*A project report on*

**AUTOMATED DETECTION OF OSTEOPOROSIS AND OSTEOPENIA IN ELDERLY PATIENTS USING KNEE X-RAY IMAGE ANALYSIS**

*Submitted in partial fulfillment for the award of the degree of*

## Bachelor of Technology in Computer

## Science and Engineering

*by*

**SARGAM CHAUDHARY (21BCE5624)**

**GAURAV BHAUMIK(21BCE1495)**

**SHASHANK AGARWAL(21BCE1586)**



**SCHOOL OF COMPUTER SCIENCE AND ENGINEERING**



**DECLARATION**

I hereby declare that the thesis entitled “**AUTOMATED DETECTION OF OSTEOPOROSIS AND OSTEOPENIA IN ELDERLY PATIENTS USING KNEE X-RAY IMAGE ANALYSIS**” submitted by GAURAV BHAUMIK(21BCE1495) for the award of the degree of Bachelor of Technology in Computer Science and Engineering, Vellore Institute of Technology, Chennai is a record of bonafide work carried out by me under the supervision of Dr. Deepa Nivethika.

I further declare that the work reported in this thesis has not been submitted and will not be submitted, either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university.

Place: Chennai

Date: **Signature of the Candidate**



**School of Computer Science and Engineering**

CERTIFICATE

This is to certify that the report entitled **“AUTOMATED DETECTION OF OSTEOPOROSIS AND OSTEOPENIA IN ELDERLY PATIENTS USING KNEE X-RAY IMAGE ANALYSIS”** is prepared and submitted by **Gaurav Bhaumik(21BCE1495)** to Vellore Institute of Technology, Chennai, in partial fulfillment of the requirement for the award of the degree of **Bachelor of Technology in Computer Science and Engineering** is a bonafide record carried out under my guidance. The project fulfills the requirements as per the regulations of this University and in my opinion meets the necessary standards for submission. The contents of this report have not been submitted and will not be submitted either in part or in full, for the award of any other degree or diploma and the same is certified.

Signature of the Guide:

Name: Dr./Prof.

Date:

Signature of the Examiner Signature of the Examiner

Name: Name:

Date: Date:

Approved by the Head of Department,

**B.Tech. CSE**

Name: Dr. Nithyanandam P

Date:

(Seal of SCOPE)

**ABSTRACT**

Osteoporosis, often asymptomatic until fractures occur, poses significant challenges for early diagnosis. This study introduces a deep learning framework using knee X-ray images to automate osteoporosis detection by integrating CNNs with transfer learning and data augmentation. A dataset of 1,947 images, categorized as Normal, Osteopenia, and Osteoporosis, was pre-processed via Min-Max scaling, resizing to 224×224 pixels, and augmentation techniques. We evaluated a custom CNN alongside pre-trained models (VGG-16, VGG-19, ResNet-50, DenseNet-121, XceptionNet) achieving accuracies between 82.1% and 94.5%. Two hybrid models were then developed: one combining VGG-19, InceptionResNetV2, and MobileNetV2 reached 97.5% accuracy, and another integrating DenseNet with EfficientNet reached 94.8%. Grad-CAM visualization provided interpretable insights, enhancing clinical relevance. The proposed approach offers a scalable solution that can be easily integrated into existing radiology workflows, and our experimental results underscore the importance of ensemble strategies in boosting model performance. Moreover, the integration of Grad-CAM visualizations adds a critical layer of interpretability that can increase clinician trust and adoption. This research sets the stage for further exploration of advanced hybrid architectures in medical imaging diagnostics, demonstrating the framework’s potential as a decision-support tool for early osteoporosis diagnosis, with future work aimed at dataset expansion and real-time implementation.

**ACKNOWLEDGEMENT**

It is my pleasure to express with deep sense of gratitude to Dr. Deepa Nivethika,

School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, for her constant guidance, continual encouragement, understanding; more than all, she taught me patience in my endeavor. My association with her is not confined to academics only, but it is a great opportunity on my part to work with an intellectual and expert in the field of Machine Learning.

It is with gratitude that I would like to extend my thanks to the visionary leader Dr. G. Viswanathan our Honorable Chancellor, Mr. Sankar Viswanathan, Dr. Sekar Viswanathan, Dr. G V Selvam Vice Presidents, Dr. Sandhya Pentareddy, Executive Director, Ms. Kadhambari S. Viswanathan, Assistant Vice-President, Dr. V. S. Kanchana Bhaaskaran Vice-Chancellor, Dr. T. Thyagarajan Pro-Vice Chancellor, VIT Chennai and Dr. P. K. Manoharan, Additional Registrar for providing an exceptional working environment and inspiring all of us during the tenure of the course.

Special mention to Dr. Ganesan R, Dean, Dr. Parvathi R, Associate Dean Academics, Dr. Geetha S, Associate Dean Research, School of Computer Science and Engineering, Vellore Institute of Technology, Chennai for spending their valuable time and efforts in sharing their knowledge and for helping us in every aspect.

In jubilant state, I express ingeniously my whole-hearted thanks to Dr. Nithyanandam P , Head of the Department, B.Tech. CSE and the Project Coordinators for their valuable support and encouragement to take up and complete the thesis.

My sincere thanks to all the faculties and staff at Vellore Institute of Technology, Chennai who helped me acquire the requisite knowledge. I would like to thank my parents for their support. It is indeed a pleasure to thank my friends who encouraged me to take up and complete this task.

Place: Chennai

Date: **GAURAV BHAUMIK**

## INDEX

1. INTRODUCTION (CHAPTER 1) 1

* 1. DOMAIN OF THE PROBLEM 1
  2. MOTIVATION 2
  3. PROBLEM STATEMENT 3
  4. OBJECTIVES 4

2. LITERATURE SURVEY (CHAPTER 2) 5

2.1 BACKGROUND RESEARCH 5

2.2 GAP ANALYSIS 11

2.3 CHALLENGES 13

3. SYSTEM DESIGN (CHAPTER 3) 15

3.1 PROBLEM DEFINITION 15

3.2 TRANSFER LEARNING IN DEEP LEARNING MODELS 15

3.3 OSTEOPOROSIS DIAGNOSIS METHODOLOGIES 16

3.4 CONVOLUTION NEURAL NETWORKS 17

3.5 DATASET 18

3.6 METHODOLOGY 20

4. IMPLEMENTATION (CHAPTER 4) 25

4.1 FUNCTION REQUIREMENTS 25

4.2 NON FUNCTIONAL REQUIREMENTS 25

4.3 COMPARATIVE ANALYSIS 26

5. RESULTS (CHAPTER 5) 32

5.1 PROPOSED HYBRID SYSTEM 32

5.2 USE OF GRAD-CAM 37

6. CONCLUSION AND FUTURE WORK (CHAPTER 6) 38

7. APPENDICES 40

8. REFERENCES 51

**LIST OF FIGURES**

1. Example of dataset 19
2. Dataset collection 19
3. Model accuracy for Custom CNN 26
4. Model loss for Custom CNN 26
5. Model accuracy scores for Custom CNN 27
6. Confusion Matrix for Custom CNN 27
7. Model accuracy for VGG19 27
8. Model loss for VGG19 27
9. Model accuracy scores for VGG19 28
10. Confusion Matrix for VGG19 28
11. Model accuracy for ResNet-50 28
12. Model loss for ResNet-50 28
13. Model accuracy scores for ResNet-50 29
14. Confusion Matrix for ResNet-50 29
15. Model accuracy for DenseNet-121 29
16. Model loss for DenseNet-121 29
17. Model accuracy scores for DenseNet-121 30
18. Confusion Matrix for DenseNet-121 30
19. Model accuracy for XceptionNet 30
20. Model loss for XceptionNet 30
21. Model accuracy scores for XceptionNet 31
22. Confusion Matrix for XceptionNet 31
23. Architecture diagram for Hybrid Model 1 33
24. Model accuracy for Hybrid Model 1 33
25. Model loss for Hybrid Model 1 33
26. Model accuracy scores for Hybrid Model 1 34
27. Confusion Matrix for Hybrid Model 1 34
28. Architecture diagram for Hybrid Model 2 35
29. Model loss for Hybrid Model 2 36
30. Model accuracy for Hybrid Model 2 36
31. Model accuracy scores for Hybrid Model 2 36
32. Confusion Matrix for Hybrid Model 2 36

**LIST OF ACRONYMS**

1. MTCNN - Multi Task Cascaded Convolution Network
2. RNN - Recurrent Neural Network
3. CV - Computer Vision
4. CNN - Convolutional Neural Network
5. GPU - Graphics Processing Unit
6. DL - Deep Learning
7. FaceNet - Face Network
8. DXA - Dual-energy X-ray absorptiometry
9. BMD – Bone Mineral Density
10. MRI – Magnetic Resonance Imaging
11. MRS – Magnetic Resonance Spectroscopy
12. DEXA - Dual-Energy X-ray Absorptiometry
13. QCT - Quantitative Computed Tomography
14. QUS - Quantitative Ultrasound
15. ReLU - Rectified Linear Unit
16. ROC - Receiver operating characteristic

**Chapter 1**

**Introduction**

* 1. DOMAIN OF THE PROBLEM

Osteoporosis is a disease that often goes undiagnosed until fractures occur, leading to pain, disability, and reduced independence in elderly patients. Early detection is crucial for effective management and fracture risk reduction, yet current diagnostic methods are time-consuming. This project presents a novel approach using transfer learning with convolutional neural networks (CNNs) for osteoporosis detection from X-ray images. The approach achieves high diagnostic accuracy and provides a revealed feature map to assist medical professionals. The innovation involves a dual strategy: integrating transfer learning from CNN architectures and employing a dataset collection augmentation mechanism. The study uses 1947 knee X-rays for training and testing, highlighting the potential of deep learning with transfer learning in aiding early osteoporosis detection and mitigating fracture risks.

The classification of osteoporosis in the knee has been the subject of numerous studies, each with its own distinct findings and methodological subtleties. For example, a 2020 study by Wani and Arora reported an impressive 91% accuracy rate; however, a closer examination of the confusion matrix revealed that the actual accuracy was lower, and the models exhibited high validation loss, suggesting potential limitations in generalization. In contrast, Kumar et al.’s study in 2023 showcased higher accuracy compared to Wani and Arora’s findings. Similarly, research by Abubakar et al. in 2022 demonstrated encouraging accuracy in binary classification, yet the study lacked reporting on validation loss and encountered challenges in correctly classifying some instances. These issues highlight the need for improved model robustness. Together, these studies underscore several significant gaps in the current state of the field, including limited dataset size, poor image quality, and the necessity for a more complete and reliable CNN-based architecture for osteoporosis identification from X-ray images. Our research efforts have shown encouraging outcomes, particularly in enhancing accuracy in both binary and multiclass classification problems.

