# DR. AMBEDKAR INSTITUTE OF TECHNOLOGY

(An Autonomous institution, affiliated to VTU, Belgaum, Aided by Government of Karnataka)
Near Jnana Bharathi Campus, Mallathahalli, Bengaluru-560056)

**Major Project – Report/21ISP706**

**On**

# "Osteoporosis Detection in Spine using Deep Learning Methods"

**BACHELOR OF ENGINEERING**

**IN**

**INFORMATION SCIENCE AND ENGINEERING**

2024-2025

**SUBMITTED BY**


| | |
|---|---|
| NITHIN S | [USN:1DA21IS033] |
| ROOPITHA G NAYAK | [USN:1DA21IS040] |
| SHREYAS D | [USN:1DA21IS048] |
| SHRINIDHI S HEGDE | [USN:1DA21IS050] |


**Under the guidance of**


Dr. K R SHYLAJA
Asst. Professor, Dept of AI/ML
Dr. AIT, Bengaluru


**Department of Information Science & Engineering**

**DR. AMBEDKAR INSTITUTE OF TECHNOLOGY**

**Outer Ring Road, Near Jnana Bharathi Campus,**
**Mallathalli, Bengaluru-560056**

2024-2025

# Dr. AMBEDKAR INSTITUTE OF TECHNOLOGY

(An Autonomous Institution, affiliated to VTU, Belgaum, Aided by Government of Karnataka) Near Jnana Bharathi Campus, Mallathalli, Bangalore – 560056

## CERTIFICATE

This is to certify that this project work entitled "**Osteoporosis Detection in Spine using Deep Learning Methods**" by **NITHIN S [1DA21IS033], ROOPITHA G NAYAK [1DA2IS040], SHREYAS D [1DA21IS048] AND SHRINIDHI S HEGDE [1DA21IS050]** submitted in partial fulfillment of the requirements of the 6th semester Project Phase 2 report for the degree of Bachelor of Engineering in **Information Science & Engineering** of **Dr. Ambedkar Institute of Technology**, Bengaluru, during the academic year 2023-24, is a bonafide record of work carried out under my guidance and supervision.

| Signature of the Guide | Signature of HOD | Signature of Principal |
|---|---|---|
| **Dr. K R SHYLAJA** | **Dr. PRERANA CHAITHRA** | **Dr. M N THIPPESWAMY** |
| **Assistant Professor** | **Professor, Head of the Department** | **Principal** |
| **Dept. of AI/ML, Dr. AIT** | **Dept. of ISE, Dr. AIT** | **Dr. AIT** |

**External viva:**

Name of the examiners — Signature with Date

1)

2)

# ACKNOWLEDGEMENT

The satisfaction that accompanies to this project would be incomplete without the mention of the people who made it possible, without whose constant guidance and encouragement would have made our efforts go in vain.

We consider ourselves privileged to express our gratitude and respect towards all those who guided us through the project, **"Osteoporosis Detection in Spine using Deep Learning Methods"**.

We would like to express our gratitude to **Dr. M N Thippeswamy, Principal, Dr. A.I.T.,** for providing us the congenial environment to work in.

We would like to express our profuse gratitude to **Dr. Prerana Chaithra, Professor HOD, Dept. of Information Science & Engineering, Dr. AIT,** for giving us the support, encouragement and providing us the required lab facilities that was necessary for the completion of this project.

As a token of gratitude, we would like to acknowledge our sincere gratefulness to our internal guide **Dr. K R Shylaja, Assistant.Prof., Dept. of AI/ML, Dr.A.I.T.,** for her unlimited support and encouragement provided throughout the process.

We also express our gratitude and sincere thanks to all the teaching and non-teaching staff of **Information Science & Engineering Department.**

Finally, yet importantly, we would like to express our heartfelt thanks to our beloved **Parents** for their blessings and our **Friends** for their help and wishes for the successful completion of this project report.

NITHIN S [1DA21IS033]
ROOPITHA G NAYAK [1DA21IS040]
SHREYAS D [1DA21IS048]
SHRINIDHI S HEGDE [1DA21IS050]

# DECLARATION

We, Nithin S [1DA21IS033], Roopitha G Nayak [1DA21IS040], Shreyas D [1DA21IS048],Shrinidhi S Hegde [1DA21IS050] students of 7th semester B.E, Department of Information Science and Engineering, Dr. AMBEDKAR INSTITUTE OF TECHNOLOGY, Bengaluru hereby declare that the project work entitled "OSTEOPOROSIS DETECTION USING DEEP LEARNING METHODS" has been duly executed by us under the guidance of Dr. K R Shylaja, Assistant Professor, Department Artificial Intelligence and Machine Learning Engineering, Dr. Ambedkar Institute of Technology, Bengaluru and submitted in partial fulfilment of the requirement.

