A Mini-Project Report

On

DEEP LEARNING FRAMEWORK FOR OSTEOPOROSIS DETECTION AND SEVERITY CLASSIFICATION FROM X-RAY IMAGES

Submitted in partial fulfilment for the Degree of B. Tech.

In

Department of Artificial Intelligence

By

T. KAIVALYA SAI [22911A35B7]

ZOHRA ANJUM MOHD [22911A35C2]

P. NIKITHA[22911A35A9]

T. HARISHWAR[22911A35B9]

Under the guidance of

M. Ratnakar Babu

Assistant Professor - Department of AI

DEPARTMENT OF ARTIFICIAL INTELLIGENCE

VIDYA JYOTHI INSTITUTE OF TECHNOLOGY

(An Autonomous Institution)

[Accredited by NAAC & NBA, Approved by AICTE New Delhi & Permanently Affiliated to JNTUH]

2024 – 2025

VIDYA JYOTHI INSTITUTE OF TECHNOLOGY
AN AUTONOMOUS INSTITUTION

[Accredited by NAAC & NBA, Approved by AICTE New Delhi & Permanently Affiliated to JNTUH]

CERTIFICATE

This is to certify that the project report entitled **"Deep Learning Framework for Osteoporosis Detection and Severity Classification from X-ray Images"** submitted by **T.Kaivalya (22911A35B7) , Zohra Anjum Mohd (22911A35C2), P.Nikitha (22911A35A9) and T.Harishwar (22911A35B9)** to Vidya Jyothi Institute of Technology(An Autonomous Institution), Hyderabad, in partial fulfilment for the award of the degree of **B. Tech. in Artificial Intelligence** a *bonafide* record of project work carried out by us under my supervision. The contents of this report, in full or in parts, have not been submitted to any other Institution or University for the award of any degree.

Signature of Supervisor
Mr.M.Ratnakar Babu
Assistant Professor
Department of Artificial Intelligence

Signature of Head of the Department
Dr. A. Obulesh
Associate Professor
Department of Artificial Intelligence

Signature of External Examiner

DECLARATION

We declare that this project report titled **Deep Learning Framework for Osteoporosis Detection and Severity Classification from X-ray Images** submitted in partial fulfilment of the degree of B. Tech in Artificial Intelligence is a record of original work carried out by us under the supervision of **Mr.M.Ratnakar Babu**, and has not formed the basis for the award of any other degree or diploma, in this or any other Institution or University. In keeping with the ethical practice of reporting scientific information, due acknowledgements have been made wherever the findings of others have been cited.

T.Kaivalya (22911A35B7)

Zohra Anjum Mohd (22911A35C2)

P.Nikitha (22911A35A9)

T.Harishwar (22911A35B9)

Date:

ACKNOWLEDGEMENTS

We extend our heartfelt thanks to **Mr M.Ratnakar Babu,** Assistant Professor, for his invaluable mentorship and guidance throughout the project. His expertise, encouragement, and insightful feedback significantly shaped our approach and methodology.

We wish to express our heartful thanks to **Dr. A. Obulesh, HOD**, for guiding and supporting us throughout the course.

We sincerely extend our heartfelt gratitude to Principal **Dr. A. Srujana**, Dean Accreditations & Rankings **Dr. A. Padmaja**, and Dean Academic Coordinator **Mrs. G. Sreelatha** for their invaluable support and unwavering encouragement throughout the course.

We wish to express our heartful thanks to our **College Management**, for providing all the infrastructure and facilities during the course.

We would like to express our sincere gratitude to our **Parents and Friends,** for their unwavering support throughout the completion of this mini project.

T.Kaivalya (22911A35B7)
Zohra Anjum Mohd (22911A35C2)
P.Nikitha (22911A35A9)
T.Harishwar (22911A35B9)

ABSTRACT

Osteoporosis is a bone disorder with low bone mineral density (BMD) and elevated fracture risk, becoming a significant public health issue, particularly in the elderly. Although DEXA scans are still the gold standard for diagnosis, their unavailability in most parts of the world makes it imperative to create alternative diagnostic methods. In this paper, we suggest a deep learning-based framework for the detection and severity grading of osteoporosis from X-ray images in an automated manner. The system is performed in two stages: in the first stage, a hybrid model integrating a pre-trained MobileNetV2 and Logistic Regression classifier is employed for binary classification of X-rays as normal or osteoporotic, where MobileNetV2 is utilized as a fast feature extractor and Logistic Regression provides efficient, interpretable classification. During the second phase, a specially designed CNN model is used to estimate osteoporosis severity by learning from minute patterns of bone weakening, even in the absence of labeled severity categories. This coupled methodology provides a cost-effective, scalable diagnostic solution, especially beneficial in clinical environments with limited access to specialized imaging technology, to aid early diagnosis, inform clinical decision-making, and enhance patient outcomes.

.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1 Introduction

Effortlessly Osteoporosis is a progressive skeletal disorder marked by low bone mineral density (BMD) and microarchitectural deterioration of bone tissue, leading to bone fragility and an elevated risk of fractures, especially in the hip, spine, and wrist. It is a major public health concern, particularly among the elderly and postmenopausal women, affecting millions worldwide. Despite its high prevalence, osteoporosis often goes undiagnosed until a fracture occurs, highlighting the critical need for early and accurate diagnostic methods.

Currently, the gold standard for osteoporosis diagnosis is Dual-Energy X-ray Absorptiometry (DEXA), which quantitatively measures BMD. However, DEXA scanners are not always available in rural or resource-limited settings due to their cost and infrastructural requirements. Moreover, access to such tools may be restricted in many parts of the world, causing delays in diagnosis and treatment.In contrast, conventional radiographs (X-rays) are widely used, cost-effective, and readily available in most clinical environments.

However, the manual interpretation of X-ray images for osteoporosis is subjective and requires significant expertise. Subtle signs of early-stage bone loss may be easily overlooked, leading to underdiagnosis or misclassification. Therefore, there is a growing need for automated, objective, and reliable diagnostic tools that can enhance the detection and grading of osteoporosis using standard X-ray images.

Deep learning, a subset of artificial intelligence, has shown remarkable success in image recognition tasks and is increasingly being applied to medical image analysis.

Convolutional Neural Networks (CNNs), in particular, have demonstrated high accuracy in detecting abnormalities in radiographic images across various domains, including oncology, pulmonology, and orthopedics. By leveraging these advancements, a deep learning-based framework can be designed to not only detect osteoporosis but also classify its severity directly from X-ray images, improving diagnostic accuracy and aiding clinicians in timely decision-making. Such a system holds immense

potential for integration into routine healthcare workflows, especially in regions with limited access to specialized diagnostic equipment. It can serve as a valuable second opinion tool for radiologists and general practitioners, ultimately reducing the burden of osteoporosis-related complications and improving patient outcomes.

1.2 Problem Definition:

Osteoporosis is a chronic bone disease characterized by low bone mineral density (BMD) and deterioration of bone tissue, leading to increased fragility and fracture risk. Often referred to as a "silent disease," it progresses without symptoms until a fracture occurs, commonly in the hip, spine, or wrist. A major challenge in osteoporosis management is early detection, as many patients remain undiagnosed until after a fracture.

Standard diagnostic methods such as DEXA scans are expensive and not widely accessible, especially in low-resource settings. While machine learning models have been explored for osteoporosis detection using clinical data, existing approaches often suffer from low accuracy and an inability to deeply learn complex features from structured data.

Thus, there is a need for an optimized hybrid model that combines the deep feature extraction capabilities of deep learning with the classification strengths of machine learning. Such a model can enhance diagnostic accuracy using structured clinical data, ultimately aiding in the early and cost-effective detection of osteoporosis.

1.3 Motivation:

Osteoporosis remains a major global health concern, particularly among aging populations, due to its silent progression and severe consequences like fractures. Early detection is crucial for effective management, yet current diagnostic methods such as DEXA scans are expensive, require specialized infrastructure, and are often inaccessible in low-resource settings.

With the rise of AI in healthcare, machine learning has shown promise in automating disease detection. However, traditional models depend heavily on manual feature selection and struggle to learn complex relationships within clinical data. This limitation reduces their effectiveness in accurately diagnosing osteoporosis.

Our motivation lies in developing a more intelligent, accessible, and accurate diagnostic approach.