Brief overview of the project:

1. **Model Architecture:** We investigate various deep learning models utilizing a transfer learning approach to ensure detection accuracy and efficiency. Models under evaluation include VGG-19, a custom CNN, ResNet-50, DenseNet-121, and XceptionNet, all aimed at achieving accurate knee osteoporosis detection. We also aim to combine two or more models to further improve accuracy.
2. **Detection Objectives:** The detection strategy is two-fold: a binary classification for osteoporosis detection and a multiclass classification for distinguishing between osteoporosis and osteopenia. This dual approach ensures that our work addresses both diagnostic directions.
3. **Dataset Use:** The model is trained and evaluated on an assembled dataset, which is a combination of four independent datasets. To address issues of class imbalance, data augmentation techniques are applied after combining the datasets.

This arranged introduction sets the stage for a detailed discussion on the implementation of our deep learning framework, its evaluation, and its potential impact on early osteoporosis detection.

* 1. MOTIVATION

Accurate measurements of bone density in the knee are crucial for reducing serious consequences. Dual-energy X-ray absorptiometry (DXA) examinations or whole leg models are ineffective for estimating knee mineral density, as bone mineral density (BMD) over the entire knee area is not of direct clinical significance. In clinical practice, the diagnosis of knee osteoarthritis is primarily based on X-ray radiographs, yet there is a growing demand for a rapid, non-invasive, and accurate diagnostic tool for this condition. Osteoporosis, a bone disease characterized by decreased bone density and strength, results in fragile, porous, and brittle bones that are more prone to fractures. This condition, common in older adults—especially postmenopausal women—is influenced by factors such as aging, hormonal changes, calcium deficiency, and lack of physical activity, and it increases the risk of fractures in critical areas like the hip, spine, and wrist. Additionally, complications such as knee degeneration, delayed treatment, prolonged post-fracture hospitalization, and secondary major depressive disorder following femoral neck fractures are frequently observed, particularly in elderly women. This study aims to develop a rapid diagnostic tool for early detection of knee osteoarthritis and osteoporosis, thereby supporting clinical decision-making and potentially reducing the serious consequences associated with these conditions.

1.3 PROBLEM STATEMENT

Accurate assessment of knee bone density is critical for the early detection and effective management of both osteoporosis and knee osteoarthritis. Current diagnostic methods, including dual-energy X-ray absorptiometry (DXA) and whole leg models, are inadequate because they evaluate bone mineral density (BMD) over the entire knee area, which does not provide clinically significant insights specific to the knee joint. In clinical practice, diagnosis largely relies on conventional X-ray radiographs; however, their interpretation can be subjective and time-consuming. This leads to delays in the diagnosis and treatment of degenerative bone conditions, especially in populations most at risk, such as elderly patients and postmenopausal women.

These delays and inaccuracies have significant implications. Osteoporosis, characterized by decreased bone density and strength, results in bones that are fragile, porous, and highly susceptible to fractures. The condition is compounded by factors such as aging, hormonal imbalances, calcium deficiency, and a sedentary lifestyle, thereby increasing the risk of fractures in critical areas like the hip, spine, and wrist. In addition to the immediate physical risks, complications from untreated or late-diagnosed osteoporosis, including knee degeneration, prolonged hospitalization, and even major depressive disorder following fractures (such as femoral neck fractures), further diminish quality of life and escalate healthcare costs. Consequently, there is an urgent need for a rapid, non-invasive, and accurate diagnostic tool specifically tailored to assess knee bone density. Such a tool would not only facilitate early intervention and improve clinical decision-making but also help mitigate the severe consequences associated with these musculoskeletal conditions.

1.4 OBJECTIVES

1. **Develop a Robust Diagnostic Model:** Design and implement a deep learning framework using convolutional neural networks (CNNs) and transfer learning to accurately classify knee X-ray images into normal, osteopenia, and osteoporosis categories.
2. **Enable Early Detection with Increased Accuracy:** Facilitate early diagnosis of osteoporosis to minimize fracture risks and improve patient outcomes by achieving high classification accuracy and reducing diagnostic delays.
3. **Address Data Challenges:** Overcome limitations posed by small and imbalanced datasets through the use of advanced data augmentation techniques and the integration of multiple independent datasets, thereby enhancing model reliability and generalizability.
4. **Advance Medical Imaging:** Enhance the application of artificial intelligence in healthcare by developing an efficient, scalable, and clinically viable diagnostic tool for knee osteoporosis detection that supports rapid and non-invasive assessments.

**Chapter 2**

**Literature Survey**

2.1 BACKGROUND RESEARCH

**[1] KONet: Toward a Weighted Ensemble Learning Model for Knee Osteoporosis Classification - M. J. A. Rasool, S. Ahmad, U. Sabina and T. K. Whangbo**

This study presents a robust detection method called KONet, designed to address the challenges in identifying knee osteoporosis, a skeletal disorder characterized by bone tissue degradation and low bone density, which increases the risk of knee fractures. Traditional diagnostic methods like knee radiography often require specialist expertise, and the large volume of X-rays, coupled with subtle variations, can result in misinterpretation. Leveraging recent advancements in deep learning, particularly convolutional neural networks (CNNs), KONet employs a weighted ensemble approach to effectively distinguish between normal and osteoporotic knee conditions, even with minor data variations. Experimental results demonstrate that the proposed model achieves superior accuracy compared to existing methods, significantly outperforming state-of-the-art CNN-based models using transfer learning.

**[2] A new superfluity deep learning model for detecting knee osteoporosis and osteopenia in X-ray images - Naguib, S.M., Saleh, M.K., Hamza, H.M. et al**

This study proposes a new deep-learning approach using a superfluity mechanism to categorize knee X-ray images into osteoporosis, osteopenia, and normal classes. The superfluity mechanism involves concatenating multiple layers, allowing features to flow into two branches. The proposed Superfluity DL model demonstrated the highest accuracy (85.42% for dataset1 and 79.39% for dataset2). Osteoporosis is a major public health problem, causing a reduction in bone mass and bone architecture disruption, increasing the risk of fracture and compromised bone strength. Early detection of osteoporosis/osteopenia using digital X-ray images would be useful and cost-effective in preventing fractures and other bone complexities. This study contributes to the development of a reliable, automated tool for initial assessment and reducing false positives and false negatives, enhancing the precision with which knee osteoporosis and osteopenia can be detected, allowing healthcare providers to prioritize patient care more effectively.

**[3] Detection of osteoporosis using image processing methods - Bo, Y., Chen, G., Li, L. et al.**

This paper highlights osteoporosis as a silent, painless illness that increases bone fragility and fracture susceptibility across various areas, emphasizing the importance of early detection for prevention and treatment. While primary osteoporosis is more common in postmenopausal women and elderly men, the disease can affect individuals of any gender or age. Characterized by decreased bone mass, structural weakness, and increased fracture risk, its assessment often relies on bone density measurements, a key indicator for evaluating osteoporosis risk, though not equivalent to the physical density of bone. Diagnostic tools such as bone densitometry, MRI, and MRS provide insights into bone structure and marrow composition but are often costly and inaccessible. To address these challenges, computer-assisted diagnosis (CAD) methods—including image processing, intelligent neural network algorithms, and genetic algorithms—offer solutions for improving image analysis, speeding up diagnoses, and reducing costs and medical errors. These advancements aim to enhance diagnostic accuracy and efficiency while mitigating the limitations of traditional methods.

**[4] Comparative analysis of vision transformers and convolutional neural networks in osteoporosis detection from X-ray images - Sarmadi, A., Razavi, Z.S., van Wijnen, A.J. et al**

This paper investigates the potential of the Vision Transformer (ViT) in medical image analysis, focusing on diagnosing osteoporosis using X-ray radio-images. Compared to traditional convolutional neural networks (CNNs), ViT demonstrates superior performance, highlighting its effectiveness when sufficient training data is available. The study underscores that both ViT and CNNs can offer robust solutions to critical medical diagnostic challenges. Deep learning, a subset of machine learning, leverages artificial neural networks (ANNs) to extract knowledge from extensive datasets and perform complex tasks. Its applications span diverse domains, including computer vision, natural language processing, speech recognition, recommender systems, autonomous vehicles, image generation, machine translation, game playing, and natural language understanding. Beyond these, deep learning is reshaping art, culture, and society by fostering innovation and creating new paradigms. In medical sciences, deep learning holds transformative potential, capable of handling large-scale and high-dimensional data while learning abstract and sophisticated functions, often with reduced dependence on extensive training datasets or predefined assumptions.

**[5] Application of a Neural Network Classifier to Radiofrequency-Based Osteopenia/Osteoporosis Screening - J. Adams, Z. Zhang, G. M. Noetscher, A. Nazarian and S. N. Makarov**

This study introduces a neural network classifier to detect osteoporotic/osteopenic conditions using non-ionizing, very-low-power radiofrequency (RF) signals transmitted through the wrist. Data was collected from 67 participants divided into two groups: Group 1 included 27 osteoporotic/osteopenic individuals (DXA T-score < -1), and Group 2 included 40 healthy individuals without major fracture risk factors. RF spectra (30 kHz–2 GHz) were measured using dual antiphase patch antennas with controlled pressure. Input features included wrist circumference and the normalized transmission coefficient (S21N). A Multi-Layer Perceptron (MLP) neural network was implemented using MATLAB, and leave-one-out cross-validation was performed. Training utilized complex spectra, which combined magnitude and phase for 201 frequency points per wrist. Among tested methods, independently processing left and right wrist spectra, then combining results manually, yielded the best performance. The model achieved a sensitivity of 83% and specificity of 94%. This approach demonstrates potential for quick, cost-effective osteoporosis screening using small, portable devices.

**[6] Using machine learning techniques to predict the risk of osteoporosis based on nationwide chronic disease data - Tu, JB., Liao, WJ., Liu, WC. et al**

This study developed a machine learning (ML) model to predict osteoporosis risk using chronic disease data from 10,000 patient records in the German Disease Analyzer database. Ten chronic conditions, including hypertension, diabetes, COPD, and cancer, were analyzed alongside demographic features like age and gender. Data was split into training (70%) and test (30%) sets using stratified random sampling, and class imbalance was addressed with oversampling. Recursive Feature Elimination (RFE) was used across nine ML algorithms to identify relevant predictors, retaining all eleven features for model development. A stacked ensemble model (Stacker) combining Logistic Regression, AdaBoost, and Gradient Boosting Classifier achieved the best performance. The Stacker model demonstrated superior accuracy, AUC-ROC (0.76), sensitivity (0.722), and specificity (0.664) compared to individual algorithms. SHAP analysis revealed the most influential features: age, gender, lipid metabolism disorders, cancer, and COPD. The model's reliability was validated using cross-validation, confusion matrix, calibration curves, and lift curves. The study underscores the potential of ML in early osteoporosis detection, enabling personalized management strategies. The findings provide a robust framework for leveraging chronic disease data to improve public health outcomes.