Date :23/12/2024

Place : Bengaluru

NITHIN S [1DA21IS033]
ROOPITHA G NAYAK [1DA21IS040]
SHREYAS D [1DA21IS048]
SHRINIDHI S HEGDE [1DA21IS050]

# ABSTRACT

Bone mineral density (BMD) is the international standard for evaluating osteoporosis/osteopenia. The success rate of BMD alone in estimating the risk of vertebral fragility fracture (VFF) is approximately 50%, making BMD far from ideal in predicting VFF. In addition, whether or not a patient has been diagnosed with osteoporosis or osteopenia, he or she may suffer a VFF. For this reason, we conducted an extensive empirical study to assess VFFs in postmenopausal women. We considered a representative dataset of spine MRI (with osteopenia or osteoporosis). Comparing the classification results of machine learning and deep learning (DL) techniques showed that DL generally achieved better results at the cost of higher computational power and hard maintainability. SVM with ResNet CNN would achieve the best results in discriminating patients from groups with and without VFFs. Our results would represent a significant step toward prospective and longitudinal studies investigating methods to achieve higher accuracy in predicting VFFs based on spine MRI features of vertebrae without fracture

# Table of Content

# Table of Figures

# INTRODUCTION

# CHAPTER 1

# INTRODUCTION

## 1.1 OVERVIEW ON MACHINE LEARNING

Machine Learning (ML) is a subset of artificial intelligence (AI) that focuses on developing algorithms and statistical models that enable computers to perform specific tasks without explicit instructions. Instead, these systems learn from data patterns and make decisions or predictions based on that data. ML has become integral in various domains, including healthcare, finance, marketing, and more, due to its ability to handle vast amounts of data and uncover insights that are often beyond human capabilities.

### 1.1.1 Types of Machine Learning:

**Supervised Learning:** Involves training a model on a labeled dataset, where the input and output are provided. The model learns to map inputs to outputs and makes predictions on new, unseen data. Examples include classification (e.g., predicting whether an email is spam or not) and regression (e.g., predicting house prices).

**Unsupervised Learning:** Involves training a model on data without labelled responses. The model tries to identify patterns and relationships within the data. Examples include clustering (e.g., customer segmentation) and association (e.g., market basket analysis).

**Semi-Supervised Learning:** Combines both labeled and unlabeled data to improve learning accuracy. It is particularly useful when labelling data is expensive or time-consuming.

**Reinforcement Learning:** Involves training an agent to make a sequence of decisions by rewarding it for good actions and penalizing it for bad ones. It is used in applications like robotics, gaming, and autonomous vehicles.

### 1.1.2 Popular Machine Learning Algorithms:

**Linear Regression:** Used for regression tasks, predicting a continuous output based on input features.

**Logistic Regression:** Used for binary classification tasks, predicting the probability of a binary outcome.

**Decision Trees:** Tree-like models used for both classification and regression, which split data into branches to make predictions.

**Random Forests:** An ensemble method that combines multiple decision trees to improve accuracy and reduce overfitting.

**Support Vector Machines (SVM):** Used for classification and regression tasks, SVMs find the optimal hyperplane that separates data points of different classes.

**Neural Networks:** Inspired by the human brain, these models are composed of layers of interconnected nodes and are particularly powerful for complex tasks like image and speech recognition.

### 1.1.3 Challenges in Machine Learning:

**Overfitting:** Occurs when a model learns the noise in the training data instead of the actual patterns, leading to poor performance on new data. Techniques like regularization, pruning, and cross-validation help mitigate overfitting.

**Underfitting:** Occurs when a model is too simple to capture the underlying patterns in the data, resulting in poor performance on both training and test data.

Data Quality: The quality and quantity of data directly impact model performance. Handling missing values, dealing with imbalanced datasets, and removing outliers are crucial steps in data preprocessing.

**Interpretability:** Some ML models, like neural networks, are often seen as "black boxes". Ensuring model transparency and interpretability is important, especially in critical applications like healthcare and finance.

## 1.2 APPLICATIONS OF MACHINE LEARNING

Machine Learning (ML) has made significant strides across various industries, transforming how tasks are performed and decisions are made. Its ability to analyse large datasets, uncover patterns, and make predictions has enabled innovative solutions in numerous fields. Here, we explore some of the prominent applications of ML.

### 1. Healthcare

**Disease Prediction and Diagnosis**

ML models can analyse patient data to predict the likelihood of diseases such as heart disease, diabetes, and cancer. By identifying patterns and risk factors, these models assist in early diagnosis and personalized treatment plans, leading to better patient outcomes.

**Medical Imaging**

ML techniques, particularly deep learning, are used to analyse medical images such as X-rays, DXA Scans , DXAs, and CT scans. They help in detecting anomalies, segmenting organs, and diagnosing conditions with high accuracy, often surpassing human performance

Drug Discovery and Development

ML accelerates the drug discovery process by predicting molecular behaviour, identifying potential drug candidates, and optimizing clinical trials. This reduces time and cost, bringing effective treatments to market faster.

**Personalized Medicine**

By analysing genetic information and patient history, ML models provide personalized treatment recommendations, ensuring that therapies are tailored to individual needs and genetic profiles.

**2. Finance**

**Fraud Detection**

ML algorithms monitor transaction patterns in real-time to detect fraudulent activities. By learning from historical data, these models identify anomalies and flag suspicious transactions, enhancing security for financial institutions and their customers.

**Algorithmic Trading**

ML models analyse market trends, historical data, and other financial indicators to make informed trading decisions. These algorithms execute trades at optimal times, maximizing profits and minimizing risks.