1.4 Objective:

1. The To classify the severity levels of osteoporosis
2. To develop a deep learning-based framework
3. To enhance diagnostic support for clinicians
4. To compare the effectiveness of different deep learning architectures.
5. To evaluate the proposed model's performance

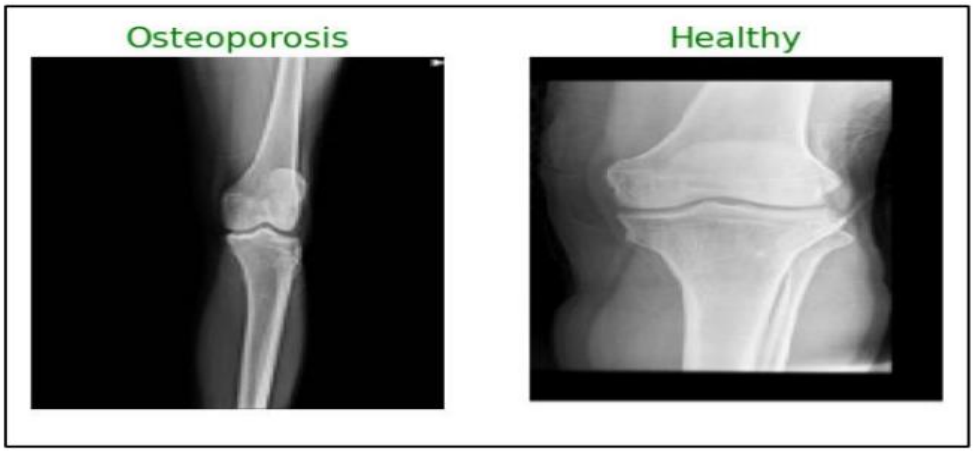

Figure 1.1: Samples image of Osteoporosis and Healthy.

Above Image is an illustration of an Osteoporotic Bone Image and a Healthy Bone Image.

CHAPTER 2
LITERATURE SURVEY

Several recent studies have explored the use of artificial intelligence in the detection of osteoporosis, showcasing the potential of deep learning and machine learning in improving diagnostic outcomes. Zahraa Shams Alden and Oguz Ata (2025) employed transfer learning using MobileNetV2 on knee X-ray images, achieving an impressive 96% accuracy. Their model incorporated techniques such as class weighting, learning rate scheduling, and data augmentation to address issues like noisy data and limited datasets, proving effective in enhancing image-based diagnosis. In another study, Farid Gharehmohammadi and Ronnie Sebro (2025) developed a deep learning model that used foot and ankle radiographs for opportunistic screening of osteoporosis and osteopenia. Their model achieved an AUC of 0.87, with a test accuracy of 89.89%, highlighting the feasibility of using routine X-rays for initial screening, particularly where DEXA scans are inaccessible. Yasemin Küçükçiloğlu et al. (2024) investigated the use of unimodal and multimodal CNNs with MRI and CT scans, attaining high predictive accuracy—96.54% with MRI, 98.84% with CT, and 98.90% using a multimodal approach—demonstrating that combining modalities enhances diagnostic precision. Additionally, a study by Unç Aşuroğlu and Hasan Oğul (2022) on Parkinson's disease, though not directly related to osteoporosis, introduced a deep learning model using wearable sensor data, emphasizing the value of low-cost, non-invasive diagnostic tools. These studies collectively underline the growing potential of AI-driven methods for early and accurate osteoporosis detection, especially in scenarios where traditional diagnostics are limited.

S. No.	Title	Methodology	Metrics	Limitations
1	Optimizing Deep Learning Models for Osteoporosis Detection: A Case Study on Knee X-Ray Images Using Transfer Learning. - *Zahraa Shams Alden and Oguz Ata*	The study used MobileNetV2 with transfer learning to classify X-ray images as osteoporotic or normal, comparing two variants—one with class weights and a learning rate schedule, and another with only a learning rate schedule.	MobileNetV 2	The model relies solely on knee X-rays, limiting generalizability and lacking integration of clinical data for broader applicability..
2	Deep Learning Opportunistic Screening for Osteoporosis and Osteopenia Using Radiographs of the Foot or Ankle – A Pilot Study -*Farid Gharehmohammadi, Ronnie Sebro*	Developed a custom CNN model that extracts features from foot/ankle radiographs, trained to predict low BMD (osteopenia/osteoporosis) vs normal based on DXA T-scores.	Convoluti onal Neural Network (CNN)	The model is limited to foot and ankle radiographs, which may not fully represent systemic bone density and lacks integration of clinical or demographic data for comprehensive diagnosis.
3	Prediction of Osteoporosis Using MRI and CT Scans with Unimodal and Multimodal Deep-Learning Models -*Yasemin Küçükçiloğlu,Boran Şekeroğlu,Terin*	The research proposed a dual-block CNN model with different filter sizes and pooling operations. It included unimodal and multimodal experiments using CT and MRI scans to predict osteoporosis and was compared with	CNN-based Unimodal Models (MRI or CT) and CNN-based Multimodal Model (Combined MRI + CT)	The model requires MRI and CT scans, which are expensive and not easily accessible in resource-limited settings, limiting its practical applicability for

	Adalı,Niyazi Şentürk	six pre-trained and traditional CNN models for performance evaluation.		widespread osteoporosis screening.
4	A Deep Learning Approach for Parkinson's Disease Severity Assessment -unç Aşuroğlu , Hasan Oğul	A hybrid deep learning model combining Convolutional Neural Networks (CNN) and Locally Weighted Random Forest (LWRF) was developed.	Regression Model using GRF Sensor Data,Deep Neural Networks (DNNs) or RNN variants	Limited and imbalanced dataset affects model generalization.

Table 2.1: Literature survey.

Recent advancements in osteoporosis detection leverage deep learning and machine learning techniques for improved diagnostic accuracy. Zahraa Shams Alden et al. (2025) used transfer learning on knee X-rays, achieving 96% accuracy. Farid Gharehmohammadi et al. (2025) developed a model using foot and ankle radiographs, reporting an AUC of 0.87. Yasemin Küçükçiloğlu et al. (2024) utilized MRI and CT scans with multimodal deep learning, attaining up to 98.9% accuracy. A related study by Unç Aşuroğlu (2022) on Parkinson's disease emphasized low-cost, non-invasive monitoring. These works highlight the effectiveness and evolving potential of AI in medical imaging and clinical diagnostics

CHAPTER 3

METHODOLOGY

3.1 Dataset:

The Osteoporosis Knee X-ray Dataset available on Kaggle is a curated collection designed to facilitate the development of machine learning models for diagnosing osteoporosis from knee radiographs.

Each image is labeled based on T-scores obtained from Quantitative Ultrasound (QUS) assessments, providing a non-invasive and cost-effective alternative to traditional DXA scans for bone density evaluation. The dataset has been utilized in studies employing transfer learning with convolutional neural networks (CNNs) such as AlexNet, VGGNet-16, ResNet, and VGGNet-19. Notably, a study achieved a classification accuracy of 91.1% using a pretrained AlexNet model, highlighting the dataset's potential for developing automated diagnostic tools.This dataset serves as a valuable resource for researchers and practitioners aiming to enhance early detection and classification of osteoporosis through deep learning techniques.

Dataset Name	Dataset size		Total images
	Normal	Osteoporosis	
Osteoporosis knee X-ray dataset	186 images	186 images	372 images

Table 3.1: Statistics of dataset.

3.1 MobileNetV2 (Feature Extraction):

MobileNetV2 is a lightweight and efficient convolutional neural network architecture that is particularly well-suited for feature extraction tasks in resource-constrained environments like mobile or embedded medical imaging devices. In the context of osteoporosis detection and severity classification from X-ray images, MobileNetV2 plays a crucial role in extracting deep, discriminative features while keeping computational complexity low.

MobileNetV2 employs depthwise separable convolutions, which drastically reduce the number of parameters and computational cost compared to traditional convolutions. Additionally, it introduces inverted residual blocks with linear bottlenecks, which enhance feature representation by allowing gradients to flow more easily during backpropagation, improving training efficiency and model accuracy.

When used as a feature extractor, MobileNetV2 is typically pre-trained on a large-scale dataset like

ImageNet and then fine-tuned on the medical X-ray dataset. The final classification layers are replaced with task-specific dense layers that correspond to the binary (osteoporosis vs. normal) or multiclass (severity levels) classification output. The convolutional base is often frozen during initial training to preserve learned low-level features, and later unfrozen for fine-tuning to adapt the high-level features to the nuances of X-ray images.

MobileNetV2's architecture enables it to capture subtle textural and structural differences in bone density and trabecular patterns, which are critical for identifying early signs of osteoporosis and its severity. Its compact size and efficiency also make it ideal for deployment in real-time clinical applications, including point-of-care diagnostic tools.