**[7] Osteoporosis Prediction Using Machine‑Learned Optical Bone Densitometry Data - Miura, K., Tanaka, S.M., Chotipanich, C. et al**

The study investigates the classification of osteoporosis using a machine learning (ML) model based on bone mineral density (BMD) measurements obtained from dual-energy X-ray absorptiometry (DXA) scans. A total of 203 participants, including 182 patients and 21 hospital staff, underwent BMD assessments of the lumbar spine, femur, and forearm. The optical measurements were taken using a system comprising two convex lenses, annular slits, a photodetector, and a laser diode operating at 850 nm wavelength. The T-scores were calculated to classify participants into osteoporosis (T-score ≤ -2.5) and non-osteoporosis groups. The ML model's parameters were optimized through grid-search with fivefold cross-validation, aiming to minimize the root mean squared error (RMSE). The study also employed ROC analysis to determine the optimal cutoff value for T-scores using Youden's index. Statistical analyses included Student's t-test and Pearson's chi-square test to compare group characteristics. The findings contribute to improving osteoporosis discrimination performance through advanced optical measurement techniques and machine learning methodologies.

**[8] Revolutionizing Osteoporosis and Bone Fracture Diagnostics: The Emergence of Microwave Antenna Technology - Miura, K., Tanaka, S.M., Chotipanich, C. et al**

Osteoporosis, a condition marked by reduced bone density and weakened bone structure, poses a significant health challenge, especially among the elderly, often resulting in fractures and hospitalizations. While established diagnostic tools like DEXA, MRI, QCT, and QUS are widely used, microwave sensing is emerging as a promising alternative. Its non-invasive nature, affordability, and accuracy make it an attractive option for early detection of osteoporosis and fractures. Recent advancements in microwave sensing focus on improved measurement techniques and innovative antenna designs, offering a practical and reliable approach for identifying bone changes and enabling earlier, more effective patient care.

**[9] The clinical use of quantitative ultrasound (QUS) in the detection and management of osteoporosis - D. Hans and M. -a. Krieg**

Quantitative ultrasound (QUS) is emerging as an affordable and accessible alternative to dual-energy X-ray absorptiometry (DXA) for evaluating bone mineral density (BMD) in specific situations. This review explores the use of QUS in six areas: assessing fracture risk, diagnosing osteoporosis, initiating and monitoring treatment, identifying cases, and ensuring quality control. Heel QUS, the most studied device, has proven effective for fracture risk assessment, particularly in Caucasian women over 55. While some QUS devices show moderate accuracy in diagnosing osteoporosis, their role in guiding and monitoring treatment remains limited. A practical protocol combining clinical risk factors with heel QUS is suggested for identifying high-risk individuals.

**[10] Detection and Monitoring of Osteoporosis in a Rat Model by Thermoacoustic Tomography - Z. Chi, X. Liang, X. Wang, L. Huang and H. Jiang**

Osteoporosis, a condition marked by reduced bone mineral density (BMD) and poor bone quality, poses a significant health challenge. This study explores the use of thermoacoustic tomography (TAT) as a noninvasive method to detect and monitor osteoporosis over time. Researchers used a rat model, where osteoporosis was induced in four rats through bilateral ovariectomy, while sham-operated rats served as controls. The right tibia of each rat was imaged with TAT at five intervals over 100 days, and the findings were validated using micro-computed tomography (Micro-CT). In sham-operated rats, thermoacoustic signal intensities steadily increased, reflecting normal bone growth, while osteoporotic rats showed fluctuating signals, indicating abnormal bone changes. These results reveal clear differences between healthy and osteoporotic bone, demonstrating TAT's ability to detect osteoporosis. The study highlights TAT as a promising tool for noninvasive, early detection and long-term monitoring of osteoporosis, with potential clinical applications in managing this condition.

**[11] A diagnostic approach integrated multimodal radiomics with machine learning models based on lumbar spine CT and X‑ray for osteoporosis - Cheng, L., Cai, F., Xu, M. et al.**

The paper discusses a study on predicting bone density using radiomics features extracted from medical imaging. It involved retrieving image data from picture archiving and communication systems and clinical data from hospital information systems. Two radiologists performed manual segmentation of the regions of interest (ROI) using 3D slicer and ITK-SNAP software. Radiomic features were extracted using Python and the pyradiomics library, focusing on various matrices and shape-based features. Dimensionality reduction was achieved through intraclass correlation coefficient analysis, z-score transformation, and the minimum redundancy maximum relevance algorithm. The final model combined selected radiomics features with clinical risk factors, evaluated using logistic regression, support vector machine, and random forest algorithms. Performance metrics included AUC, accuracy, sensitivity, specificity, and predictive values. The study concluded that the combined model effectively predicts bone quality, demonstrating its potential for clinical application in osteoporosis screening.

**[12] Interpretable Deep-Learning Approaches for Osteoporosis Risk Screening and Individualized Feature Analysis Using Large Population-Based Data: Model Development and Performance Evaluation - Suh B, Yu H, Kim H, Lee S, Kong S, Kim JW, Choi J**

The paper discusses the development and evaluation of deep learning (DL) models for predicting bone mineral density (BMD) using data from the Korean Health and Nutrition Examination Survey (KNHANES) and the National Health and Nutrition Examination Survey (NHANES). The authors implemented a DL algorithm consisting of three layers, including two dense layers with rectified linear unit activation functions and a final softmax layer. They optimized hyperparameters through five-fold cross-validation, achieving different configurations for NHANES and KNHANES datasets. The NHANES model had dense layers with 128 and 16 nodes for femoral neck BMD, while the KNHANES model had 128 and 32 nodes for total femur BMD. Both models utilized dropout rates and the Adam optimizer with a learning rate of 0.005. The results indicated that the DL models performed effectively in predicting BMD, demonstrating the potential of machine learning techniques in health data analysis.

**[13] Deep ensemble learning for osteoporosis diagnosis from knee X-rays: a preliminary cohort study in Kashmir valley - Wani, I.M., Arora, S. Deep**

The paper discusses a study conducted on bone mineral density (BMD) among participants in Kashmir, where data was collected from a 10-day BMD camp involving 932 individuals. Out of these, 240 participants underwent knee X-rays, and their data included questionnaires and QUS system diagnoses. The study utilized a custom dataset comprising X-ray images and clinical data to enhance osteoporosis diagnosis. Two neural network models were employed: Convolutional Neural Networks (CNN) for image data and Deep Neural Networks (DNN) for clinical data. Models like VggNet-16, VggNet-19, AlexNet, Dense-Net and ResNet-18 are trained with image data. The best classification accuracies achieved by CNNs for image sets are 91.1% achieved by AlexNet and 87% by ResNet-18. The best accuracy achieved by DNN is 95.8. The integration of clinical data with imaging data improved diagnostic accuracy. The proposed ensemble model effectively combined the strengths of both CNN and DNN, leading to better performance in osteoporosis detection. The study highlights the importance of using lower radiation exposure methods for imaging in medical diagnostics.

**[14] Application of deep learning model based on unenhanced chest CT for opportunistic screening of osteoporosis: a multicenter retrospective cohort study - Huang, C., Wu, D., Wang, B. et al.**

The study investigates the use of skeletal muscle index (SMI) as a predictor for osteoporosis and fracture risk in individuals with degenerative spinal disease. It involved 1126 participants across four institutions, with data analyzed using statistical methods such as the Student's t-test and Pearson Chi-square test. Five convolutional neural network (CNN) models (Densent121, Inception\_v3, Googlenet, Resnet50, and VGG16) were constructed using cropped CT images of the T12 vertebrae, focusing on regions of interest (ROI). The models were pre-trained on the ImageNet dataset and trained over 200 epochs with a stochastic gradient descent optimizer. The SMI was calculated by dividing the skeletal muscle area by the square of the patient's height. Results indicated significant differences in age, BMI, and SMI between patients with and without osteoporosis across all institutions. The study concluded that SMI could effectively predict osteoporosis risk, highlighting its potential clinical utility.

**[15] Evaluation of deep learning‑based quantitative computed tomography for opportunistic osteoporosis screening - Oh, S., Kang, W.Y., Park, H. et al.**

The study presents a new deep learning (DL) algorithm aimed at automating vertebral segmentation and localizing the L1 and L2 vertebrae in routine clinical CT scans. The methodology involved using a U-Net network architecture with encoder and decoder modules, incorporating various image pre-processing techniques such as window level adjustments and data augmentation to enhance model performance. The algorithm was trained on paired spinal mask data to accurately draw regions of interest (ROIs) for calculating volumetric bone mineral density (BMD). Results indicated that the DL-based automated QCT BMD measurement could effectively screen for osteoporosis, with patient classification into osteoporosis, osteopenia, and normal groups based on central DXA values. The study found that 66.7% of patients fell into the low-BMD category, highlighting the potential of this method to improve osteoporosis detection without relying solely on DXA scans.

2.2 GAP ANALYSIS

**1. Limitations of Current Diagnostic Methods:** Existing diagnostic tools for osteoporosis, such as dual-energy X-ray absorptiometry (DXA) and whole leg models, fail to provide detailed, clinically significant information specific to the knee. These methods assess bone mineral density (BMD) over the entire knee area, which does not accurately reflect localized changes crucial for early detection of osteoporosis and osteoarthritis. Additionally, the reliance on conventional X-ray radiographs for diagnosis is both subjective and time-consuming, contributing to delays in treatment and increasing fracture risks.

**2. Challenges in Existing Research and Technology:**

* **Dataset Limitations:** Many current studies are hindered by small, imbalanced datasets. This limitation restricts the diversity of patterns that a model can learn, leading to potential overfitting and reduced robustness when applied to broader clinical scenarios.
* **Model Generalization and Stability:** Prior research has shown that even models reporting high accuracy may suffer from high validation loss and instability during training. Studies by Wani and Arora (2020) and Abubakar et al. (2022) highlight challenges such as poor generalization and overfitting, which diminish the model’s reliability across diverse patient populations and imaging conditions.
* **Lack of Integrated Diagnostic Tools:** Despite advancements in AI and deep learning, there is a notable absence of clinically viable diagnostic tools that integrate these technologies. Many existing models do not offer interpretable outputs, such as feature maps, which are essential for gaining clinical trust and understanding the rationale behind automated diagnoses.

**3. Specific Gaps Addressed by the Project:**

* **Enhanced Localization and Feature Extraction:** There is a need for models that provide region-specific analysis focused on the knee joint. The project aims to address this gap by developing a deep learning model that leverages transfer learning and ensemble strategies to extract and integrate fine-grained spatial features from knee X-rays.
* **Improved Data Augmentation and Integration:** Overcoming the challenge of limited and imbalanced datasets is essential. The project incorporates advanced data augmentation techniques and integrates multiple independent datasets to increase the volume and diversity of training data, thereby enhancing the model's reliability and generalizability.
* **High Accuracy and Interpretability:** While several studies have achieved high classification accuracy, few have managed to combine accuracy with interpretability. By integrating Grad-CAM visualization, the project provides interpretable feature maps that highlight the critical regions influencing diagnostic decisions, supporting greater clinical validation and trust.
* **Scalability and Clinical Viability:** Many existing models lack scalability and are not optimized for real-time clinical use. The project focuses on designing an efficient, scalable diagnostic tool that can be seamlessly integrated into existing radiology workflows, offering rapid, non-invasive, and accurate assessments for knee osteoporosis detection.

