.

YOLOv7 was the best object detection algorithm available after its release in 2022 by Chien-Yao Wang et al. The accuracy obtained on the YOLOv7 model for this particular application was 90.9% for mAP@0.5 (mean average precision) as evident in Table III.

The various losses as well as the evaluation metrics are plotted for training and validation sets.

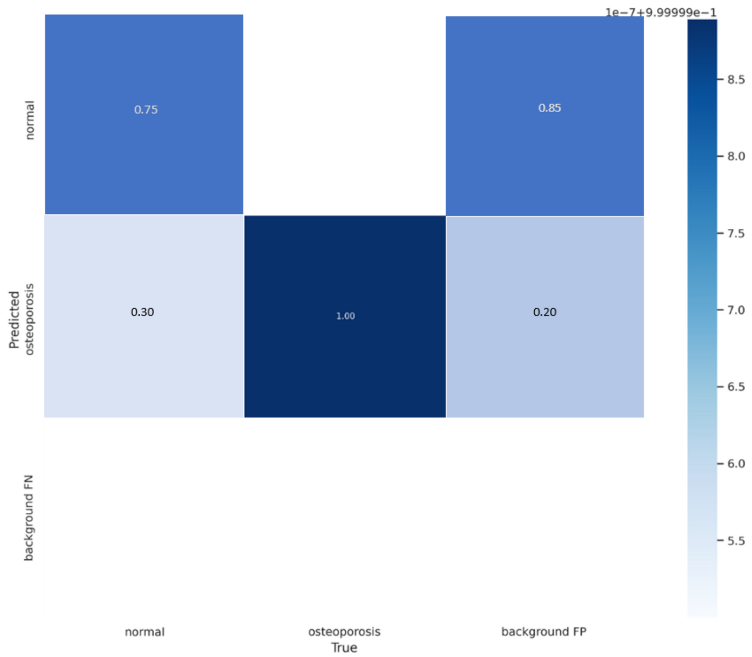


Fig. 3. Confusion matrix for X-Ray data of hips two classes: a) Normal, and b) Osteoporosis, along with an additional background class trained on the YOLO v7 model.

Classes	Evaluation Metrics			
	Precision	***Recall***	***mAP @0.5***	***mAP @0.5:0.95***
All	0.820	0.788	0.909	0.822
Normal	0.939	0.700	0.908	0.849
Osteoporotic	0.700	0.875	0.910	0.795

TABLE 3. PRECISION, RECALL, AND MEAN AVERAGE PRECISION VALUES OF YOLOV7 ON HIP X-RAY DATASET

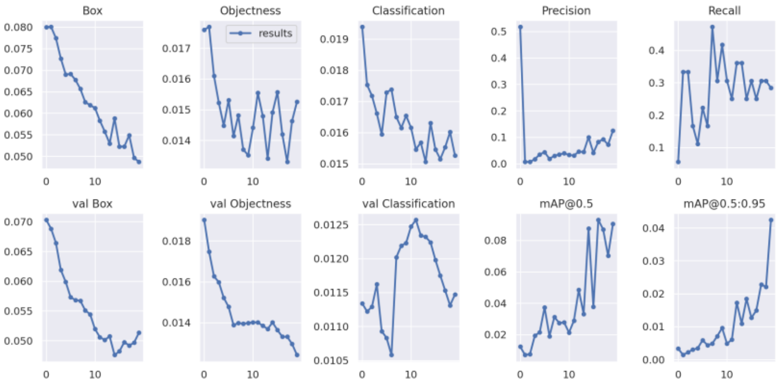


Fig. 4. Graphs for box loss, objectness loss, classification loss, precision, recall and mean average precision (mAP@0.5 & mAP@0.95) over training and validation sets on YOLO v7

4.1.4. Ultralytics YOLO v8

Ultralytics YOLOv8 is the most recent version of the well-known real-time object identification and image segmentation model. Due to cutting-edge developments in computer vision and deep learning, YOLOv8 offers unparalleled speed and accuracy. Because of its straightforward engineering, it is great for a large number of utilizations and is effectively versatile to numerous equipment stages, including edge gadgets and cloud APIs.

The precision obtained after training the YOLOv8 model on the hips X-ray dataset was 98.8% mAP@0.5, the highest accuracy obtained during the research. The results for evaluation metrics used in this study are shown in Table IV and the corresponding graphs for losses and evaluation metrics are plotted .

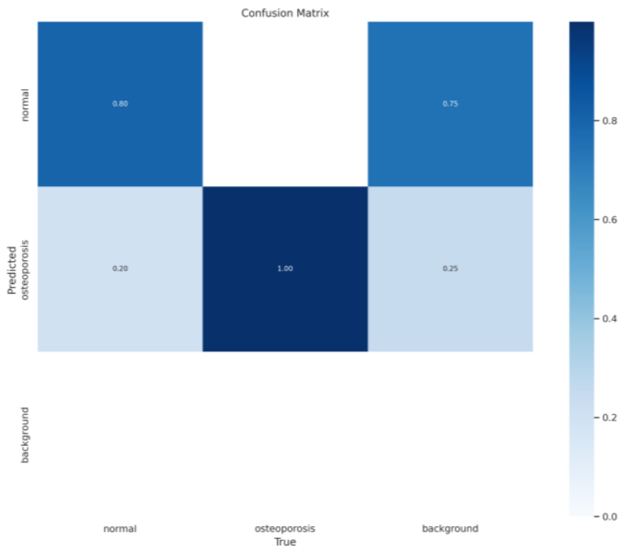


Fig. 5. Confusion matrix for X-Ray data of hips two classes: a) Normal, and b) Osteoporosis, along with an additional background class trained on the YOLO v8 model.