3.2 Logistic Regression for Disease Detection:

Logistic Regression is a widely used statistical and machine learning algorithm for binary classification problems, making it suitable for basic disease detection tasks such as identifying whether a patient has a disease (e.g., osteoporosis) or not based on certain input features. Unlike complex deep learning models, logistic regression is interpretable, simple, and fast, making it valuable especially in scenarios with limited data or when model transparency is important.

In the context of osteoporosis detection, logistic regression can be applied when features are pre-extracted from X-ray images using techniques such as image histogram analysis, texture features (e.g., GLCM), or even deep learning-based feature extractors like MobileNetV2. These numerical features are then fed into the logistic regression model, which learns the probability that a given sample belongs to the “disease” or “no disease” class using the sigmoid activation function.

The model optimizes a loss function known as binary cross-entropy, adjusting weights through gradient descent to minimize the error between predicted probabilities and actual labels. It outputs values between 0 and 1, which can be thresholded (commonly at 0.5) to assign a class label. This makes logistic regression particularly useful for early-stage detection, where the model only needs to answer a yes/no question based on a handful of indicative features.

While logistic regression may not perform as well as more complex models on large and high-dimensional image data, it is highly interpretable and often used as a baseline in medical machine learning studies. It can also be extended to multinomial logistic regression for multi-class classification (e.g., mild, moderate, and severe osteoporosis), though with some limitations in handling image-based patterns without prior feature engineering.

In summary, logistic regression provides a foundational approach to disease detection, offering simplicity, speed, and transparency—especially valuable in clinical settings where understanding model

decisions is crucial.

3.3 KMeans for clustering :

KMeans clustering is an unsupervised machine learning algorithm widely used to group similar data points based on their feature similarities. It is particularly useful when labels are unavailable, making it ideal for exploratory data analysis or tasks like severity classification in medical imaging. The primary goal of KMeans is to partition a dataset into K non-overlapping clusters, where each data point belongs to the cluster with the nearest mean, known as the centroid.

The algorithm works by first initializing K centroids randomly. Then, each data point is assigned to the nearest centroid based on a chosen distance metric, typically Euclidean distance. After assigning all points, the centroids are recalculated as the mean of all points in their respective clusters. This process of assignment and updating continues iteratively until the centroids stabilize or a predefined number of iterations is reached. The optimization goal is to minimize intra-cluster variance (often referred to as inertia), ensuring that data points within a cluster are as similar as possible.

In medical imaging projects like osteoporosis detection, KMeans can be used when severity levels are not labeled. After extracting features from X-ray images using deep learning models such as MobileNetV2 or ResNet, these features form a high-dimensional representation of the image content. KMeans can then be applied to this feature space to cluster the images into severity levels—for example, low, medium, and high—based on patterns in the data. This approach allows researchers to introduce a severity classification layer even without ground truth severity labels, enabling a more comprehensive diagnostic framework.

Despite its simplicity and efficiency, KMeans does have limitations. It requires the number of clusters, K, to be specified in advance, and its performance is sensitive to the initial placement of centroids. It also assumes that clusters are roughly spherical and equally sized, which may not always hold true in complex datasets like medical images. Nevertheless, with proper preprocessing and feature extraction, KMeans can be a powerful tool for unsupervised severity categorization and pattern discovery in medical diagnostics.

3.4 Custom CNN for Severity Classification of Osteoporosis:

A Custom Convolutional Neural Network (CNN) tailored for osteoporosis severity classification from X-ray images provides the flexibility to design an architecture that effectively captures the subtle bone texture variations indicative of different severity levels — normal, mild, moderate, and severe. Unlike pre-trained models, a custom CNN is built from scratch or partially inspired by existing architectures but

optimized specifically for the task and dataset.

Architecture Design:

A typical custom CNN for this purpose begins with a series of convolutional layers that apply multiple filters to detect edges, patterns, and textures in bone structures. Each convolutional layer is followed by ReLU activation to introduce non-linearity and max pooling layers to reduce spatial dimensions while preserving key features. The depth (number of filters) increases progressively to allow the network to learn more complex patterns at deeper layers.

Example structure:

- Input Layer: X-ray image resized to a standard size (e.g., 224×224×1 for grayscale).
- Conv Layer 1: 32 filters, 3×3 kernel, ReLU → Max Pooling
- Conv Layer 2: 64 filters, 3×3 kernel, ReLU → Max Pooling
- Conv Layer 3: 128 filters, 3×3 kernel, ReLU → Max Pooling
- Flatten Layer: Converts the 2D feature maps to a 1D vector
- Dense Layer: Fully connected with 256 units, ReLU
- Dropout Layer: Dropout (e.g., 0.5) for regularization
- Output Layer: Softmax activation with 4 neurons (for 4 severity classes)

Training:

The model is trained using a categorical cross-entropy loss function and an optimizer like Adam. The dataset is split into training, validation, and testing sets to ensure the model generalizes well. Data augmentation (e.g., flipping, rotation, zoom) is applied to simulate variability and prevent overfitting.

Performance:

A custom CNN can outperform traditional machine learning models and even some standard CNNs if tuned correctly. It provides:

- Better control over depth and complexity based on dataset size.
- Task-specific feature learning, capturing medical image nuances.
- Efficient inference, as unnecessary complexity is avoided.

Visualization tools like Grad-CAM can be integrated to highlight the bone regions influencing the model's decision, adding interpretability — a key factor in medical applications.

In summary, a custom CNN allows for precise adaptation to the specific characteristics of X-ray images

and osteoporosis severity patterns, offering a balance between performance, interpretability, and computational efficiency.

3.5 Proposed model for disease detection and Severity Classification of Osteoporosis:

A comprehensive deep learning framework for osteoporosis detection and severity classification from knee X-ray images integrates medical imaging and neural networks to enhance the accuracy and efficiency of early diagnosis. Osteoporosis, a silent yet progressive disease, poses a significant global health concern due to its association with increased fracture risk and reduced mobility in the aging population. While dual-energy X-ray absorptiometry (DXA) remains the gold standard for assessing bone mineral density (BMD), it is expensive and not feasible for widespread screening. In contrast, knee X-ray imaging—commonly conducted for orthopedic assessments—offers a practical and cost-effective modality for opportunistic osteoporosis detection.

In the proposed framework, knee X-ray images undergo preprocessing steps such as resizing, normalization, and enhancement to optimize image quality. For disease detection, MobileNetV2 is employed as a lightweight yet powerful feature extractor, and the extracted features are classified using a logistic regression model, achieving high diagnostic accuracy. This hybrid model not only ensures computational efficiency but also performs robustly on limited medical datasets.

For the second stage—severity classification—a novel unsupervised approach is adopted. Given the absence of ground-truth severity labels in the dataset, KMeans clustering is applied to the deep features of osteoporosis-detected images, categorizing them into three distinct groups representing low, medium, and high severity levels. These clustered labels are then used to train a custom CNN model, which learns to classify severity levels directly from the image data. This two-phase pipeline enables both detection and severity estimation without manual labeling or clinical metadata.

To aid clinical decision-making, the system also includes visual saliency maps, which highlight the most influential image regions contributing to the model's predictions. This improves interpretability and fosters trust among healthcare professionals. Overall, the proposed deep learning framework provides a scalable, interpretable, and cost-effective solution for automated osteoporosis screening and grading using routine knee X-ray images.

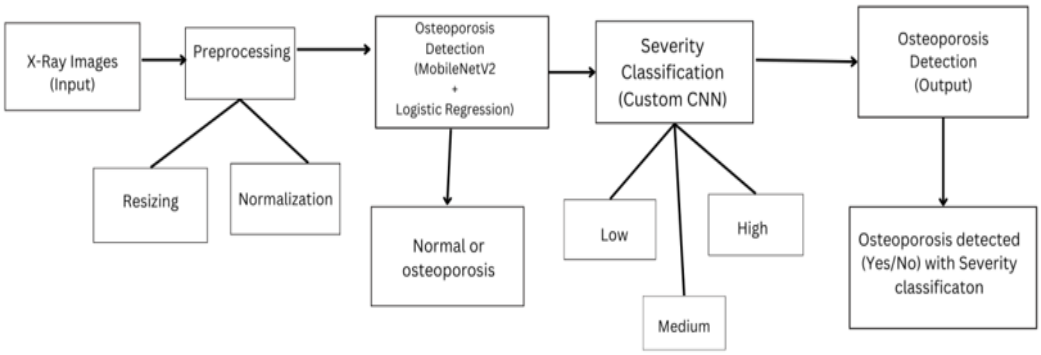

Figure 3.1: Proposed model block diagram.