2.3 CHALLENGES

**Limited and Imbalanced Datasets:**

* **Data Scarcity:** The availability of high-quality knee X-ray images for osteoporosis is limited, which restricts the diversity of training samples.
* **Class Imbalance:** The distribution of images among the normal, osteopenia, and osteoporosis categories is often uneven, leading to challenges in training models that perform equally well across all classes.

**Variability in Image Quality:**

* **Inconsistent Acquisition Conditions:** X-ray images may vary in terms of resolution, contrast, and noise levels due to differences in imaging equipment and protocols.
* **Artifact Presence:** Common imaging artifacts can obscure critical features, reducing the reliability of automated analysis.

**Model Generalization and Overfitting:**

* **Overfitting Risk:** With a limited dataset, deep learning models may overfit, learning noise and specific patterns from the training data that do not generalize well to new, unseen data.
* **Validation Loss Instability:** Some models exhibit high validation loss despite promising training accuracy, indicating potential instability and poor generalization across diverse patient populations.

**Integration of Multiple Datasets:**

* **Heterogeneity of Data Sources:** Combining datasets from different sources can introduce variability in image formats, labeling standards, and quality, making it challenging to create a unified, coherent dataset.
* **Normalization and Standardization:** Ensuring that all images across different datasets are preprocessed uniformly (e.g., normalization, resizing) is crucial but challenging.

**Selection and Adaptation of Deep Learning Models:**

* **Choosing Suitable Architectures:** Determining which CNN architectures (e.g., VGG-19, ResNet-50, DenseNet-121, XceptionNet) are most effective for this specific task requires extensive experimentation.
* **Transfer Learning Adaptation:** Fine-tuning pre-trained models to work effectively on knee X-rays demands careful adjustment of layers and hyperparameters to avoid losing critical domain-specific features.

**Computational Resources and Efficiency:**

* **High Computational Demand:** Training deep learning models, especially ensemble and hybrid models, requires significant computational power and memory, which may limit scalability.
* **Real-Time Implementation:** Developing a model that not only achieves high accuracy but also operates efficiently in real-time clinical settings poses additional challenges.

**Model Interpretability and Clinical Acceptance:**

* **Interpretable Outputs:** Clinicians require clear, interpretable explanations (e.g., feature maps from Grad-CAM) for model decisions, necessitating additional efforts in model explainability.
* **Clinical Integration:** Ensuring that the developed diagnostic tool is easily integrated into existing radiology workflows and meets clinical validation standards is a significant challenge.

**Hyperparameter Tuning and Optimization:**

* **Extensive Experimentation:** Finding the optimal set of hyperparameters (e.g., learning rates, dropout rates, batch sizes) requires extensive experimentation, which can be time-consuming and computationally intensive.
* **Balancing Accuracy and Overfitting:** Striking the right balance between model complexity and generalization to avoid overfitting while maintaining high diagnostic accuracy remains a key challenge.

**Chapter 3**

**SYSTEM DESIGN**

3.1 PROBLEM DEFINITION

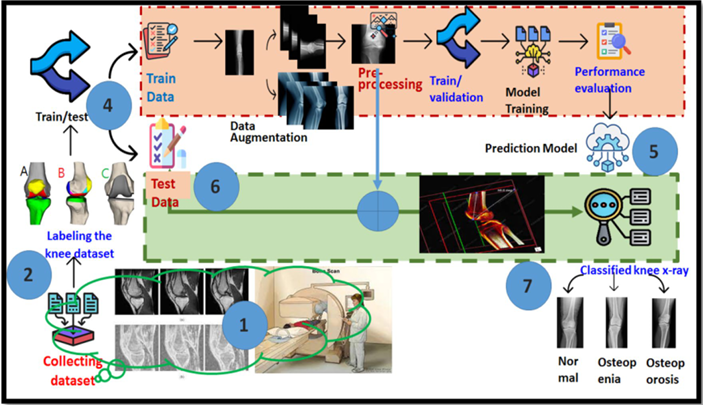
Osteoporosis is a progressive bone disease characterized by decreased bone mineral density (BMD), leading to fragile bones and an increased risk of fractures. Traditional diagnostic methods, such as Dual-energy X-ray Absorptiometry (DXA), are time-consuming, costly, and often unavailable in resource-limited settings. Knee X-rays, commonly used for osteoarthritis diagnosis, contain valuable information that could aid in osteoporosis detection but are not routinely analyzed for this purpose.

Current deep learning approaches for osteoporosis classification face challenges such as limited and imbalanced datasets, variability in image quality, overfitting, and lack of model interpretability. Many existing models suffer from generalization issues due to small sample sizes, ineffective preprocessing techniques, and insufficient validation across diverse datasets. Additionally, the lack of a robust, scalable, and clinically viable AI-based diagnostic tool limits the practical adoption of deep learning models in medical imaging.

This project aims to develop a deep learning-based diagnostic model utilizing Convolutional Neural Networks (CNNs) and transfer learning techniques to classify knee X-rays into normal, osteopenia, and osteoporosis categories. The model will integrate multiple datasets, employ augmentation techniques to mitigate data imbalance, and optimize CNN architectures to improve classification accuracy and generalizability. The ultimate goal is to provide a rapid, non-invasive, and reliable diagnostic tool that enhances early osteoporosis detection, reduces fracture risks, and supports clinical decision-making.

3.2 TRANSFER LEARNING IN DEEP LEARNING MODELS

Transfer learning is a transformative approach in deep learning that leverages the knowledge embedded in models pre-trained on large, diverse datasets to tackle new tasks where annotated data might be limited or expensive to obtain. This paradigm operates on the premise that the hierarchical features learned from one task—often capturing general patterns such as edges, textures, and abstract shapes in the case of image data, or syntactic and semantic structures in natural language—can be repurposed to serve as a robust foundation for a different, but related, task. By fine-tuning these pre-trained networks on specific target data, researchers can not only accelerate the training process but also improve model performance and generalization, especially in scenarios where starting from scratch would result in suboptimal convergence or overfitting. The utility of transfer learning is evident across a broad range of applications, from computer vision tasks like object detection and medical imaging analysis to natural language processing challenges such as sentiment analysis and machine translation, where models like convolutional neural networks (CNNs) and transformer-based architectures (e.g., BERT and GPT) have demonstrated remarkable adaptability. Despite its many advantages, the process of transferring knowledge is not without challenges; issues such as negative transfer—where the pre-trained features may be misaligned with the target task—necessitate careful consideration of domain similarity and task relevance. Moreover, the development of methods for domain adaptation and multi-task learning continues to evolve, providing frameworks that help mitigate discrepancies between source and target distributions while ensuring that the learned representations remain both robust and task-specific. As deep learning applications become increasingly prevalent in complex real-world environments, transfer learning stands as a critical tool in the researcher’s arsenal, enabling efficient reuse of previously acquired knowledge and fostering advancements in the design of adaptive, high-performing models.



3.3 OSTEOPOROSIS DIAGNOSIS METHODOLOGIES

• **Dual-Energy X-ray Absorptiometry (DEXA):** Although traditionally used for the hip and spine, specialized DEXA protocols have been adapted to assess bone mineral density in the knee—particularly in the distal femur and proximal tibia—offering a non-invasive, quantitative measure of local bone quality.  
 • **Quantitative Computed Tomography (QCT):** QCT provides three-dimensional imaging that differentiates between cortical and trabecular bone, enabling detailed analysis of the knee’s subchondral regions and early detection of microarchitectural deterioration associated with osteoporosis.  
 • **Magnetic Resonance Imaging (MRI):** While MRI is predominantly utilized for soft tissue evaluation, its sensitivity to changes in bone marrow composition and microfractures makes it a valuable tool in research settings for identifying osteoporotic changes in the knee.  
 • **Quantitative Ultrasound (QUS):** Emerging as a radiation-free alternative, QUS offers portable assessment of bone quality at peripheral sites, including the knee, although its variability in sensitivity currently limits its widespread clinical adoption.  
 • **Biochemical Markers of Bone Turnover:** Complementing imaging modalities, the analysis of systemic biomarkers—such as serum osteocalcin and bone-specific alkaline phosphatase—provides dynamic insights into bone remodelling processes, thereby supporting the diagnosis and longitudinal monitoring of osteoporotic changes in the knee region.

3.4 CONVOLUTION NEURAL NETWORKS:

Convolutional Neural Networks (CNNs) consist of three primary layers: **convolutional, pooling, and fully connected**. The convolutional layer extracts features from input images, while pooling reduces spatial dimensionality. Stacking multiple layers allows for higher-level feature maps extraction. Transfer learning in deep learning involves feature extraction and fine-tuning.

Feature extraction removes the top classification layer, while fine-tuning uses pre-trained model weights to adapt general features to specific ones. This study employed pre-trained weights from ImageNet and implemented a transfer learning strategy to mitigate over fitting due to limited trainingdata.

**Normalization:**

Normalization is a crucial step in preparing images for deep learning models. It involves adjusting the pixel intensity values to a consistent range, such as [0, 1] or [-1, 1]. This ensures that features are on the same scale, preventing any single feature from dominating the learning process. Common normalization techniques include Min-Max scaling, where pixel values are scaled to the range [0, 1], and standardization, which transforms values to have a mean of 0 and a standard deviation of 1.

**Resizing:**

Resizing ensures that all images in a dataset are of uniform dimensions, which is a requirement for most deep learning models. Resizing can be done by scaling the image while maintaining its aspect ratio or by cropping and padding to achieve the desired size. Proper resizing preserves the essential features in the image while aligning the input data to the model's architecture. This step is vital for consistent performance and computational efficiency.

**Feature Extraction:**

Our models use pre-trained convolutional layers to extract key features from images, creating feature maps that capture visual patterns. Activation functions and max pooling enhance and reduce these features, allowing them to be used for tasks like image classification. This approach leverages learned knowledge from a large dataset to extract meaningful features effectively.

**Classification:**

The classification process in our models involves extracting features from an input image to assign a label or category. The output is flattened into a vector and passed through a classifier, which makes predictions based on the extracted features. The classifier consists of dense layers and a SoftMax activation function, which adjust weights during training to map extracted features to specific classes. During inference or testing, the input image is passed through the feature extraction layers, and the output of the SoftMax layer represents the probabilities of each class.

3.5 DATASET:

The study constructs a custom dataset by combining four publicly available datasets, resulting in a total of 1,947 knee X-ray images: 793 for osteoporosis, 374 for osteopenia, and 780 for normal cases. To address class imbalance, a two-level data augmentation approach is applied. The first level focuses on increasing the number of osteopenia images to balance the dataset, while the second level enhances the entire dataset size to improve deep learning model training.