Credit Scoring

ML improves the accuracy of credit scoring by assessing the creditworthiness of individuals based on a variety of data sources. This leads to more reliable lending decisions and better risk management for financial institutions.

Risk Management

ML helps in assessing and mitigating risks by analyzing market data, economic indicators, and other variables. It supports financial institutions in developing strategies to manage potential losses and maintain stability.

### 3. Marketing

**Customer Segmentation**

ML algorithms segment customers based on their behaviour, preferences, and demographics. This enables businesses to tailor marketing campaigns to specific segments, increasing engagement and conversion rates.

Recommendation Systems

ML powers recommendation engines used by platforms like Amazon, Netflix, and Spotify. These systems analyse user behaviour to suggest products, movies, music, and more, enhancing user experience and boosting sales.

**Sentiment Analysis**

ML techniques analyse social media posts, reviews, and feedback to gauge customer sentiment. This helps businesses understand public perception, manage their brand reputation, and respond proactively to customer concerns.

Targeted Advertising

ML models optimize ad targeting by analysing user behaviour and preferences. This ensures that advertisements reach the right audience, increasing the effectiveness of marketing efforts and reducing costs.

### 4. Transportation

**Autonomous Vehicles**

ML is at the core of self-driving car technology, enabling vehicles to perceive their environment, make decisions, and navigate safely. Algorithms process data from sensors, cameras, and LIDAR to detect obstacles, follow traffic rules, and ensure passenger safety.

Traffic Prediction and Management

ML models analyse traffic patterns, historical data, and real-time information to predict congestion and optimize traffic flow. This helps in reducing travel time, fuel consumption, and emissions.

**Route Optimization**

ML algorithms optimize delivery routes for logistics companies, taking into account factors such as traffic, weather, and road conditions. This enhances efficiency, reduces costs, and improves customer satisfaction.

### 5. Manufacturing

**Predictive Maintenance**

ML models predict equipment failures by analysing sensor data and historical maintenance records. This allows for timely maintenance, reducing downtime and extending the lifespan of machinery.

**Quality Control**

ML techniques inspect products for defects by analysing images and sensor data. This ensures high-quality production and minimizes the chances of defective products reaching customers.

**Supply Chain Optimization**

ML optimizes supply chain operations by forecasting demand, managing inventory, and predicting delays. This leads to cost savings, improved efficiency, and better customer service.

### 6. Retail

**Inventory Management**

ML algorithms forecast demand for products, helping retailers maintain optimal inventory levels. This reduces overstocking and stockouts, leading to better customer satisfaction and reduced costs.

**Dynamic Pricing**

ML models adjust prices in real-time based on demand, competition, and other factors. This maximizes revenue by setting optimal prices that attract customers and improve profitability.

**7. Energy**

**Smart Grid Management**

ML enhances the efficiency and reliability of smart grids by predicting energy demand, optimizing distribution, and integrating renewable energy sources. This leads to more sustainable and cost-effective energy management.

**Energy Consumption Prediction**

ML models analyse consumption patterns to predict future energy usage. This helps in planning and managing energy resources more effectively, reducing waste and lowering costs.

**8. Education**

**Personalized Learning**

ML systems tailor educational content to individual students' learning styles and paces. This ensures a more effective learning experience and better academic outcomes.

Automated Grading

ML algorithms grade assignments and exams, providing instant feedback to students. This reduces the workload on educators and allows for more timely interventions.

## 1.3 Overview on Deep Learning

Deep learning, a subset of machine learning, focuses on algorithms inspired by the structure and function of the human brain, known as artificial neural networks. Unlike traditional machine learning models, deep learning methods automatically discover representations from raw data, eliminating the need for manual feature engineering. This makes deep learning particularly effective for tasks involving large, complex datasets.

At the core of deep learning are neural networks, which consist of layers of interconnected nodes (neurons). These layers include an input layer, one or more hidden layers, and an output layer. Each neuron processes data by applying a mathematical transformation, such as a weighted sum followed by an activation function. The network learns by adjusting these weights through backpropagation, a process that minimizes the error between predicted and actual outputs using optimization techniques like gradient descent.

Deep learning has achieved remarkable success in various domains, including image and speech recognition, natural language processing, autonomous systems, and healthcare. Techniques such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have revolutionized computer vision and sequence-based tasks, respectively. Furthermore, advancements in hardware, particularly GPUs and TPUs, have accelerated the training and deployment of deep learning models.