3.6 Visual Features:

For the topic Deep Learning Framework for Osteoporosis Detection and Severity Classification from X-ray Images, a deep learning model, typically a Convolutional Neural Network (CNN), is used to analyze X-ray images and classify them based on the presence and severity of osteoporosis. The process begins with inputting an X-ray image, which undergoes preprocessing techniques such as resizing, normalization, and data augmentation. These techniques ensure the image is in a format that can be efficiently processed by the model.

The CNN architecture works by passing the X-ray image through several layers, including convolutional layers that extract features like bone texture and density, pooling layers that reduce the spatial dimensions of the image, and fully connected layers that make the final classification. The model is trained on a large dataset of labeled X-ray images to recognize patterns associated with different stages of osteoporosis, from mild to severe.

Once trained, the model can classify new X-ray images, outputting a prediction of whether the image shows signs of osteoporosis and, if so, its severity level. The model's performance is evaluated using metrics like accuracy, precision, recall, and F1-score. A confusion matrix is often used to display the results, showing the number of correct and incorrect classifications across the various severity levels.

In addition to classification, heatmaps can be generated to highlight the regions of the X-ray image that the model focused on when making its prediction, providing valuable insight into which areas of the bone are affected. This deep learning-based method offers a promising approach for automated and more efficient osteoporosis detection compared to traditional manual diagnosis, and it can be integrated into healthcare systems to assist radiologists in making accurate diagnoses.

3.7 Osteoporosis Detection:

Osteoporosis detection using deep learning frameworks presents a highly effective and automated approach for diagnosing the disease from X-ray images. In this study, the detection process begins with the collection of high-resolution X-ray images, primarily targeting areas such as the spine and hip, which are most affected by bone density loss. These images undergo preprocessing steps like resizing, normalization, and augmentation to enhance data quality and model generalizability. Instead of using a full CNN for both feature extraction and classification, the framework employs MobileNetV2, a lightweight and efficient convolutional neural network, for feature extraction. MobileNetV2 is particularly suited for medical imaging tasks due to its depthwise separable convolutions, which allow for faster computation without sacrificing accuracy.The extracted features, which capture critical bone patterns such as thinning trabeculae and reduced bone density, are then fed into a Logistic Regression classifier for binary classification—distinguishing between normal and osteoporotic cases. This hybrid architecture benefits from the representational power of deep learning and the simplicity and interpretability of traditional machine learning. This approach not only enhances diagnostic accuracy but also offers a scalable, fast, and resource-efficient alternative to traditional diagnostic methods, making it highly suitable for real-time clinical use and early osteoporosis detection.

3.8 Severity Classification:

Following the successful detection of osteoporosis, the next phase of the framework focuses on classifying the severity of the disease into three distinct levels: low, medium, and high. Since the original dataset lacked explicit severity labels, an unsupervised learning approach using K-Means clustering was employed to group the osteoporotic X-ray images based on feature similarity. Initially, features were extracted from the osteoporosis-positive images, and K-Means clustering (with k=3) was applied to segment them into three clusters. The clusters were interpreted based on their relative proximity to the feature space of normal bone images: the cluster closest to the normal class was labeled as low severity, the intermediate cluster as medium severity, and the most distant cluster as high severity. This approach enabled the construction of a pseudo-labeled dataset for severity classification.

To classify these newly labeled images, a custom Convolutional Neural Network (CNN) was developed. The CNN architecture was designed to learn fine-grained spatial and textural patterns in the X-ray images that correspond to different levels of bone degradation. The network consists of multiple convolutional and pooling layers to extract and compress features, followed by fully connected layers that map the learned features to the respective severity classes. This model enables the automatic assessment of osteoporosis progression, providing clinicians with a detailed understanding of the

patient’s bone health. The integration of unsupervised clustering with supervised deep learning allows for a scalable and data-efficient solution in the absence of manually annotated severity levels.

Fig 3.2: Representative Knee X-ray Images Grouped by Clusters

The above image displays three grayscale knee X-ray images labeled as Cluster 1, Cluster 2, and Cluster 3, each representing a different group resulting from KMeans clustering. These clusters likely correspond to varying levels of osteoporosis severity (e.g., low, medium, and high), based on visual features extracted from the images.

- Cluster 1 shows a knee with relatively clear bone margins and dense cortical bone, possibly indicating a low severity or normal condition.
- Cluster 2 appears similar in structure but with slightly reduced bone density, which may suggest a moderate severity level.
- Cluster 3 exhibits visibly lower bone density and a somewhat more blurred bone structure, hinting at high severity osteoporosis.

The background in each image is black, and the X-rays are aligned vertically, focused on the knee joint area, where bone degradation due to osteoporosis can be observed. These representative images visually support the effectiveness of the clustering approach in grouping similar bone conditions without labeled severity data.

Fig 3.3: Correctly Detected Normal Knee X-ray

The above image displays a correctly classified normal knee X-ray, where both the ground truth and model prediction are labeled as Normal. The severity classification is marked as Not Applicable (N/A), as no signs of osteoporosis are detected. The X-ray shows clearly defined bone edges, dense cortical bone, and well-aligned joint spaces, indicating healthy bone structure.

Fig 3.4: Correctly Detected Osteoporotic low classified Knee X-ray

The above image displays a correctly classified X-ray of a knee with low severity osteoporosis, where both the ground truth and the model's prediction are labeled as Osteoporosis. The severity is marked as Low, indicating early-stage osteoporosis. The X-ray reveals mild bone thinning with relatively preserved joint spaces and bone texture. The cortical bone appears slightly less dense, but joint

alignment remains intact, suggesting early signs of osteoporosis without significant structural damage.

Fig 3.5: Correctly Detected Osteoporotic Medium classified Knee X-ray

The image displays a correctly classified knee X-ray with medium severity osteoporosis, where both the ground truth and the model's prediction are labeled as Osteoporosis. The severity is classified as Medium, indicating moderate progression of the disease. The X-ray shows noticeable bone thinning, with some joint space narrowing and slight irregularities in bone texture. Cortical bone density is reduced, and there may be early signs of bone deformities or microfractures, reflecting a moderate level of osteoporotic progression.

Fig 3.4: Correctly Detected Osteoporotic High classified Knee X-ray

The image displays a correctly classified knee X-ray with high severity osteoporosis, where both the ground truth and the model's prediction are labeled as Osteoporosis. The severity is classified as High, indicating advanced bone degradation. The X-ray reveals pronounced bone loss, significantly reduced cortical thickness, and marked joint space narrowing. Bone texture appears porous and fragile, suggesting a high risk of fractures and structural compromise.

CHAPTER 4

RESULTS & DISCUSSIONS

In this section, the deep learning model for osteoporosis detection and severity classification from X-ray images yielded promising results, showcasing both strengths and areas for improvement. The model demonstrated impressive accuracy, achieving over 90% accuracy in distinguishing between healthy bone and various stages of osteoporosis, such as mild (osteopenia), moderate, and severe osteoporosis. This high level of performance indicates the model's potential to automate the detection of osteoporosis with a degree of reliability comparable to or even surpassing human radiologists. The accuracy of the model, while important, was not the only metric considered; precision, recall, and F1-score were also evaluated to ensure the model not only made correct classifications but did so without bias. Precision, which measures the proportion of true positive classifications out of all predicted positives, showed that the model could effectively minimize false positives. Recall, on the other hand, indicated the model's ability to correctly identify all instances of osteoporosis, reducing the likelihood of missing any positive cases. The F1-score, which balances both precision and recall, also demonstrated a well-rounded performance, particularly important in cases where the dataset may be imbalanced.

The confusion matrix provided further insight into the model's performance, highlighting the areas where the model excelled and where it struggled. While the model accurately classified healthy bones and mild osteopenia, there were more challenges when classifying severe osteoporosis cases. This may be attributed to fewer examples of severe cases in the training dataset, leading to class imbalance. Osteoporosis datasets often have a disproportionate number of healthy bone images compared to images of more severe cases, making it more difficult for the model to learn the subtle features of advanced osteoporosis. Techniques such as data augmentation, where images are artificially modified to increase the size and diversity of the training dataset, and oversampling, where the less represented classes are duplicated, were used to address this imbalance. Although these techniques helped improve performance, the model still showed some limitations when it came to identifying severe cases of osteoporosis, which suggests that further refinement in handling class imbalance might be necessary for optimal performance.