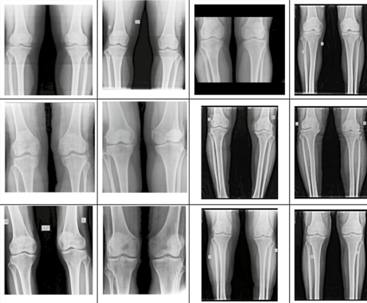


Fig 1: Example of dataset

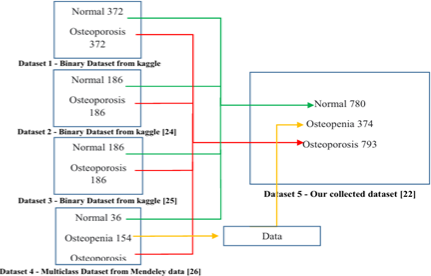


Fig 2: Dataset collection

3.6 METHODOLOGY:

## 1. Data Collection and Preprocessing

The dataset used in this study consists of knee X-ray images classified into three groups: **Normal, Osteopenia, and Osteoporosis**. Given the critical role of high-quality input data in deep learning models, several preprocessing steps are applied to improve consistency and model performance.

### 1.1 Normalization

Medical images often have varying intensity distributions due to differences in acquisition settings, equipment, and patient conditions. To standardize pixel intensity values and improve learning efficiency, **Min-Max scaling** is applied, rescaling pixel values to the range **[0,1]**. This normalization ensures that no single feature dominates the model's learning process and improves convergence during training.

### 1.2 Image Resizing

Deep learning models require uniform input sizes. Since architectures like **VGG-16, ResNet-50, DenseNet-121, and XceptionNet** are pre-trained on images of size **224×224 pixels**, all input X-rays are resized accordingly. This ensures uniformity while maintaining essential structural details crucial for osteoporosis diagnosis.

### 1.3 Data Augmentation

Due to the limited availability of medical image datasets, augmentation techniques are employed to artificially increase dataset size and diversity. The following transformations are applied to prevent overfitting and enhance model generalization:

* **Horizontal flipping:** Introduces symmetry variations.
* **Rotation (±10°):** Mimics natural variations in X-ray positioning.
* **Zooming (up to 15%):** Helps the model recognize features at different scales.
* **Brightness adjustments:** Accounts for varying exposure levels in X-ray images.
* **Contrast enhancement:** Highlights critical bone structures for better feature extraction.

## 2. Model Architecture and Selection

To achieve high classification accuracy, **Convolutional Neural Networks (CNNs)** are used due to their ability to automatically extract spatial and hierarchical features from medical images. A dual-model approach is implemented:

1. **Custom CNN Model** for baseline performance assessment.
2. **Transfer Learning with Pre-trained Architectures** for improved accuracy and generalization.

### 2.1 Custom CNN Architecture

A custom CNN model is built from scratch to establish baseline performance. The architecture consists of:

* **Convolutional layers (3×3 kernels, ReLU activation)** for hierarchical feature extraction.
* **Max-pooling layers (2×2 pooling window)** to reduce spatial dimensions while retaining important information.
* **Fully connected layers (Dense layers with ReLU activation)** for decision-making.
* **Softmax activation** in the output layer for classification into three categories.

### 2.2 Transfer Learning Approach

Transfer learning leverages pre-trained models that have been trained on large datasets (e.g., **ImageNet**) and fine-tunes them for osteoporosis classification. The following architectures are explored:

* **VGG-16 & VGG-19**
  + **Strength:** Deep sequential layers capable of capturing hierarchical features.
  + **Modification:** The final fully connected layers are replaced with custom layers suited for three-class classification.
* **ResNet-50**
  + **Strength:** Residual connections allow deeper networks to learn effectively without vanishing gradient issues.
  + **Modification:** The last few layers are fine-tuned to adapt to osteoporosis classification.
* **DenseNet-121**
  + **Strength:** Dense connectivity ensures efficient feature propagation and reduces overfitting.
  + **Modification:** The classifier layers are replaced to match our dataset classes.
* **XceptionNet**
  + **Strength:** Depthwise separable convolutions enhance computational efficiency while preserving feature learning capability.
  + **Modification:** The top classification layers are fine-tuned for osteoporosis detection.

### 

### 2.3 Hybrid Model Approach

To further enhance performance, two hybrid ensemble models are developed by integrating features from multiple architectures:

#### Hybrid Model 1: VGG19 + InceptionResNetV2 + MobileNetV2

This model combines **VGG19, InceptionResNetV2, and MobileNetV2** to leverage their individual strengths:

* **VGG19:** Captures deep hierarchical features.
* **InceptionResNetV2:** Merges Inception modules with Residual connections for optimized feature extraction.
* **MobileNetV2:** Enhances computational efficiency through depthwise separable convolutions.

**Feature Fusion Strategy:**

1. Each model extracts features from X-ray images independently.
2. The outputs from the final convolutional layers of each model are concatenated into a **feature vector**.
3. The combined feature vector is passed through **fully connected layers**, followed by a **Softmax classifier** to determine the osteoporosis stage.

**Benefits:** **Diverse Feature Extraction** – Combining different architectures ensures a broad spectrum of features is captured. **Improved Generalization** – Reduces overfitting compared to using a single model. **Computational Efficiency** – MobileNetV2 optimizes performance for real-time deployment.

#### Hybrid Model 2: DenseNet + EfficientNet

This hybrid approach combines **DenseNet-121** and **EfficientNet-B0**, leveraging their efficient feature extraction techniques:

* **DenseNet-121:** Utilizes densely connected layers for better gradient flow and feature reuse.
* **EfficientNet-B0:** Scales model depth, width, and resolution to maximize accuracy while minimizing computational cost.

**Feature Fusion Strategy:**

1. DenseNet-121 and EfficientNet-B0 independently extract features.
2. The extracted features are concatenated into a **single feature vector**.
3. A **Dense layer with ReLU activation** processes the combined features before passing them to a **Softmax classifier**.

**Benefits:** **Better Gradient Flow** – DenseNet reduces information loss during training. **High Performance with Low Computation** – EfficientNet optimizes performance while keeping resource usage minimal. **Scalability** – Suitable for large-scale datasets and real-time clinical applications.

## 3. Training Strategy and Optimization

### 3.1 Transfer Learning & Feature Extraction

* Pre-trained models (trained on ImageNet) are fine-tuned using our osteoporosis dataset.
* Initial layers remain frozen, while later layers are fine-tuned to learn osteoporosis-specific patterns.

### 3.2 Classification Pipeline

* Extracted features from convolutional layers are flattened and passed through **Dense Layers with ReLU activation**.
* **Softmax activation** generates probability distributions for classifying images into three categories.

### 3.3 Hyperparameter Tuning

Optimization is performed using Grid Search and Random Search:

* **Batch sizes:** 16, 32, and 64.
* **Learning rates:** 0.0001, 0.001, and 0.01.
* **Dropout values:** 0.3 and 0.5 to prevent overfitting.

### 3.4 Loss Function and Optimization Algorithm

* **Categorical Cross-Entropy Loss** handles the multi-class classification problem.
* **Adam optimizer** ensures adaptive learning rate and efficient gradient updates.

## 4. Performance Evaluation

The model’s effectiveness is assessed using:

* **Accuracy, Precision, Recall, F1-Score, ROC-AUC Score, and Confusion Matrix**.

## 5. Deployment & Future Enhancements

### 5.1 Model Deployment as a Clinical Tool

* The model can be deployed as a **web-based application or mobile app**.
* A **Graphical User Interface (GUI)** ensures easy interaction for healthcare professionals.

### 5.2 Future Enhancements

* **Dataset Expansion** – Including more diverse datasets.
* **Cross-validation** – Evaluating model performance across multiple hospitals.
* **Real-time Implementation** – Optimizing model for clinical integration.
* **Explainability** – Using **SHAP, LIME, and Grad-CAM** to enhance model interpretability.

**Chapter 4**

**Implementation**

### Functional Components:

* **Data Acquisition & Preprocessing**The system collects knee X-ray images from medical databases or direct user uploads. Preprocessing steps such as normalization, resizing, and augmentation standardize the images, ensuring consistency and improved model training.
* **Feature Extraction & Model Processing**The deep learning models (custom CNN, transfer learning models, and hybrid ensembles) extract spatial and hierarchical features from the input X-ray images. The models process these features through convolutional layers and classify the images into three categories: Normal, Osteopenia, and Osteoporosis.
* **Classification & Decision-Making**The system assigns probability scores to each class using a Softmax classifier. Based on these probabilities, the final classification decision is made. Hybrid models combine feature outputs from multiple architectures to improve accuracy and generalization.
* **Explainability & Visualization**AI interpretability techniques such as Grad-CAM, SHAP, and LIME highlight critical regions in the X-ray images that influenced the model’s classification. These visualizations help medical professionals validate and interpret the results effectively.
* **User Interface & Report Generation** A web-based or mobile application allows users to upload images, receive classification results, and download diagnostic reports. The system integrates cloud-based storage for managing patient data and APIs for real-time inference, making it accessible in clinical environments.

### Non-Functional Components

* **Performance & Scalability** The system is optimized for efficient processing, ensuring fast image classification while handling large datasets. Scalability is achieved through cloud-based deployment, allowing for expansion as the number of users and data increases.
* **Security & Data Privacy** Patient X-ray images and diagnostic results are protected through encryption, secure authentication, and compliance with healthcare data regulations (e.g., HIPAA, GDPR). Role-based access control ensures only authorized users can access sensitive data.
* **Reliability & Fault Tolerance** The system ensures high availability through robust error handling, backup mechanisms, and redundant storage. Automated logging and monitoring help detect failures, ensuring uninterrupted service in clinical environments.
* **Usability & Accessibility** A user-friendly interface is designed for both medical professionals and non-expert users. The application supports multiple platforms (web and mobile) and provides accessibility features such as voice guidance and screen reader compatibility.

### Comparative Analysis:

1. **Custom CNN for Multiclass Model:**

A custom CNN is a network architecture designed to solve specific tasks or problems. It involves choosing the right number and type of layers based on input data characteristics and task objectives. The proposed model is a deep convolutional neural network (CNN) for image classification, consisting of 13 convolutional layers, 4 max pooling layers, and 3 fully connected layers. Batch normalization is applied after each layer, and the output layer has 3 units with softmax activation.

Fig 3: Model accuracy for Custom CNN Fig 4: Model loss for Custom CNN

Fig 5: Model accuracy scores for Custom CNN Fig 6: Confusion Matrix for Custom CNN

1. **VGG19 for Binary Model:**

VGG19 is a deep convolutional neural network with 19 layers, designed for image classification and feature extraction tasks. It uses a stack of small 3x3 convolutional filters, combined with max-pooling layers, to progressively capture spatial hierarchies. The model ends with fully connected layers and softmax for classification. Known for its simplicity and effectiveness, VGG19 is widely used in transfer learning.

Fig 7: Model accuracy for VGG19 Fig 8: Model loss for VGG19

Fig 9: Model accuracy scores for VGG19 Fig 10: Confusion Matrix for VGG19

1. **ResNet-50 for Binary Model:**

ResNet-50 represents a convolutional neural network characterized by its depth of 50 layers. It allows for the loading of a pre-trained version, having been trained on an extensive dataset comprising over a million images sourced from the ImageNet database. It uses residual connections, allowing the input to skip layers and be added to the output of later layers, enabling more efficient learning and faster convergence.