## 1.4 INTRODUCTION TO OSTEOPOROSIS

Osteoporosis is the prevailing bone's disease, and its features are low bone density mass and the modification of their micro-architecture structure, so that bones' tolerance is reduced and the risk of fracture is increased. In osteoporosis, the Bone Mineral Density (BMD) is reduced; the bone micro-architecture is disrupted whereas the concentration and the variety of proteins in bones are altered. The classic Osteoporotic fractures are hip, vertebral and wrist fractures. Osteoporotic fractures are defined as occurring at a site associated with low BMD and which at the same time increased in incidence after the age of 50 years. Often the first apparent symptom of osteoporosis is a broken bone, which is why the condition is also known as "the silent crippler", as people do not realize they have osteoporosis until it's too late. However early detection and treatment of osteoporosis can decrease the fracture risk of a person to a minimum. For these reasons, there are studies which use Artificial Intelligence for predicting and detecting if a person has osteoporosis or not. Osteoporosis slowly evolves as an asymptomatic condition with an increasing possibility of bone fragility fractures, in special Vertebral Fragility Fractures (VFF). According to the international osteoporosis foundation, 20% of men and 60% of women over fifty years old will suffer a bone fracture due to osteoporosis worldwide. Thus, osteoporosis is a public health issue that may increase the risk of VFF, decreasing physical function, causing immobility, social isolation, and depression, contributing to a lower quality of life . Thus, VFF's early treatment and prevention are essential. Osteoporosis has been one of the most common diseases in the US and has affected over 3 million patients every year. Numerous efforts have been made in recent years to provide early diagnosis of osteoporosis to reduce bone fracture risks. Currently, a golden standard is to utilize the Dual-energy Xray Absorptiometry (DXA) for osteoporosis diagnosis, which is, however, cost inefficient and impractical for routine examination.

# 1.5 EXISTING SYSTEM

The existing system for osteoporosis risk assessment on vertebral spine without fracture may involve traditional methods of diagnosis and assessment. Typically, the primary diagnostic tool for osteoporosis is Dual X-ray Absorptiometry (DXA) examination, which assesses bone mineral density (BMD). DXA is considered the golden standard for diagnosing osteoporosis and predicting vertebral fragility fractures (VFF) in postmenopausal women. However, the sensitivity of BMD alone is limited, around 50%, in evaluating the risk of VFF.

**1. Overview of Traditional Osteoporosis Risk Assessment Systems**

- **Dual X-ray Absorptiometry (DXA):**
  DXA is widely regarded as the gold standard for diagnosing osteoporosis and assessing fracture risk. It works by measuring Bone Mineral Density (BMD) at specific sites, such as the spine and hip.
- **How DXA Works:**
  DXA involves passing low-energy X-rays through the bone to measure its density. The results are presented as a T-score, which compares the patient's bone density to a healthy young adult population.
- **Limitations of DXA:**
  While effective, DXA is limited in its ability to predict vertebral fragility fractures (VFF) accurately. Studies suggest its sensitivity in evaluating VFF risk is approximately 50%. This means that many individuals at risk of fractures may not be identified through BMD alone.
  DXA does not account for other critical factors influencing bone strength, such as bone geometry, micro-architecture, or material properties.

**2. Emerging Concerns with Traditional Systems**

- **Undetected Risks in Early Stages:**
  Traditional systems may fail to identify patients with normal BMD who are still at significant risk due to deteriorating bone quality or other factors.

- **Inaccessibility of Advanced Imaging:**
  While tools like QCT and DXA provide richer insights, they are not widely available or cost-effective for routine screening.
- **Lack of Integration:**
  Current systems often treat BMD and clinical risk factors as separate entities, leading to fragmented risk assessments.

## 1.6 PROPOSED SYSTEM

Our envisioned osteoporosis risk assessment system for the vertebral spine not only enhances diagnostic accuracy but also extends its impact through an integrated web- based platform. The system incorporates advanced technologies, including radiomics models and texture attributes (TAs) extracted from lumbar spine DXA scans, optimizing predictive models through a hybrid approach that includes Convolution Neural Networks (CNNs), Visual Geometry Group (VGG-16), RESnet with RF Classifier ensembled to a produce single hybrid model. The user-friendly web interface seamlessly integrates into clinical workflows, offering transparent insights into osteoporosis risk. Alongside diagnostic capabilities, the platform aims to connect users with nearby healthcare professionals specializing in musculoskeletal health, facilitating early intervention.

Users accessing the website can upload their DXA images, initiating a comprehensive analysis that goes beyond traditional diagnostics. The system's output includes personalized risk assessments and recommendations.

To enhance user engagement, the platform provides a user-friendly dashboard displaying results and suggested actions. Integration with a database of nearby doctors specializing in bone health enables users to connect with relevant healthcare professionals, fostering a proactive approach to osteoporosis management. This holistic web-based solution prioritizes accuracy, accessibility, and early intervention, marking a significant step toward improving musculoskeletal healthcare.

### 1.6.1 SCOPE OF THE PROJECT

The main scope of this proposed system is to:

- Increase the accuracy of the existing system from 83% to more by using more dataset and more optimized algorithms.

- To train large datasets of DXA scans to increase the efficiency.
- To consider bone texture as a feature for parameter analysis

The main scope of this proposed system is as follows: To present a system to predict the Osteoporosis condition to prevent bone fracture risks in the future of the patients. To gain knowledge about DXA dataset collection and its feature extraction using most efficient algorithms. To develop a suitable deep learning model that would be able to analyse the medical condition with best possible accuracy. Support Vector Machine and Decision Tree methods are used and suitable classification techniques are selected based on accuracy value.

### 1.6.2 PROBLEM STATEMENT OF THE PROJECT

Develop a deep learning model to predict the risk of osteoporosis in non-fractured vertebral bones by analyzing DXA textural attributes.This project aims to bridge this gap by improving prediction accuracy through hybrid models like VGG16, Residual Networks (ResNet), and Convolutional Neural Networks (CNN) and ensemble all these trained models into a single hybrid model for better accuracy and precision in detection of osteoporosis.