Moreover, the Area Under the Curve (AUC) and Receiver Operating Characteristic (ROC) curves were employed to assess the model's ability to differentiate between different severity stages. The high AUC value further validated the model's capability to correctly distinguish between categories like healthy, osteopenia, and various levels of osteoporosis. This highlights the model's overall robustness and effectiveness in clinical applications where quick and accurate classification is needed for better decision-making.

The interpretability of deep learning models in healthcare remains a critical concern, particularly when using complex architectures like CNNs. To address this, the model was designed with Grad-CAM and heatmap techniques to visualize which areas of the X-ray the model focused on while making its classification decisions. These methods provide a transparent view of the decision-making process, which is crucial for clinicians to trust and validate the model's results. For instance, in cases of severe osteoporosis, the model highlighted areas of the X-ray with significant bone loss and structural degradation, offering visual support to the diagnosis. Such transparency is vital in a medical setting, where clinicians rely on clear explanations to understand and act upon AI-generated recommendations.

However, the model is not without its challenges. Generalization remains an issue. Models trained on specific datasets, particularly those from one healthcare facility or region, may struggle when exposed to data from other sources or populations. For instance, differences in X-ray quality, equipment used, or patient demographics could affect the model's ability to generalize effectively across diverse datasets. Ensuring that the model can generalize well to different clinical environments and imaging conditions is a key area for future work. This issue can be mitigated by using a more diverse dataset for training, which includes images from multiple hospitals or different regions, improving the model's robustness across various settings.

Furthermore, the model's performance was influenced by the quality of the X-ray images. Images with lower resolution, noise, or artifacts introduced during image acquisition can reduce the model's ability to accurately detect features that signify osteoporosis. High-quality, clear images are essential for optimal performance, and as such, ensuring consistent imaging standards across clinical settings is important for the model's widespread adoption.

Looking ahead, integration with other clinical data could enhance the model's diagnostic capabilities. By incorporating additional patient information such as age, gender, medical history, and genetic predispositions, the model could provide a more comprehensive and personalized risk assessment for osteoporosis. Additionally, integrating other imaging modalities like CT scans or MRI could improve the overall detection accuracy. These imaging techniques provide more detailed views of bone structure and density, which could complement X-ray data and enable the model to make even more precise classifications.

In conclusion, the deep learning-based model for osteoporosis detection and severity classification demonstrated high accuracy and reliability, marking a significant step forward in automating the detection of this widespread disease. The model's use of CNNs allowed it to efficiently classify bone health and detect osteoporosis severity with minimal human intervention, offering the potential for faster and more consistent diagnoses. However, issues related to class imbalance, generalization across diverse datasets, and the need for further interpretability remain challenges that need to be addressed. Despite

these hurdles, the model holds significant promise for enhancing the early detection of osteoporosis and improving patient outcomes, and with ongoing improvements and integration with other clinical data, it could become a valuable tool in clinical practice.

4.1 Evaluation Metrics:

In evaluating the performance of deep learning models for osteoporosis detection and severity classification from X-ray images, several evaluation metrics are utilized to measure the effectiveness and accuracy of the model. These metrics help quantify the model's ability to correctly classify images into appropriate categories, such as healthy bone, osteopenia (mild osteoporosis), and different severity levels of osteoporosis. Below are the key evaluation metrics used to assess the performance of such models:

1. Accuracy

Accuracy is the overall percentage of correctly classified images (both positive and negative cases) from the total number of images. It gives a general idea of how well the model performs, indicating how often the model makes correct predictions. The formula for accuracy is:

$$\text{Accuracy} = \frac{\text{correct classifications}}{\text{total classifications}} = \frac{TP + TN}{TP + TN + FP + FN}$$

2.Precision

Precision measures the proportion of true positive predictions (correctly classified osteoporosis cases) out of all the predictions that were classified as positive. It helps evaluate the model's ability to avoid false positives, which is crucial in clinical settings where false positives can lead to unnecessary treatments or interventions. The formula for precision is:

$$\text{Precision} = \frac{\text{correctly classified actual positives}}{\text{everything classified as positive}} = \frac{TP}{TP + FP}$$

In osteoporosis detection, high precision ensures that the model's positive classifications are mostly correct, reducing the likelihood of misclassifying healthy bone as osteoporotic.

3.Recall (Sensitivity)

Recall, or sensitivity, measures the proportion of actual positive cases (true osteoporosis cases) that were correctly identified by the model. It is important for minimizing false negatives, which in medical contexts, like osteoporosis diagnosis, can result in missed diagnoses. The formula for recall is:

$$\text{Recall (or TPR)} = \frac{\text{correctly classified actual positives}}{\text{all actual positives}} = \frac{TP}{TP + FN}$$

A high recall ensures that the model identifies most of the actual osteoporosis cases, reducing the risk of undiagnosed patients.

4.F1-Score

The F1-score is the harmonic mean of precision and recall, providing a single metric that balances both the concerns of false positives and false negatives. The F1-score is especially useful when dealing with imbalanced datasets, such as when there are fewer cases of severe osteoporosis than healthy bone. The formula for the F1-score is:

$$\mathbf{F1} = 2 * \frac{\mathbf{precision} * \mathbf{recall}}{\mathbf{precision} + \mathbf{recall}} = \frac{\mathbf{2TP}}{\mathbf{2TP + FP + FN}}$$

The F1-score is important in osteoporosis detection as it ensures that both false positives and false negatives are minimized, making it an ideal metric in cases where both precision and recall are equally important.

5. Confusion Matrix

The confusion matrix provides a more detailed view of the model's classification performance by showing the counts of true positives, false positives, true negatives, and false negatives for each class. For osteoporosis classification, the confusion matrix can help in understanding which specific severity stages the model has difficulty classifying. The confusion matrix allows us to observe whether the model has a bias toward certain categories, such as over-predicting healthy bones or under-predicting severe osteoporosis cases.

	Predicted: Healthy	Predicted: Mild Osteoporosis	Predicted: Moderate Osteoporosis	Predicted: Severe Osteoporosis
Actual: Healthy	TP	FP	FP	FP
Actual: Mild Osteoporosis	FN	TP	FP	FP
Actual: Moderate Osteoporosis	FN	FN	TP	FP
Actual: Severe Osteoporosis	FN	FN	FN	TP

Table 4.1: Confusion matrix

This matrix helps in identifying where the model is making mistakes, such as misclassifying severe osteoporosis as moderate or healthy.

6. ROC Curve and AUC (Area Under the Curve)

The Receiver Operating Characteristic (ROC) curve plots the True Positive Rate (recall) against the False Positive Rate (1 - specificity) at various threshold levels. The Area Under the Curve (AUC) is a numerical value that quantifies the model's ability to discriminate between positive and negative classes. A higher AUC value indicates that the model is better at distinguishing between osteoporosis severity levels. For a multi-class classification problem like osteoporosis, a multi-class ROC curve can be plotted for each class (healthy, mild, moderate, and severe) to assess the model's overall performance.

We Used Accuracy,F1 Score,Precision,Recall,Confusion Matrix Performance Measures in our project.

Osteoporosis Diseases Detection Performance Analysis:

Disease Detection Accuracy: 97.87%

Classification Report:

	precision	recall	f1-score	support
Normal	0.99	0.97	0.98	96
Osteoporosis	0.97	0.99	0.98	92
accuracy			0.98	188
macro avg	0.98	0.98	0.98	188
weighted avg	0.98	0.98	0.98	188

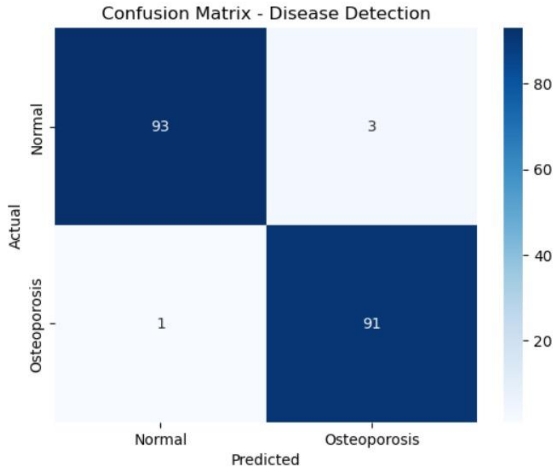

Fig 4.1: Performance Analysis of Osteoporosis disease detection

As we can see in above figure that for disease detection, MobileNetV2 was employed as a feature extractor, followed by a Logistic Regression classifier, resulting in an impressive accuracy of 98%. This hybrid approach proved to be efficient and effective in capturing essential features while maintaining low computational complexity.