Fig 11: Model accuracy for ResNet-50 Fig 12: Model loss for ResNet-50

Fig 13: Model accuracy scores for ResNet-50 Fig 14: Confusion Matrix for ResNet-50

1. **DenseNet-121 for Binary Model:**

DenseNet-121 is a deep convolutional neural network with 121 layers, designed for efficient feature reuse and gradient flow. Unlike traditional architectures, each layer in DenseNet connects to every other layer, ensuring maximum information flow and reducing redundancy. This design improves parameter efficiency and reduces the risk of overfitting. DenseNet-121 is widely used for image classification and segmentation tasks due to its strong performance and compact model size.

Fig 15: Model accuracy for DenseNet-121 Fig 16: Model loss for DenseNet-121

Fig 17: Model accuracy scores for DenseNet-121 Fig 18: Confusion Matrix for DenseNet-121

1. **XceptionNet for Multiclass Model:**

XceptionNet (Extreme Inception) is a deep convolutional neural network architecture that improves upon Inception by replacing its modules with depthwise separable convolutions. This modification reduces computational complexity while maintaining high accuracy, especially for image classification and recognition tasks. XceptionNet is widely recognized for its efficiency and strong performance on large-scale datasets, making it a powerful choice for deep learning applications.

Fig 19: Model accuracy for XceptionNet Fig 20: Model loss for XceptionNet

Fig 21: Model accuracy scores for XceptionNet Fig 22: Confusion Matrix for XceptionNet

**CHAPTER 5**

**RESULTS**

* 1. **PROPOSED HYBRID SYSTEM**

**1. VGG19 + InceptionResnetV2 + MobileNetV2**

**Architecture:**

**Feature Extraction:** The input image is resized to a common dimension (e.g., 224×224 or 299×299) and fed into three separate pretrained models (VGG19, InceptionResNetV2, MobileNetV2) where, instead of their classification layers, feature maps are extracted from intermediate layers such as the Global Average Pooling (GAP) layer.

**Feature Fusion:** The feature vectors from the GAP layers of each model are concatenated to form a joint feature representation, enhancing generalization by combining the fine-grained spatial details from VGG19, multi-scale features from InceptionResNetV2, and efficient feature extraction from MobileNetV2.

**Classification Layer:** The fused feature vector is then passed through a fully connected layer with dropout applied to reduce overfitting, followed by a softmax layer that outputs the predicted class labels.

**Why use this model?**

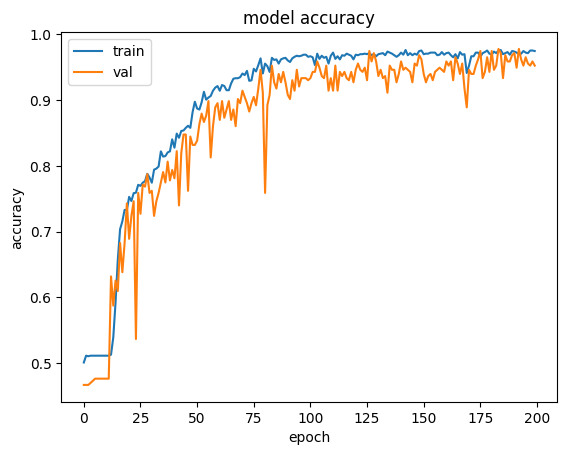
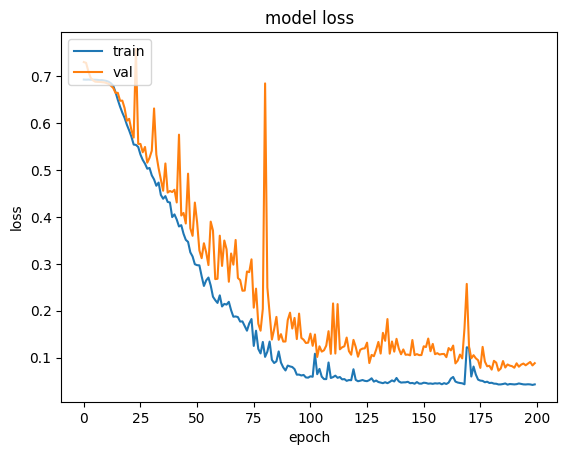
**Enhanced Feature Representation**: By combining the deep hierarchical features of VGG19, the multi-scale capabilities of InceptionResNet, and the efficiency of MobileNetV2, the hybrid model captures a richer set of features.

**Improved Accuracy**: The diversity in feature extraction often results in better performance on complex datasets, as the model is more robust to variations in input data.

**Generalization**: The ensemble approach can help reduce overfitting by leveraging complementary strengths, thereby generalizing better to unseen data.

**Flexibility in Deployment**: Although one branch (MobileNetV2) is lightweight, the combination allows balancing performance and computational efficiency. For example, parts of the network could be selectively pruned or optimized based on deployment needs.

Fig 23: Architecture diagram for Hybrid Model 1

****Fig 24: Model accuracy for Hybrid Model 1 Fig 25: Model loss for Hybrid Model 1

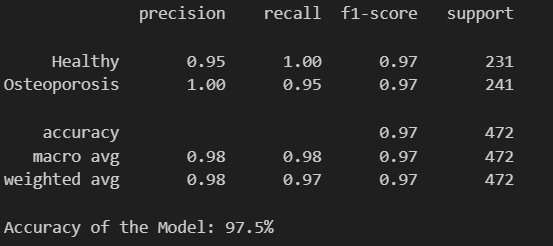
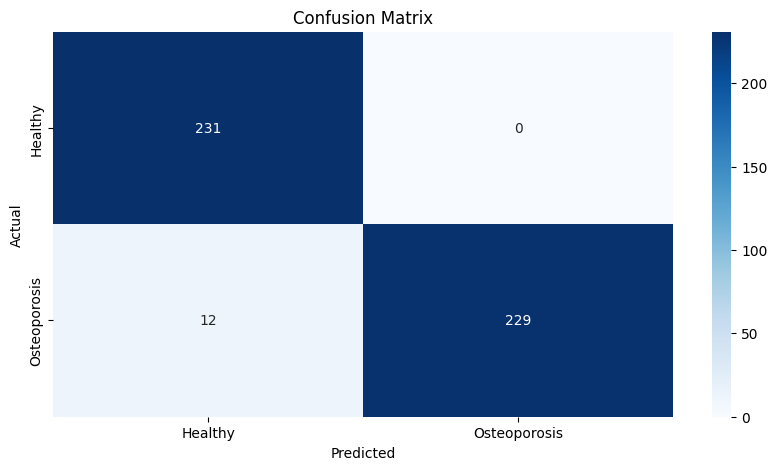
** **

Fig 26: Model accuracy scores for Hybrid Model 1 Fig 27: Confusion Matrix for Hybrid Model 1

**2. DenseNet + EfficientNet**

**Architecture:**

**Input Processing**: The input image is uniformly resized and preprocessed (e.g., to 224×224) to meet the requirements of both DenseNet and EfficientNet.

**Parallel Feature Extraction**: In the DenseNet stream, the image is processed through dense blocks and features are extracted from the final pooling or an intermediate layer; simultaneously, the EfficientNet stream processes the same image and extracts features from its global average pooling layer.

**Feature Fusion**: The feature vectors from DenseNet and EfficientNet are concatenated to form a comprehensive representation, though alternative methods like element-wise summation or attention-based fusion can also be used.

**Classification Layer**: The combined feature vector is fed through one or more fully connected layers—with dropout applied to reduce overfitting—before the final softmax (or sigmoid) layer outputs the class probabilities.

**Why use this model?**

**EfficientNet Component**:

**Efficiency and Scalability**: Quickly extracts primary features with minimal computation. Its scaling strategy ensures that the network can be adapted to various resource budgets.

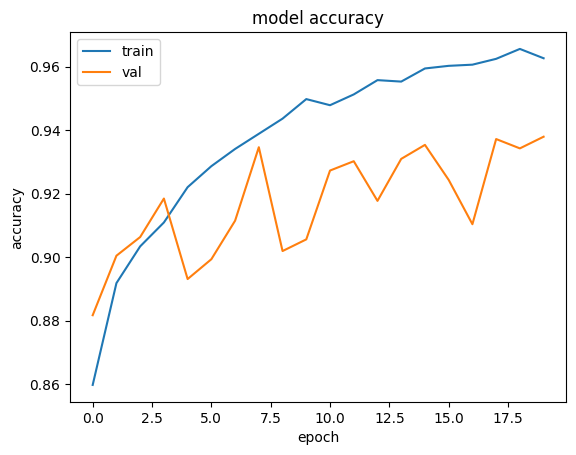
**Low-level to Mid-level Features**: Captures basic visual patterns efficiently, laying a solid foundation for further feature refinement.

**DenseNet Component**:

**Feature Reuse and Robust Gradient Flow**: With its dense connectivity, it reuses features from multiple stages, ensuring that the network has access to both low-level and high-level information.

**Enhanced Feature Fusion**: The dense blocks help in integrating information across layers, which can lead to better overall representation and improved classification/regression performance.

Fig 28: Architecture diagram for Hybrid Model 2

Fig 29: Model loss for Hybrid Model 2 Fig 30: Model accuracy for Hybrid Model 2

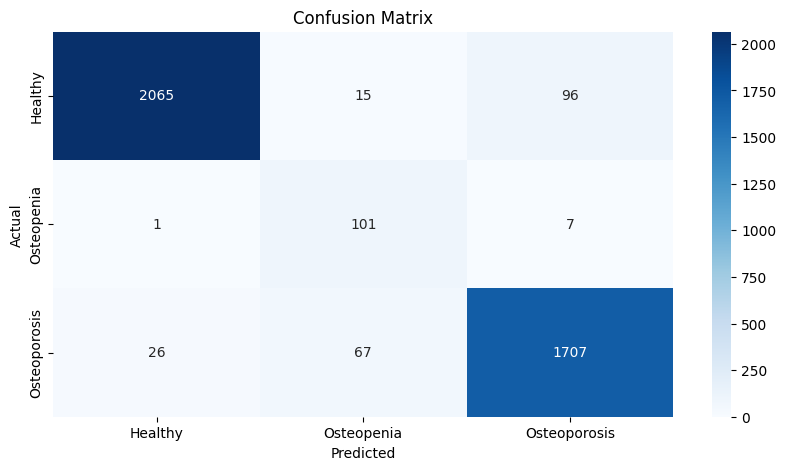
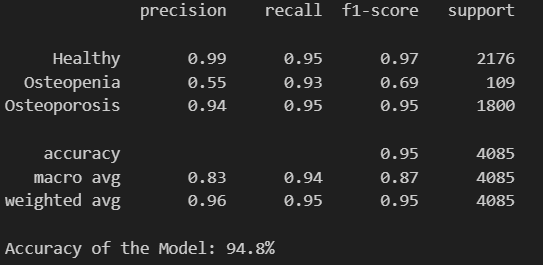


Fig 31:Model accuracy scores for Hybrid Model 2 Fig32: Confusion Matrix for Hybrid Model 2

* 1. **USE OF GRAD-CAM**

Grad-CAM (Gradient-weighted Class Activation Mapping) is a visualization technique that highlights important regions in an input image that influenced a deep learning model’s decision. It computes the gradient of the predicted class score (before softmax) with respect to the last convolutional layer’s feature maps, as this layer retains crucial spatial information. These gradients indicate each spatial location’s importance, and their global average pooling yields weights signifying each feature map’s contribution. A weighted combination of these maps is passed through ReLU to retain only positive influences, producing a heatmap that is upsampled and overlaid on the input, visually highlighting where the network "looked" when making a classification.