# LITERATURE SURVEY

# CHAPTER 2

# LITERATURE SURVEY

**Radiomics Based on Lumbar Spine Magnetic Resonance Imaging to Detect Osteoporosis [1**] signal intensity of the lumbar spine in magnetic resonance imaging (DXA) correlates to bone mineral density (BMD). This study aims to explore a lumbar spine magnetic resonance imaging based on the radiomics model for detecting osteoporosis. Radiomic models established based on lumbar spine DXA can be used to detect osteoporosis.

- **Normalization and Alignment:**
  The learning-free method uses predefined rules and anatomical landmarks to align spine images. This ensures that variations in image orientation and scale are accounted for, which is critical for DXA scans due to patient-specific posture differences during imaging.
- **Segmentation Techniques:**
  They likely employ morphological operations, edge detection, or template matching to isolate the vertebrae from surrounding tissues. These preprocessing steps aim to provide clear regions of interest (ROIs) for localization.
- **Importance:**
  This approach contrasts with deep learning-based models like CNNs, which learn segmentation patterns automatically during training. Highlighting this reference shows how traditional methods paved the way for modern automated solutions and provides a baseline for evaluating preprocessing techniques.

**Frequency of normal bone measurement in postmenopausal women with fracture: a registry-based cohort study [2]** registry-based cohort study assessed the percentage of women with prior or incident fracture who had normal bone defined as a normal bone mineral density T-score and normal trabecular bone score (TBS). Inclusion of TBS reduced the percentage with normal bone. Normal bone measurement is rare in women with fracture. Including TBS with BMD increases identification of abnormal bone in women with fracture compared with BMD alone. Normal bone is present in < 6% of women with any fracture and < 1% of those with hip fracture. What is thought to be normal bone in women with fracture is rarely normal.

- **Ultrasound Radio Frequency Signals:**

  The research utilizes ultrasound radio frequency signals as the primary data source for diagnosis. This is a significant aspect as it explores an alternative approach to traditional methods like dual-energy X-ray absorptiometry (DXA).

- **Multi-channel Convolutional Neural Networks (MCNN)**:

  The study employs a deep learning architecture (MCNN) for analyzing and interpreting the complex patterns within the ultrasound signals. This demonstrates the potential of artificial intelligence in improving osteoporosis diagnosis.The study aims to explore the potential of ultrasound-based techniques and deep learning to improve the accuracy, accessibility, and cost-effectiveness of osteoporosis diagnosis.

**Association of bone mineral density with bone texture attributes extracted using routine magnetic resonance imaging [3**] Dual-energy X-ray absorptiometry (DXA)- derived bone mineral density (BMD) often fails to predict fragility fractures. Quantitative textural analysis using magnetic resonance imaging (DXA) may potentially yield useful radiomic features to predict fractures. We aimed to investigate the correlation between BMD and texture attributes (TAs) extracted from DXA scans and the interobserver reproducibility of the analysis. Specific TAs could be reliably extracted from routine DXA and correlated with BMD. Our results encourage future evaluation of the potential usefulness of quantitative texture measurements from DXA scans for predicting fragility fractures.

- **Radiomics and Feature Extraction:**

  The study explores the use of radiomics a field combining medical imaging and data science to identify potential lumbar fractures using DXA scans. Radiomics involves extracting a large amount of quantitative features from medical images, which are then analyzed using computational techniques to aid in diagnosis or prognosis.

The researchers developed a methodology or system named BEAUT (likely an acronym standing for the proposed approach). The goal of BEAUT is to improve the identification of lumbar fractures, which can often be challenging to diagnose accurately due to the complex anatomy of the lumbar spine and variability in imaging quality.

**Data-driven diagnosis of spinal abnormalities using feature selection , machine learning algorithms[4]** This paper focuses on the application of machine learning algorithms for predicting spinal abnormalities. As a data preprocessing step, univariate feature selection as a filter based feature selection, and principal component analysis (PCA) as a feature extraction algorithm are considered. A number of machine learning approaches namely support vector machine (SVM), logistic regression (LR), bagging ensemble methods are considered for the diagnosis of spinal abnormality. The SVM, LR, bagging SVM and bagging LR models are Lokesh M. Giripunje, Suchita Sudke, Pradnya Wadkar, Krishna Ambure worked on applied on a dataset of 310 samples publicly available in Kaggle repository.

- **Bone Texture Analysis:**
  This study investigates the correlation between BMD and TAs from DXA scans and assesses interobserver reproducibility.
- **DXA-Derived Attributes in Machine Learning:**
  TAs from routine DXA scans correlate with BMD and can be reliably extracted.Quantitative texture measurements show potential for predicting fragility fractures and merit further investigation..
- **Non-Invasive Diagnostic Techniques:**
  Use their findings to emphasize the potential of DXA as a non-invasive alternative to DXA for osteoporosis detection.