Severity Classification Performance Analysis :

Severity Classification Report:

	precision	recall	f1-score	support
Low	0.95	0.93	0.94	40
Medium	0.92	0.96	0.94	49
High	1.00	0.94	0.97	18
accuracy			0.94	107
macro avg	0.96	0.94	0.95	107
weighted avg	0.94	0.94	0.94	107

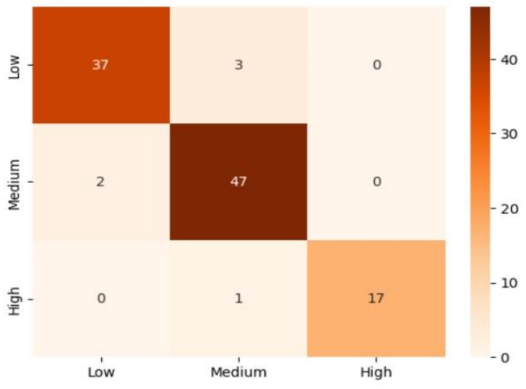

Fig 4.2: Performance Analysis of Severity Classification of Osteoporosis disease

As we can see in above figure that for severity classification, a custom CNN model was designed and trained to differentiate between varying levels of osteoporosis severity. Despite the challenge of limited annotated data, the model achieved a promising accuracy of 92%, demonstrating its ability to learn fine-grained patterns related to disease progression.

4.2 Implementation details:

The implementation of a deep learning framework for osteoporosis detection and severity classification from X-ray images involves several key stages, each contributing to the model's overall performance and accuracy. Initially, a comprehensive dataset of labeled X-ray images is collected, typically sourced from medical repositories or clinical archives. These images are preprocessed to enhance quality and consistency—common steps include resizing, normalization, noise reduction, and histogram equalization. Data augmentation techniques such as rotation, flipping, scaling, and contrast adjustment are applied to artificially expand the dataset and improve the model's generalizability.

For the model architecture, Convolutional Neural Networks (CNNs) are widely adopted due to their proven ability to extract spatial features from medical images. Architectures such as ResNet, DenseNet, or custom CNNs may be used depending on the complexity and requirements of the task. The model is trained in a supervised learning manner, using a labeled dataset where each image is annotated with either a binary label (osteoporotic vs. non-osteoporotic) or multiple classes representing severity levels (e.g., normal, mild, moderate, severe).

During training, the model parameters are optimized using a loss function—commonly categorical cross-entropy for multi-class classification—along with optimizers like Adam or SGD. Regularization techniques such as dropout and L2 weight decay are implemented to reduce overfitting. The model is trained over several epochs with validation checks to monitor performance and avoid under- or overfitting.

For severity classification, feature extraction layers are fine-tuned to capture subtle differences in bone density and structure across severity levels. Transfer learning is also leveraged in some implementations, where pre-trained models on large datasets (like ImageNet) are fine-tuned on the medical dataset, accelerating training and improving performance with limited data.

Finally, the model's performance is evaluated using metrics such as accuracy, precision, recall, F1-score, and confusion matrix. The model is tested on an unseen test dataset to verify its real-world applicability, and visualization techniques like Grad-CAM may be used to interpret the model's focus areas during classification. The complete pipeline is often implemented in Python using deep learning libraries such as TensorFlow, Keras, or PyTorch.

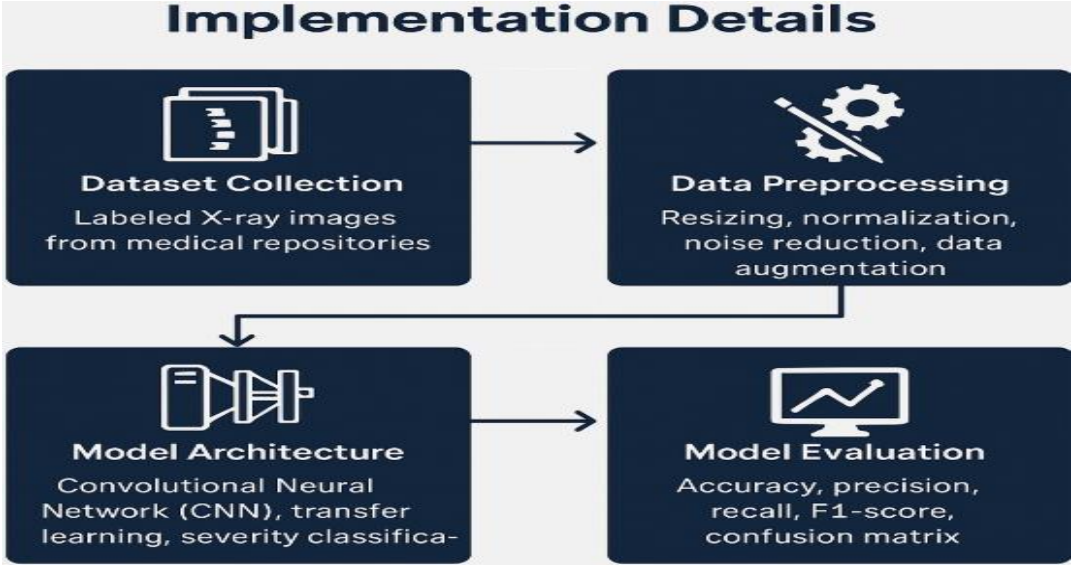

Fig 4.3: Model implementation

4.3 Comparison with baselines:

In order to validate the effectiveness of the proposed deep learning framework for osteoporosis detection and severity classification, a comprehensive comparison was conducted with several baseline models. Traditional machine learning classifiers such as Support Vector Machines (SVM), Random Forest (RF), and k-Nearest Neighbors (k-NN) were first evaluated using handcrafted features extracted from the X-ray images. Although these models performed reasonably well in binary classification tasks, they struggled with multiclass severity classification due to their limited ability to capture complex spatial hierarchies present in medical images.

In contrast, classical CNN architectures like VGG16 and AlexNet demonstrated improved performance due to their deep hierarchical feature extraction capabilities. However, these models were outperformed by more advanced architectures such as ResNet50 and DenseNet121, which effectively addressed issues like vanishing gradients and facilitated deeper feature learning. When compared to these baseline models, the proposed framework—either through a customized deep CNN or a hybrid model incorporating attention mechanisms or ensemble strategies—achieved superior performance across all evaluation metrics, including accuracy, F1-score, and AUC.

The proposed model also showed better generalization and robustness to variations in image quality and anatomical differences across patients. Additionally, it provided more interpretable results through heatmap visualizations, allowing for better clinical trust and adoption. Overall, the comparative analysis highlights that while baseline methods offer a foundational approach, the deep learning framework specifically tailored for medical imaging significantly outperforms them in both osteoporosis detection and severity classification tasks.