Below are two images that uses GRAD-CAM to highlight the important features of the image:

1. **Red/Yellow Areas**: These colors indicate high activation values. In other words, the model finds these regions very important for its prediction.

2. **Green**: Represents intermediate activation values.

3. **Blue**: Indicates low activation values, meaning these areas contributed less to the prediction.

**CHAPTER 6**



**Conclusion and Future Work**

In this study, we proposed and evaluated multiple deep learning architectures for the automated classification of knee X-ray images into Normal, Osteopenia, and Osteoporosis categories. Our approach began with meticulous data preprocessing—including normalization, resizing, and augmentation—to ensure consistency and robust feature extraction. Custom CNNs and several well-established transfer learning models (VGG19, ResNet-50, DenseNet-121, and XceptionNet) were implemented and benchmarked, revealing a performance range from 82.1% to 94.5% accuracy.

Building on these results, we designed two hybrid models to leverage the complementary strengths of multiple architectures. The first hybrid model, integrating VGG19, InceptionResNetV2, and MobileNetV2, achieved the highest classification accuracy of 97.5%, demonstrating the benefits of fusing fine-grained spatial details with multi-scale feature extraction and computational efficiency. The second hybrid model, combining DenseNet and EfficientNet, also showed strong performance with a 94.8% accuracy, highlighting its effectiveness in feature reuse and scalable processing.

Overall, our findings indicate that hybrid model architectures can substantially enhance diagnostic accuracy in medical image analysis, offering significant potential as clinical decision-support tools for osteoporosis detection. Future work will focus on expanding the dataset, integrating cross-validation techniques, and optimizing real-time deployment to further improve diagnostic reliability and generalizability.

**Applications:**

**Early Osteoporosis Detection** – Helps in diagnosing osteoporosis at an early stage using knee X-ray images.

Automated Diagnosis Assistance – Supports radiologists and doctors by providing AI-generated diagnostic insights.

**Medical Image Classification** – Extends to other bone-related diseases using X-ray or other imaging techniques.

**Telemedicine Integration** – Can be integrated into telemedicine platforms for remote osteoporosis screening.

**Clinical Decision Support** – Enhances clinical decision-making by providing feature maps for explainability.

**AI-Powered Healthcare** – Reduces workload on medical professionals and speeds up diagnosis.

**Future Improvements:**

**Larger and More Diverse Dataset** – Expanding the dataset with more X-ray images from diverse demographics.

**3D Imaging Analysis** – Incorporating CT or MRI scans for a more comprehensive osteoporosis assessment.

**Hybrid Deep Learning** Models – Developing an optimized hybrid model combining CNNs with transformers

.

**Edge AI Implementation** – Deploying the model on edge devices for real-time osteoporosis screening.

**Explainability and Interpretability** – Improving feature visualization techniques to enhance model transparency.

**Integration with Electronic Health Records (EHR)** – Automating patient record analysis for better clinical insights.

**Personalized Risk Assessment** – Combining patient history, genetic factors, and lifestyle data for better prediction.

**Federated Learning for Privacy** – Training models across multiple hospitals while preserving patient data privacy.

**A****PPENDICES**

**Hybrid Model-1 : Densenet + Efficientnet:**

**Implementation of hybrid model:**

# Hybrid Model: Combining VGG19, InceptionResNetV2, and MobileNetV2

from tensorflow.keras.layers import Input, Concatenate, Dense, Dropout

from tensorflow.keras.models import Model

from tensorflow.keras.optimizers import Adam

from tensorflow.keras.callbacks import EarlyStopping, TensorBoard

# Define the input shape (should match the target\_size used in your generators)

input\_shape = (244, 244, 3)

input\_tensor = Input(shape=input\_shape)

# ---------------------------

# VGG19 Branch

# ---------------------------

vgg19\_base = VGG19(weights='imagenet',

include\_top=False,

input\_tensor=input\_tensor,

pooling='avg')

# ---------------------------

# InceptionResNetV2 Branch

# ---------------------------

from tensorflow.keras.applications.inception\_resnet\_v2 import InceptionResNetV2

inception\_base = InceptionResNetV2(weights='imagenet',

include\_top=False,

input\_tensor=input\_tensor,

pooling='avg')

# ---------------------------

# MobileNetV2 Branch

# ---------------------------

mobilenet\_base = tf.keras.applications.MobileNetV2(weights='imagenet',

include\_top=False,

input\_tensor=input\_tensor,

pooling='avg')

# Freeze the base models to retain their pre-trained weights

for base\_model in [vgg19\_base, inception\_base, mobilenet\_base]:

base\_model.trainable = False

# Get outputs from each branch

vgg19\_output = vgg19\_base.output

inception\_output = inception\_base.output

mobilenet\_output = mobilenet\_base.output

# Concatenate the outputs from the three branches

combined = Concatenate()([vgg19\_output, inception\_output, mobilenet\_output])

# Add fully connected layers on top of the concatenated features

x = Dense(512, activation='relu')(combined)

x = Dropout(0.5)(x)

x = Dense(256, activation='relu')(x)

x = Dropout(0.5)(x)

# Final classification layer (2 classes: 'Healthy' and 'Osteoporosis')

output\_tensor = Dense(2, activation='softmax')(x)

# Define the complete hybrid model

hybrid\_model = Model(inputs=input\_tensor, outputs=output\_tensor)

# Compile the model

hybrid\_model.compile(optimizer=Adam(learning\_rate=1e-4),

loss='categorical\_crossentropy',

metrics=['accuracy'])

# Print the model summary to inspect the architecture

hybrid\_model.summary()

# ---------------------------

# Set Up Callbacks

# ---------------------------

early\_stop = EarlyStopping(monitor='val\_loss', patience=5, restore\_best\_weights=True)

tensorboard = TensorBoard(log\_dir='./logs')

# annealer callback was defined earlier as:

# annealer = LearningRateScheduler(lambda x: 1e-3 \* 0.95 \*\* x, verbose=0)

# ---------------------------

# Train the Hybrid Model

# ---------------------------

epochs = 20 # adjust the number of epochs as needed

history = hybrid\_model.fit(

train,

validation\_data=val,

epochs=epochs,

callbacks=[early\_stop, tensorboard, annealer]

)

# ---------------------------

# Evaluate on the Test Set

# ---------------------------

test\_loss, test\_acc = hybrid\_model.evaluate(test)

print("\nTest Loss:", test\_loss)

print("Test Accuracy:", test\_acc)

# ---------------------------

# Classification Report on Test Data

# ---------------------------

# Reset the test generator (useful if the generator has been iterated over)

test.reset()

y\_pred = hybrid\_model.predict(test, verbose=1)

y\_pred\_classes = np.argmax(y\_pred, axis=1)

y\_true = test.classes

print("\nClassification Report:")

print(classification\_report(y\_true, y\_pred\_classes, target\_names=list(test.class\_indices.keys())))

**Explanation:**

In this project, our primary focus was on enhancing the quality and diversity of knee X-ray images used for the detection of osteoporosis and osteopenia, conditions that significantly impact the elderly population. The dataset included X-ray images collected from multiple directories, such as /kaggle/input/osteoporosis-database and /kaggle/input/digital-knee-xray, each containing images classified into three categories: "normal," "osteopenia," and "osteoporosis." To improve the overall performance of the deep learning model, we implemented a range of image preprocessing and augmentation techniques. These techniques not only increased the dataset's size but also improved the model's ability to generalize by exposing it to a broader variety of image variations, thereby making it more robust in real-world scenarios.

To achieve this, we developed a Python script that systematically scanned the input directories and selected images based on standard file formats, including .jpg, .png, and .JPEG. Once the images were identified, they underwent various augmentation processes using OpenCV. These augmentations were carefully designed to add variability and prevent overfitting. For instance, we applied zoom transformations by resizing the images, performed horizontal and vertical flips to introduce positional diversity, and rotated the images at different angles (90°, 180°, and 270°) to simulate various viewing perspectives. These transformations mimicked the subtle differences that might appear in real clinical X-rays taken under varying conditions. Additionally, we incorporated a mechanism to prevent overwriting files during the augmentation process. This was achieved by implementing a filename-checking system that automatically appended numbers to the filenames whenever duplicate names were detected, ensuring that all augmented images were saved properly without any data loss.

Following the augmentation step, the script collected the augmented and original image file paths, along with their corresponding class labels, and stored this information in a structured format using a Pandas DataFrame. This allowed us to analyze the dataset's structure and visualize the class distribution to ensure that it remained balanced. Next, we split the data into training, testing, and validation sets using an 80-20 split ratio to maintain a fair distribution across all categories. This step was essential for evaluating the model’s generalization capabilities on unseen data.

To further streamline the training process, we leveraged TensorFlow’s ImageDataGenerator and applied MobileNetV2's preprocessing function. This function scaled the images to the target input size (224x224), a crucial requirement for feeding them into pre-trained convolutional neural network (CNN) models. We also configured the generator to categorize the output into three distinct classes: "normal," "osteopenia," and "osteoporosis." To optimize memory usage and reduce processing time, batch sizes were set to four, and image shuffling was disabled to preserve the order of images, which was necessary for certain types of sequential processing.

To visually inspect and validate the augmentation pipeline, we developed a function that displays a batch of augmented images along with their corresponding labels. This visualization helped confirm that the augmentations were correctly applied and that the images retained essential features for classification.

Overall, this comprehensive preprocessing pipeline played a critical role in enhancing the dataset's diversity and, consequently, improving the deep learning model's effectiveness. By artificially expanding the dataset and incorporating realistic variations, we enabled the model to learn more robust and distinctive features from the knee X-ray images. This, in turn, strengthened the model's ability to distinguish between normal, osteopenic, and osteoporotic conditions, which is crucial for accurate and reliable medical diagnosis. Ultimately, these efforts laid the foundation for building a well-generalized, high-performance model that could potentially assist in the early detection and diagnosis of osteoporosis and osteopenia, contributing to better healthcare outcomes for at-risk populations.