**Osteoporosis Detection using machine learning techniques and feature selection[5]** All classifiers have been evaluated using the well-known 10-fold cross validation method and the results are reported analytically. In addition, a feature selection method identified that with the use of only two diagnostic factors (age and weight), similar performance could be achieved. The scope of the proposed exhaustive methodology is to assist therapists in osteoporosis prediction, avoiding unnecessary further testing with bone densitometry.

This study introduces a new method to predict osteoporosis by addressing data imbalance during pre-processing, using a sine–cosine algorithm combined with an information gain fuzzy-rough set for selecting the most relevant features. Classifiers then predict osteoporosis based on these selected features. The method was tested on benchmark datasets and real osteoporosis data, showing competitive results compared to other approaches. The selected features were highly correlated with osteoporosis, highlighting their effectiveness in prediction.

**Data-driven diagnosis of spinal abnormalities using feature selection , machine learning algorithms[4]** This paper focuses on the application of machine learning algorithms for predicting spinal abnormalities. As a data preprocessing step, univariate feature selection as a filter based feature selection, and principal component analysis (PCA) as a feature extraction algorithm are considered. A number of machine learning approaches namely support vector machine (SVM), logistic regression (LR), bagging ensemble methods are considered for the diagnosis of spinal abnormality. The SVM, LR, bagging SVM and bagging LR models are Lokesh M. Giripunje, Suchita Sudke, Pradnya Wadkar, Krishna Ambure worked on applied on a dataset of 310 samples publicly available in Kaggle repository. The performance of classification of abnormal and normal spinal patients is evaluated in terms of a number of factors including training and testing accuracy, recall, and miss rate.

**Correlation of bone mineral density with DXA T2* values in quantitative analysis of lumbar osteoporosis [6]** the DXA T2* value is moderately negatively correlated with the bone mineral density assessed with quantitative computed tomography in evaluating osteoporosis in postmenopausal women and may have some potential in assessing severity of lumbar osteoporosis for scientific research. Purpose: To investigate the T2* quantitative measurement in magnetic resonance imaging (DXA) and its correlation with the bone mineral density (BMD) values evaluated with quantitative computed tomography (QCT) in women with postmenopausal lumbar vertebrae osteoporosis.

**The study aimed to develop a predictive model using ensemble artificial neural networks (ANNs) to identify key risk factors for hip bone fractures in elderly individuals. [7]**

- **Ensemble Artificial Neural Networks:**

The study employs a powerful machine learning technique, specifically an ensemble of artificial neural networks, for the prediction task. Ensemble methods combine multiple individual models to improve overall prediction accuracy and robustness.

- **Key Risk Factors:**

The research aims to identify and utilize key risk factors associated with hip bone fractures. These factors could include age, medical history, lifestyle habits, and other relevant parameters.

- **Biomedical Signal Processing and Control:**

This research falls under the broader field of biomedical signal processing and control, which involves the application of signal processing and control theory to biomedical problems.

**Risk Group Identification: It specifically targets the identification of individuals at high risk of developing osteoporosis, which can guide preventative measures [8]** This is for informational purposes only. For medical advice or diagnosis, consult a professional.Osteoporosis is a silent disease that weakens bones and increases the risk of fractures. Early detection is crucial for preventative measures. This paper reviews the use of artificial intelligence (AI) in identifying individuals at high risk of developing osteoporosis. AI models can analyze various factors, such as medical history, lifestyle, and genetic predispositions, to predict the likelihood of developing osteoporosis. This information can be used to implement targeted preventative measures for high-risk individuals, potentially reducing the incidence of fractures and improving bone health.

## 2.1 Comparative Study of the Literature Survey Papers

| Sl. No. | Title of the Paper | Methodology | Contributions | Limitations |
|---|---|---|---|---|
| 1 | Radiomics Based on Lumbar Spine Magnetic Resonance Imaging to Detect Osteoporosis, May 2021 | Lumbar Magnetic Resonance Imaging | Radiomic models based on lumbar spine DXA were used to detect osteoporosis. The performance of the classification models was evaluated through the estimated area under the receiver using only the characteristic curve. | |
| 2 | Frequency of Normal Bone Measurement in Postmenopausal Women with Fracture: A Registry-Based Cohort Study, August 2021 | Dual-energy X-ray Absorptiometry, Trabecular Bone Score (TBS) | Including TBS with BMD increases the identification of abnormal bone in women with fractures compared with BMD alone. | What is thought to be normal bone in women with fractures is rarely normal. |
| 3 | Association of Bone Mineral Density with Bone Texture Attributes Extracted Using Routine Magnetic Resonance Imaging, March 2020 | Quantitative Textural Analysis Using Magnetic Resonance Imaging (DXA) | Texture measurements were used to predict fragility fractures. | The cervical and thoracic spinal segments were not studied. |
| 4 | Data-Driven Diagnosis of Spinal Abnormalities Using Feature Selection and Machine Learning Algorithms, February 2020 | SVM, LR, Bagging SVM, and Bagging LR Models | Machine learning algorithms were used for predicting spinal abnormalities. | The computation time of the classifiers was not studied. |
| 5 | Osteoporosis Detection Using Machine Learning Techniques and Feature Selection, May 2022 | Support Vector Machines (SVM), Random Forests, Artificial Neural Networks (ANN), Logistic Regression (LR) | Feature selection was used to accurately identify the risk of osteoporosis in postmenopausal women. | Structural design was not used for constructing the models. No comparative study of various analytical methods was carried out. |
| 6 | Correlation of Bone Mineral Density with DXA T2* Values in Quantitative Analysis of Lumbar Osteoporosis | Quantitative Computed Tomography T2* Mapping | The DXA T2* value is moderately negatively correlated with the bone mineral density assessed with quantitative computed tomography in evaluating osteoporosis in postmenopausal | Gender-specific study, single imaging modality, limited sample size. |