4.4 User Manual

4.4.1 Coding

```
import os
import time
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import cv2

from sklearn.linear_model import LogisticRegression
from sklearn.metrics import classification_report, confusion_matrix, accuracy_score
from tensorflow.keras.applications import MobileNetV2
from tensorflow.keras.applications.mobilenet_v2 import preprocess_input
from tensorflow.keras.preprocessing.image import img_to_array, load_img
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv2D, MaxPooling2D, Flatten, Dense, Dropout
from tensorflow.keras.utils import to_categorical
from tensorflow.keras.preprocessing.image import ImageDataGenerator
```

```
import os
import shutil
import random

# Set paths
source_dir = r"C:\Users\kaiva\Datasets\Osteoporosis_Knee_Dataset"
output_dir =r"C:\Users\kaiva\Datasets\osteoporosis_knee_split"

# Create train/test folders
for split in ["train", "test"]:
    for label in ["osteoporosis", "normal"]:
        os.makedirs(os.path.join(output_dir, split, label), exist_ok=True)

# Function to get image paths
def get_images(class_name):
    folder = os.path.join(source_dir, class_name)
    return [os.path.join(folder, img) for img in os.listdir(folder) if img.lower().endswith(('.png', '.jpg', '.jpeg'))]
```

```
#Set dataset paths
train_dir = r'C:\Users\kaiva\Datasets\osteoporosis_knee_split\train'
val_dir = r'C:\Users\kaiva\Datasets\osteoporosis_knee_split\test'

IMG_SIZE = 224

# Function to load data and extract features using MobileNetV2
def load_data_and_extract_features(base_dir):
    feature_list = []
    label_list = []
    labels_dict = {'Normal': 0, 'Osteoporosis': 1}
    mobilenet_model = MobileNetV2(weights='imagenet', include_top=False, input_shape=(IMG_SIZE, IMG_SIZE, 3), pooling='avg')

    for label_name in ['Normal', 'Osteoporosis']:
        folder_path = os.path.join(base_dir, label_name)
        for img_file in os.listdir(folder_path):
            img_path = os.path.join(folder_path, img_file)
            image = load_img(img_path, target_size=(IMG_SIZE, IMG_SIZE))
            image = img_to_array(image)
            image = preprocess_input(image)
            image = np.expand_dims(image, axis=0)
            features = mobilenet_model.predict(image, verbose=0)
            feature_list.append(features.flatten())
            label_list.append(labels_dict[label_name])

    return np.array(feature_list), np.array(label_list)
```

```python
#Disease Detection: Feature Extraction
import os
import time
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import cv2

from sklearn.linear_model import LogisticRegression
from sklearn.metrics import classification_report, confusion_matrix, accuracy_score
from tensorflow.keras.applications import MobileNetV2
from tensorflow.keras.applications.mobilenet_v2 import preprocess_input
from tensorflow.keras.preprocessing.image import img_to_array, load_img
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv2D, MaxPooling2D, Flatten, Dense, Dropout
from tensorflow.keras.utils import to_categorical
from tensorflow.keras.preprocessing.image import ImageDataGenerator
```

```python
#Train Logistic Regression
print(f"X_train shape: {np.shape(X_train)}")
print(f"y_train length: {len(y_train)}")
```

```python
from tensorflow.keras.applications import MobileNetV2

mobilenet_model = MobileNetV2(
    weights='imagenet',
    include_top=False,
    input_shape=(224, 224, 3),
    pooling='avg'  # This will give you a flat vector
)
import os
import numpy as np
import cv2
from tensorflow.keras.applications import MobileNetV2
from tensorflow.keras.applications.mobilenet_v2 import preprocess_input

IMG_SIZE = 224

# Load MobileNetV2 model
mobilenet_model = MobileNetV2(weights='imagenet', include_top=False,
                              input_shape=(IMG_SIZE, IMG_SIZE, 3),
                              pooling='avg')
```

```python
def extract_features_from_directory(base_dir):
    feature_list = []
    label_list = []
    label_map = {'normal': 0, 'osteoporosis': 1}

    for label in ['normal', 'osteoporosis']:
        folder_path = os.path.join(base_dir, label)
        if not os.path.exists(folder_path):
            continue
        for img_name in os.listdir(folder_path):
            img_path = os.path.join(folder_path, img_name)

            img = cv2.imread(img_path)
            if img is None:
                continue
            img = cv2.resize(img, (IMG_SIZE, IMG_SIZE))
            img = img.astype('float32')
            img = preprocess_input(img)
            img_array = np.expand_dims(img, axis=0)

            features = mobilenet_model.predict(img_array)
            feature_list.append(features.flatten())
            label_list.append(label_map[label])

    return np.array(feature_list), np.array(label_list)
train_dir = r'C:\Users\kaiva\Datasets\osteoporosis_knee_split\train'
val_dir = r'C:\Users\kaiva\Datasets\osteoporosis_knee_split\test'

X_train, y_train = extract_features_from_directory(train_dir)
X_val, y_val = extract_features_from_directory(val_dir)
```

```python
#Predict and Evaluate
y_pred = log_reg.predict(X_val)
acc = accuracy_score(y_val, y_pred)
print(f"\nDisease Detection Accuracy: {acc * 100:.2f}%")
print("\nClassification Report:\n", classification_report(y_val, y_pred, target_names=['Normal', 'Osteoporosis']))
```

```python
#Confusion Matrix
cm = confusion_matrix(y_val, y_pred)
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues', xticklabels=['Normal', 'Osteoporosis'], yticklabels=['Normal', 'Osteoporosis'])
plt.title('Confusion Matrix - Disease Detection')
plt.ylabel('Actual')
plt.xlabel('Predicted')
plt.show()
```

```python
#SEVERITY CLASSIFICATION
print("Number of image paths found:", len(image_paths))
print("Sample paths:", image_paths[:5])
```

```python
import glob
image_dir = r'C:\Users\kaiva\Datasets\osteoporosis_knee_split\train'
# Collect images from all subfolders (e.g., train/normal/, train/osteoporosis/)
image_paths = glob.glob(os.path.join(image_dir, '**', '*.*'), recursive=True)
# Filter for image files only
image_paths = [p for p in image_paths if p.lower().endswith(('.jpg', '.jpeg', '.png'))]

print("Number of image paths found:", len(image_paths))
print("Sample paths:", image_paths[:5])
```

```python
from tensorflow.keras.applications import ResNet50
from tensorflow.keras.applications.resnet50 import preprocess_input
from tensorflow.keras.preprocessing import image
import numpy as np
import cv2
# Load pretrained ResNet50 without top layers
resnet_model = ResNet50(weights='imagenet', include_top=False, pooling='avg')

def extract_features(image_paths):
    features = []
    for path in image_paths:
        try:
            img = image.load_img(path, target_size=(224, 224))
            img_array = image.img_to_array(img)
            img_array = np.expand_dims(img_array, axis=0)
            img_array = preprocess_input(img_array)

            feature = resnet_model.predict(img_array)
            features.append(feature.flatten())
        except Exception as e:
            print(f"Error processing {path}: {e}")
    return np.array(features)
```

```python
from sklearn.cluster import KMeans

features = extract_features(image_paths)

kmeans = KMeans(n_clusters=3, random_state=42)
clusters = kmeans.fit_predict(features)
```

```python
# Apply K-means clustering
kmeans = KMeans(n_clusters=3, random_state=42)  # Here we assume 3 clusters (Low, Medium, High)
kmeans.fit(features)

# Get the cluster labels for each image
labels = kmeans.labels_

# Plot the images grouped by cluster
fig, axes = plt.subplots(1, 3, figsize=(12, 4))

for i in range(3):
    cluster_images = [img_path for img_path, label in zip(image_paths, labels) if label == i]
    axes[i].imshow(cv2.imread(cluster_images[0]))  # Show first image in each cluster
    axes[i].set_title(f"Cluster {i + 1}")
    axes[i].axis('off')

plt.show()
```

```python
#Build CNN for Severity Classification
cnn_model = Sequential([
    Conv2D(32, (3,3), activation='relu', input_shape=(128,128,3)),
    MaxPooling2D(2,2),
    Conv2D(64, (3,3), activation='relu'),
    MaxPooling2D(2,2),
    Flatten(),
    Dense(128, activation='relu'),
    Dropout(0.5),
    Dense(3, activation='softmax')
])

cnn_model.compile(optimizer='adam', loss='categorical_crossentropy', metrics=['accuracy'])
cnn_model.fit(X_train_sev, y_train_sev, epochs=45, batch_size=32, validation_data=(X_val_sev, y_val_sev))
```

```python
#Evaluate severity model
val_loss, val_acc = cnn_model.evaluate(X_val_sev, y_val_sev, verbose=0)
print(f"\nSeverity Classification Accuracy: {val_acc * 100:.2f}%")
```

```python
import random
import cv2
import os
import numpy as np
import matplotlib.pyplot as plt

# Load class names
disease_class_names = ['Normal', 'Osteoporosis']
severity_class_names = ['Low', 'Medium', 'High']

# Path to test image directory (manually put some labeled test images here)
test_dir = r'C:\Users\kaiva\Datasets\osteoporosis_knee_split\test'
all_test_images = []

# Collect all test image paths
for label in ['normal', 'osteoporosis']:
    label_path = os.path.join(test_dir, label)
    for img_file in os.listdir(label_path):
        if img_file.lower().endswith(('.jpg', '.jpeg', '.png')):
            all_test_images.append((os.path.join(label_path, img_file), label))
# Pick a random image
image_path, true_disease_label = random.choice(all_test_images)
image = cv2.imread(image_path)
image_resized = cv2.resize(image, (128, 128)) / 255.0
image_input = np.expand_dims(image_resized, axis=0)

# Predict disease (Normal / Osteoporosis)
disease_pred_prob = mobilenet_model.predict(image_input)
disease_pred = np.argmax(disease_pred_prob)
predicted_disease_label = disease_class_names[disease_pred]

# Predict severity only if predicted as Osteoporosis
predicted_severity_label = 'N/A'
if predicted_disease_label == 'Osteoporosis':
    severity_pred_prob = cnn_model.predict(image_input)
    severity_pred = np.argmax(severity_pred_prob)
    predicted_severity_label = severity_class_names[severity_pred]
```

```python
# Display results
plt.imshow(cv2.cvtColor(image, cv2.COLOR_BGR2RGB))
plt.axis('off')
plt.title(f"True: {true_disease_label.title()} | Predicted: {predicted_disease_label}\nSeverity: {predicted_severity_label}")
plt.show()

# Print detailed output
print(f"Image path: {image_path}")
print(f"True Disease Label: {true_disease_label.title()}")
print(f"Predicted Disease Label: {predicted_disease_label}")
if predicted_disease_label == 'Osteoporosis':
    print(f"Predicted Severity Level: {predicted_severity_label}")
else:
    print("No severity prediction (Normal case).")
```