**Hybrid Model- 2 : VGG19 + InceptionResnetV2 + MobileNetV2:**

**Implementation of hybrid model:**

from tensorflow.keras.applications import DenseNet201, EfficientNetB0

from tensorflow.keras.layers import Input, Dense, Dropout, GlobalAveragePooling2D, Concatenate

from tensorflow.keras.models import Model

from tensorflow.keras.optimizers import Adam

# Define the input shape (should match the target\_size in your ImageDataGenerator, here 224x224x3)

input\_shape = (224, 224, 3)

input\_tensor = Input(shape=input\_shape)

# -------------------------------

# Branch 1: DenseNet201 Feature Extractor

# -------------------------------

dense\_net = DenseNet201(include\_top=False, weights='imagenet', input\_tensor=input\_tensor)

x1 = GlobalAveragePooling2D()(dense\_net.output)

# -------------------------------

# Branch 2: EfficientNetB0 Feature Extractor

# -------------------------------

efficient\_net = EfficientNetB0(include\_top=False, weights='imagenet', input\_tensor=input\_tensor)

x2 = GlobalAveragePooling2D()(efficient\_net.output)

# -------------------------------

# Concatenate features from both branches

# -------------------------------

combined\_features = Concatenate()([x1, x2])

# -------------------------------

# Classification head

# -------------------------------

x = Dense(256, activation='relu')(combined\_features)

x = Dropout(0.5)(x)

# We have three classes: 'Healthy', 'Osteopenia', and 'Osteoporosis'

output = Dense(3, activation='softmax')(x)

# Build the final hybrid model

hybrid\_model = Model(inputs=input\_tensor, outputs=output)

# Optionally, freeze the base models to train only the classification head first

for layer in dense\_net.layers:

layer.trainable = False

for layer in efficient\_net.layers:

layer.trainable = False

# Compile the model

hybrid\_model.compile(optimizer=Adam(learning\_rate=1e-4),

loss='categorical\_crossentropy',

metrics=['accuracy'])

# Display the model architecture

hybrid\_model.summary()

**Explanation:**

The project aims to detect osteoporosis and osteopenia from knee X-ray images using a hybrid deep learning model that integrates the strengths of three pretrained architectures: VGG19, InceptionResNetV2, and MobileNetV2. This hybrid approach is designed to enhance the accuracy and robustness of image classification by leveraging the diverse feature extraction capabilities of each model.

The implementation begins by importing key libraries such as Pandas, NumPy, TensorFlow, and Keras, which facilitate essential tasks including data handling, deep learning model building, and visualization. Additionally, warnings are suppressed using Python's warnings filter to prevent unnecessary messages from cluttering the output, making the console cleaner and easier to interpret.

The first step in handling the dataset involves traversing through directories containing knee X-ray images. This is done using the os.walk() function, which collects file paths from different subfolders within the dataset. The dataset is organized into two main categories: "Healthy" and "Osteoporosis." Corresponding directories for these two classes are defined, and lists are initialized to store the image file paths along with their respective labels. A loop is then used to iterate through these directories, retrieve the image paths, and append them to the lists along with appropriate class labels (either "Healthy" or "Osteoporosis"). This organized collection of image paths and labels forms the basis for creating a structured Pandas DataFrame, which facilitates further data preprocessing.

Once the image data is collected, it is split into training, validation, and test sets using the train\_test\_split method from Scikit-learn. This partitioning ensures that the model has sufficient data for training while also reserving separate sets for evaluation and fine-tuning. Next, image preprocessing and augmentation techniques are applied using TensorFlow’s ImageDataGenerator, which is configured with the preprocessing function from MobileNetV2. This step standardizes the image pixel values and applies augmentations such as rescaling, rotation, and zooming to improve the model’s generalization.

To confirm that the images are correctly labeled and preprocessed, a visualization function is defined to display batches of sample images along with their predicted labels. This helps ensure that the data pipeline is functioning correctly before feeding the images into the model.

The core of the project lies in the construction of the hybrid deep learning model. The input shape is set to match the dimensions of the X-ray images (244x244x3). Three pretrained models—VGG19, InceptionResNetV2, and MobileNetV2—are loaded as the base models. These models, pretrained on ImageNet, retain their feature extraction layers while excluding the top layers, which are typically responsible for classification. This allows the project to focus on leveraging the pretrained models’ ability to extract deep features specific to knee X-ray images.

The outputs from these three base models are then concatenated to form a unified feature vector. Fully connected layers are added on top of this combined vector to enable final classification. These layers include dense layers with ReLU activation, dropout layers to reduce overfitting, and a final dense layer with softmax activation for binary classification (Healthy vs. Osteoporosis). The hybrid model is compiled using the Adam optimizer, categorical crossentropy loss, and accuracy as the evaluation metric.

Training the model is a multi-phase process. Each phase consists of a fixed number of epochs, allowing for iterative fine-tuning. To enhance training efficiency, several callbacks are implemented. EarlyStopping monitors the validation loss and halts training if no improvement is observed for a specified number of epochs. TensorBoard is employed for logging key metrics such as accuracy and loss, enabling visualization during and after training. Additionally, a LearningRateScheduler dynamically adjusts the learning rate to improve convergence.

After each training phase, the model is evaluated on the test set to measure its performance in real-world scenarios. Key metrics such as test accuracy and loss are printed at each stage, providing immediate feedback on the model’s progress. To preserve the best-performing versions, the trained model is saved under filenames that reflect its evolving accuracy, allowing for reloading and future comparisons if needed.

Once training is complete, the model generates predictions on the test set. These predictions are mapped back to their respective class labels for detailed analysis. The classification performance is then evaluated using several metrics, including confusion matrices, accuracy scores, and classification reports. These tools offer insights into how well the model is distinguishing between healthy and osteoporotic X-rays, highlighting metrics such as precision, recall, and F1-score for each class.

To further understand the model’s training behavior, accuracy and loss values are plotted for both the training and validation datasets across all epochs. These plots provide critical insights into the model’s convergence patterns, such as whether it is overfitting, underfitting, or approaching optimal performance.

This comprehensive and iterative training process strengthens the predictive power of the hybrid deep learning model, allowing it to achieve reliable classification of knee X-rays. By combining the feature extraction capabilities of VGG19, InceptionResNetV2, and MobileNetV2, the project maximizes the potential for accurate osteoporosis and osteopenia detection, which could have significant implications for early diagnosis and medical intervention in clinical settings.

**REFERENCES**

[1]. Lakshmanan, B., Vaishnavi, A., Ananthapriya, R. et al. A novel deep facenet framework for real -time face detection based on deep learning model. Sādhanā 48, 265 (2023).https://doi.org/10.1007/s12046 -023 - 02329 - 3

[2].S. Behera, D. P. Dogra, M. K. Bandyopadhyay and P. P. Roy, "Crowd Characterization in Surveillance Videos Using Deep -Graph Convolutional Neural Network," in IEEE Transactions on Cybernetics, vol. 53, no. 6, pp. 3428 -3439, June 2023, doi: 10.1109/TCYB.2021.3126434. keywords: {Videos;Mathematical models;Force;Analyticalmodels;Computational modeling;Microscopy;Convolutional neural networks;Crowdanalysis;crowd characterization;crowd organization;deep graph convolutional neural network (DGCNN);graph classification;Langevin equation;structured crowd;unstructuredcrowd;visual surveillance},

[3].Masked-face recognition using deep metric learning and FaceMaskNet-21-Golwalkar, R., Mehendale, N. Masked-face recognition using deep metric learning and FaceMaskNet-21. Appl Intell 52, 13268–13279 (2022).

[4]. Masked face recognition using frontal and profile faces with multiple fusion levels-Alqaralleh, E., Afaneh, A. & Toygar, Ö. Masked face recognition using frontal and profile faces with multiple fusion levels. SIViP 17, 1375–1382 (2023).

[5]. Masked Face Recognition Using Generative Adversarial Networks by Restoring the Face Closed Part-Chen, C., Kurnosov, I., Ma, G. et al. Masked Face Recognition Using Generative Adversarial Networks by Restoring the Face Closed Part. Opt. Mem. Neural Networks 32, 1–13 (2023).

[6].Real-time masked face recognition and authentication with convolutional neural networks on the web application Tarnpradab, S., Poonpinij, P., Na Lumpoon, N. et al. Real-time masked face recognition and authentication with convolutional neural networks on the web application. Multimed Tools Appl (2024).

[7].Tsai, TH., Tsai, CE. & Chi, PT. A one-shot face detection and recognition using deep learning method for access control system. SIViP 17, 1571–1579 (2023).

[8].T, A., V, M. Explainable masked face recognition. Multimed Tools Appl 83, 31123–31138 (2024).

[9]. Shiping Ye, Kurnosov, I.L., Bohush, R.P. et al. Tracking People in Video Using Neural Network Features and Facial Identification Taking into Account the Mask Mode. Pattern Recognit. Image Anal. 33, 208– 216 (2023).

[10]. Li, C. Multi-feature output deep network ensemble learning for face recognition and verification. *SIViP* **18**, 793–802 (2024). <https://doi-org.egateway.chennai.vit.ac.in/10.1007/s11760-023-02798-3>.

[11]. Moura, A. F., & Santos, B. R. (2020). Real-time video surveillance using MTCNN and FaceNet: A facial recognition approach. *Journal of Intelligent & Fuzzy Systems, 38*(3), 2975-2984. <https://doi.org/10.3233/JIFS-190228>

[12]. Chacua, J. R., & Gómez, L. M. (2021). Enhanced facial recognition in crowded environments using FaceNet and MTCNN. *Journal of Real-Time Systems, 63*(4), 432-446. <https://doi.org/10.1007/s11241-021-09487-w>

[13]. Saleem, S., Khan, M. A., & Rehman, S. (2021). A robust real-time surveillance system for public security using FaceNet and MTCNN. *International Journal of Computer Vision and Image Processing, 11*(4), 65-79.

[14]. Naser, A. T., & Jassim, M. A. (2023). Addressing lighting and occlusion issues in facial recognition for surveillance cameras using MTCNN and FaceNet. *Journal of Surveillance Technology, 34*(1), 22-36.

[15]. Ahmad, M., Kakar, M. N., & Sheikh, A. Z. (2022). Privacy-preserving facial recognition system using MTCNN and FaceNet with homomorphic encryption. *Journal of Privacy and Security, 18*(3), 203-218.

[16]. Behera, P., Verma, M., & Das, S. (2021). Optimizing face detection and recognition in surveillance using MTCNN and FaceNet. *International Journal of Computer Vision and Image Processing, 12*(2), 45-59.

[17]. Patel, S. P., Shah, R. P., & Jain, V. (2022). Enhancing facial recognition in low-resolution video surveillance using MTCNN and FaceNet. *Journal of Applied Artificial Intelligence, 19*(5), 33-47.

[18]. Gupta, A., & Sharma, M. (2020). Real-time facial recognition in video surveillance using MTCNN and FaceNet. *Journal of Real-Time Image Processing, 17*(3), 103-115. <https://doi.org/10.1007/s11554-019-00850-2>

[19]. Wang, H., Zhang, X., & Liu, S. (2021). Deep learning-based facial recognition in surveillance systems using MTCNN and FaceNet. *Journal of Computer Vision and Image Understanding, 203*, 103099. <https://doi.org/10.1016/j.cviu.2020.103099>

[20]. Li, Q., Xu, T., & Zhang, Y. (2023). Improved facial recognition in surveillance systems with MTCNN and FaceNet. *IEEE Transactions on Image Processing, 32*(4), 2125-2137. https://doi.org/10.1109/TIP.2023.3276549