| Sl. No. | Title of the Paper | Methodology | Contributions | Limitations |
|---|---|---|---|---|
| | | | women and may have some potential in assessing the severity of lumbar osteoporosis for scientific research. | |
| 7 | Ensemble Artificial Neural Networks Applied to Predict the Key Risk Factors of Hip Bone Fracture for Elders | Clinical Data Analysis and Three-Layer Back-Propagation ANN Models | Three-layer back-propagation ANN models were developed separately for female and male subjects to predict hip bone fracture risk. | The study's findings are based on the specific dataset used, which may limit the generalizability of the results to broader populations or different demographic groups. |
| 8 | Artificial Intelligence on the Identification of Risk Groups for Osteoporosis | Hybrid Ensembles Combining Neural Networks and Decision Trees | The selected systems were evaluated based on diagnostic coverage, cost-effectiveness, and their ability to identify significant somatic factors. A comprehensive review of various AI methodologies applied in screening for osteoporosis risk. | Combining various databases and clinical attributes presents complexities, particularly in developing standardized, evidence-based criteria for evaluating significant risk factors. |

Table 2.1.1 Literature Survey Papers

# REQUIREMENT SPECIFICATION

# CHAPTER 3

# REQUIREMENT SPECIFICATION

The Requirement Specification details the system's functional and non-functional requirements, guiding the development process. It includes user, system, and interface requirements, specifying expected operations and quality attributes like performance and security. This ensures clear understanding among stakeholders and provides a baseline for validation throughout development.

## 3.1 FUNCTIONAL REQUIREMENTS

The functional requirements for a system describe what the system should do. These requirements depend on the type of software being developed and the overall approach taken by the organization when writing the requirements. Functional system requirements describe the system's functions in detail, including its inputs, outputs, exceptions, and other related aspects.

Functional requirements are as follows:

- Time delay should be minimized.
- Extracting the features efficiently.
- Provides faster GUI interaction.

**Hardware Requirements**

- Processor: i5 Processor
- RAM: 512Mb
- Hard Disk: 10GB
- Input Device: Standard keyboard and Mouse
- Output Device: VGA and High-Resolution Monitor
- DXA (Magnetic Resonance Imaging) Scans basically DXA scans.

**Software Requirements**

- TensorFlow Python (version:3.10)
- Pytorch Library(version:1.13)
- Anaconda Environment(version:2.3.1)

## 3.2 NON-FUNCTIONAL REQUIREMENTS

Non-functional requirements are requirements that aren't directly concerned with the precise functions delivered by the system and relate to the emergent system properties like reliability, reaction time and store occupancy.

Alternatively, they'll define constraints on the system like capabilities of I/O devices and therefore the data representations utilized in the system interfaces. The non-functional requirements are as follows:

- Provide an easy interface for physically challenged and disabled people.
- Provides maximum accuracy.
- Handles errors efficiently.

# METHODOLOGY AND SYSTEM DESIGN

# CHAPTER 4

# METHODOLOGY AND SYSTEM DESIGN

The Methodology and System Design section outlines the approach and framework used to develop the system. It includes the selection of appropriate models, such as CNN, VGG16, and RESNET with RF, and describes the data preprocessing, model training, and evaluation processes. This section details the design architecture, including the layers and components of the models, and explains how they interact to achieve the desired outcomes.

## 4.1 Data Pre-processing

In this phase, we gather the dataset of DXA scan images of the spine of the patients from scan centers.

**Data Collection Phase**

In this phase, we gather the dataset of DXA scan images of the spine of the patients from scan centers.

**Augmentation:**

Augmentation is the process of artificially increasing the amount of data by generating new data points from existing data. It involves creating new images based on the existing dataset to improve its size, thereby reducing overfitting.

**Standardization:**

Standardization involves scaling data to fit a standard normal distribution. This ensures that all variables contribute equally. It is the process of creating standards and transforming data taken from different sources into a consistent format that adheres to these standards.

**Normalization:**

Normalization is the process of transforming features to be on a similar scale. This improves the performance and training stability of the model, dramatically enhancing its accuracy.

**Annotation:**

Data annotation is the process of labeling content to make it recognizable by machines through computer vision. This ensures systems deliver accurate results and helps modules identify elements to train computer vision and speech recognition models.

**Outlier Rejection:**

Outliers are data points significantly different from the rest of the dataset. They increase error variance and reduce the power of statistical tests.

**De-noising:**

De-noising is an advanced technique used to decrease grainy spots and decolorization in images while minimizing the loss of quality.