Classes	**Evaluation Metrics**			
	Precision	***Recall***	***mAP @0.5***	***mAP @0.5:0.95***
All	0.929	0.904	0.988	0.888
Normal	1.000	0.807	0.995	0.889
Osteoporotic	0.858	1.000	0.982	0.887

TABLE 4. PRECISION, RECALL, AND MEAN AVERAGE PRECISION VALUES OF YOLOv8 ULTRALYTICS ON HIP X-RAY DATASET

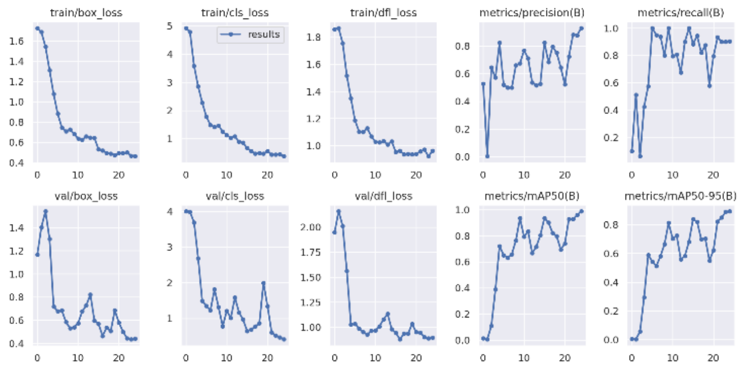


Fig. 6. Graphs for box loss, objectness loss, classification loss, precision, recall and mean average precision (mAP@0.5 & mAP@0.95) over training and validation sets on YOLO v8

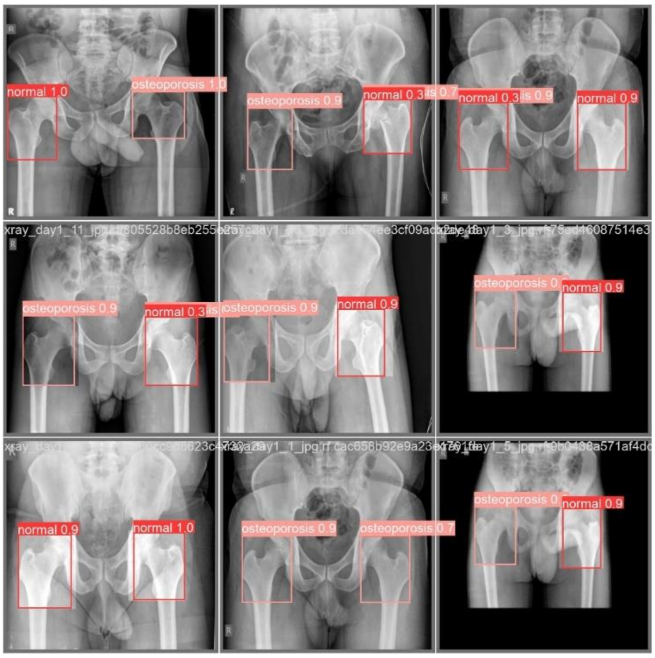


Fig. 7. Ultralytics YOLOv8.0.20 prediction results on the validation set of hip X-Ray dataset labeling the confidence on the bone being predicted as normal or osteoporotic.

4) Conclusion

The bone disorder called osteoporosis is characterized by decreased bone mineral density (BMD) and an increased risk of fracture. The Dual-Energy X-Ray Absorptiometry (DXA) scan, one of the most recent techniques for diagnosing and treating osteoporosis, has become the gold standard for reliable osteoporosis detection. The sole disadvantage of this procedure is that it is pricey and out of reach for individuals in lower socioeconomic brackets. The authors, therefore, suggest using deep learning techniques on X-ray pictures in place of DEXA by conducting this study. On an X-Ray dataset of hips consisting of 117 images after augmentation, the paper summarises and compares four deep learning algorithms used for the same task: ResNet-50, Inception Net, YOLOv7, and Ultralytics YOLO v8 models.

Among the four algorithms evaluated, the YOLO v8 model demonstrated the highest accuracy in the conducted study. This superior performance can be attributed to several key factors inherent to the YOLO v8 architecture. Firstly, the YOLO (You Only Look Once) model family is known for its real-time object detection capabilities. YOLO v8, in particular, leverages a single unified neural network architecture that predicts bounding boxes and class probabilities directly on a dense grid, thereby enabling faster inference speed compared to other object detection models. Additionally, YOLO v8 incorporates numerous architectural enhancements and advancements over its predecessors, such as the inclusion of feature pyramid networks, anchor-based predictions, and feature fusion techniques. These enhancements facilitate more accurate localization and classification of objects within images. Considering these factors, the YOLO v8 model's superior accuracy can be attributed to its real-time object detection capabilities, architectural enhancements, robustness against scale variations and occlusions, and the availability of comprehensive training data. These advantages collectively contribute to its remarkable performance and make it a preferred choice for accurate object detection tasks.