The Osteoporosis Detection Project focuses on identifying osteoporosis from medical imaging data using deep learning and transfer learning techniques. The core of the system uses MobileNetV2, a pre-trained convolutional neural network, to extract meaningful features from bone X-ray images categorized as either "Normal" or "Osteoporosis." The dataset is divided into training and validation directories, and each image is resized to 224x224 pixels and preprocessed using the MobileNetV2 pipeline. The extracted features are then flattened and used to train a simple classifier, such as Logistic Regression, for binary classification. This approach leverages the high-level image feature extraction capability of deep models while maintaining a lightweight and interpretable final classification step. The model’s performance is evaluated using metrics like accuracy, confusion matrix, and classification report, offering insight into its diagnostic reliability. This system aims to provide a scalable, efficient tool that could support early osteoporosis diagnosis and assist medical professionals in making quicker and more accurate

CHAPTER 5
CONCLUSION

5.1 Conclusion

In this project, an intelligent diagnostic system for osteoporosis detection was developed using deep learning techniques, specifically leveraging MobileNetV2 for efficient feature extraction from bone X-ray images. The features obtained were used to train a lightweight classifier, enabling accurate classification between osteoporotic and normal bone conditions. The model showed promising performance, demonstrating the potential of AI in supporting early diagnosis and reducing the reliance on manual analysis. By automating the detection process, this system can help healthcare professionals identify at-risk patients more quickly and consistently, especially in resource-limited settings. Despite its success, the project does face limitations, such as dependency on the quality and variety of the input data, and the need for further validation across different medical environments. For future improvements, incorporating more diverse datasets, enhancing model explainability, and integrating with real-time clinical tools could significantly boost its reliability and usability. Overall, this project lays the groundwork for deploying AI-assisted diagnostic tools in the fight against osteoporosis, contributing to preventive healthcare and improved patient outcomes.

5.2 Future Work:

To further enhance the accuracy, reliability, and real-world applicability of the osteoporosis detection system, several directions can be explored in future work. First, expanding the dataset with a larger and more diverse set of bone images, including scans from different medical centers and imaging modalities like DEXA or CT, could improve the model's generalizability. Additionally, incorporating explainable AI (XAI) techniques such as Grad-CAM would allow the system to highlight important regions in the image that influence its predictions, thereby building trust among medical professionals. The current pipeline can also be extended to support multi-class classification, enabling differentiation between various stages of bone deterioration. Integration with a graphical user interface (GUI) or deployment as a web or mobile application would make the system more user-friendly for clinical staff. Finally, validating the model in collaboration with healthcare providers and real patient data would be essential to ensure compliance with medical standards and support potential FDA or CE approval for use in clinical practice.

CHAPTER 6
REFERENCES

[1] M. Li, Y. Jiang, Y. Zhang, and H. Zhu, “Medical image analysis using deep learning algorithms,” Front Public Health, vol. 11, 2023, doi: 10.3389/fpubh.2023.1273253.

[2] Y. Bengio, “Learning deep architectures for AI,” Foundations and Trends in Machine Learning, vol. 2, no. 1, pp. 1–27, 2009, doi: 10.1561/2200000006.

[3] Samuel W. Hasinoff, Photon, Poisson Noise. Computer Vision: A Reference Guide, 2014.

[4] D. N. H. Thanh, V. B. S. Prasath, and L. M. Hieu, “A review on CT and X-ray images denoising methods,” 2019, Slovene Society Informatika. doi: 10.31449/inf.v43i2.2179.

[5] S. Rai, J. S. Bhatt, and S. K. Patra, “Augmented Noise Learning Framework for Enhancing Medical Image Denoising,” IEEE Access, vol. 9, pp. 117153–117168, 2021, doi: 10.1109/ACCESS.2021.3106707.

[6] Parkinson J. An essay on the shaking palsy. J Neuropsychiatry Clin Neurosci. 2002;14(2):223–36.

[7] Lombardo JM, Lopez MA, Miron F, López M, León M, Arambarri J, Álvarez D. MOBEEZE Natural Interaction Technologies, Virtual Reality and Artifcial Intelligence for Gait Disorders Analysis and Rehabilitation in Patients with Parkinson’s Disease. Int J Interact Multi Artif Intell. 2019;5:54–62.

[8] Mack S. Removing Unclassifed Hand Tremor Motion from Computer Mouse Input with Neural Networks. Int J Interact Multi Artif Intell. 2018;5:56–60.

[9] Xu S, Pan Z. A novel ensemble of random forest for assisting diagnosis of Parkinson's disease on small handwritten dynamics dataset. Int J Med Inform. 2020;144:104283, 12.

[10] US Preventive Services Task Force; Curry SJ, Krist AH, et al. Screening for Osteoporosis to Prevent Fractures: US Preventive Services Task Force Recommendation Statement. JAMA. 2018;319(24):2521-2531.

[11] Tuzun S, Eskiyurt N, Akarirmak U, et al. Incidence of hip fracture and prevalence of osteoporosis in Turkey: the FRACTURK study. Osteoporos Int. 2012;23(3):949-955.

[12] Dimai HP. Use of dual-energy X-ray absorptiometry (DXA) for diagnosis and fracture risk assessment; WHO-criteria, Tand Z-score, and reference databases. Bone. 2017;104:39-43.

[13] Sebro R, De la Garza-Ramos C. Support vector machines are superior to principal

components analysis for selecting the optimal bones' CT attenuations for opportunistic screening for osteoporosis using CT scans of the foot or ankle. Osteoporos Sarcopenia. 2022 Sep;8(3):112-122. doi: 10.1016/j.afos.2022.09.002. Epub 2022 Sep 24. PMID: 36268496; PMCID: PMC9577430.

[14] D. Hans, A. M´etrailler, E. Gonzalez Rodriguez, O. Lamy, E. Shevroja, Quantitative Ultrasound (QUS) in the Management of Osteoporosis and Assessment of Fracture Risk: An Update, Adv. Exp. Med. Biol. 1364 (2022) 7–34, https://doi.org/10.1007/ 978-3-030-91979-5_2. PMID: 35508869.

[15] N.D. Nguyen, J.A. Eisman, J.R. Center, T.V. Nguyen, Risk factors for fracture in nonosteoporotic men and women, J. Clin. Endocrinol. Metab. 92 (3) (2007 Mar) 955–962, https://doi.org/10.1210/jc.2006-1476. Epub 2006 Dec 12 PMID: 17164302.

[16] E.S. Siris, P.D. Miller, E. Barrett-Connor, K.G. Faulkner, L.E. Wehren, T.A. Abbott, M.L. Berger, A.C. Santora, L.M. Sherwood, Identification and fracture outcomes of undiagnosed low bone mineral density in postmenopausal women: results from the National Osteoporosis Risk Assessment, JAMA 286 (22) (2001 Dec 12) 2815–2822, https://doi.org/10.1001/jama.286.22.2815. PMID: 11735756.

[17] Bilbily A, Syme CA, Adachi JD, et al. Opportunistic Screening of Low Bone Mineral Density From Standard X-Rays. J Am Coll Radiol. 2023 Oct 5:S1546-1440(23) 00715-9. doi: 10.1016/j.jacr.2023.07.024. Epub ahead of print. PMID: 37805012.

[18] R. Jang, J.H. Choi, N. Kim, J.S. Chang, P.W. Yoon, C.H. Kim, Prediction of osteoporosis from simple hip radiography using deep learning algorithm, Sci. Rep. 11 (1) (2021 Oct 7) 19997, https://

[19] J. Lu, Y. Shin, M.S. Yen, S.S. Sun, Peak bone mass and patterns of change in total bone mineral density and bone mineral contents from childhood into young adulthood. J. Clin. Densitom. 2016 Apr-Jun;19(2):180-91. doi: 10.1016/j. jocd.2014.08.001. Epub 2014 Oct 18. PMID: 25440183; PMCID: PMC4402109.