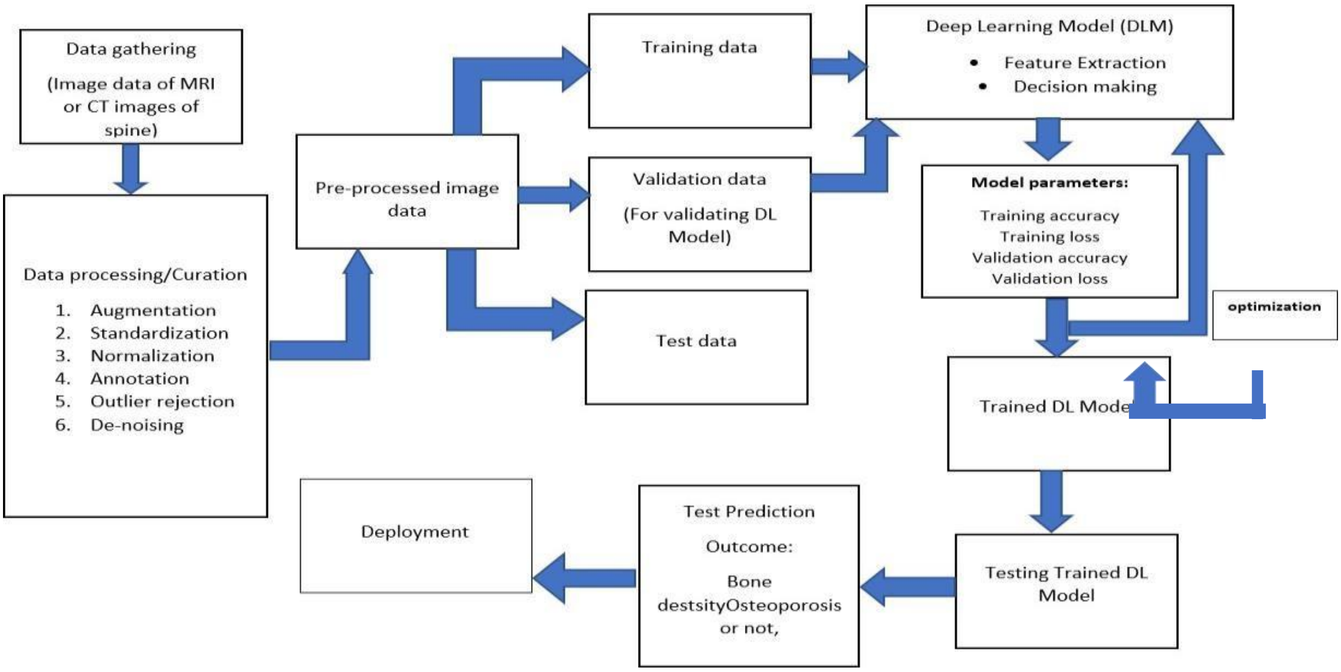


Fig 4.1 System Architecture

## 4.2 Classification

In this phase, we will divide the dataset into three sets: **Training data**, **Validation data**, and **Testing data**.

- **Training Dataset:**
  Training data is the initial dataset used to train the deep learning model. It teaches the DL algorithms how to make predictions or perform a desired task. In our system, 70% of the main dataset will be used as training data.
- **Validation Dataset:**
  Validation data provides the first test against unseen data, allowing us to evaluate how well the model performs on new data. In our system, 10% of the overall dataset will be used as validation data.
- **Testing Dataset:**
  Testing data is a subset of the original dataset used to check the model's accuracy. It ensures validation of all boundary conditions, including data with all possible combinations of boundary values. In our system, 20% of the overall dataset will be used as testing data.

### 4.3 Deep Learning Model

In this phase, we will be using hybrid algorithm that is SVM with RESNET CNN to implement our Deep Learning model to achieve the purpose of diagnosis of Osteoporosis.

**VGG16**-The VGG16 model is a deep convolutional neural network architecture that has made significant contributions to the field of computer vision. It consists of 16 convolutional layers followed by max-pooling layers, with small 3x3 filters applied throughout the network. This uniform architecture makes it easy to understand and implement, while also achieving impressive performance on various visual recognition tasks, such as image classification and object detection. Despite being surpassed by newer architectures in terms of accuracy, VGG16 remains a popular choice for its balance between performance and simplicity, serving as a foundational model in the development of more advanced convolutional neural networks.

**RESNET**– Residual Neural Networks: A residual neural network (ResNet) is an artificial neural network (ANN). It is a gateless or open-gated variant of the HighwayNet, the first working very deep feedforward neural network with hundreds of layers, much deeper than previous neural networks. Residual network (ResNet) is a Convolution Neural Network (CNN) architecture that overcame the “Vanishing gradient” problem, making it possible to construct networks with upto thousands of Convolutional layers, which out perform shallower networks.

**CNN** – Convolution Neural Networks: A CNN is a kind of network architecture for deep learning algorithms and is specifically used for image recognition and tasks that involve the processing of pixel data. There are other types of neural networks in deep learning, but for identifying and recognizing objects, CNNs are the network architecture of choice. The Convolutional Neural Network (CNN or ConvNet) is a subtype of Neural Networks that is mainly used for applications in image and speech recognition. Its built-in convolutional layer reduces the high dimensionality of images without losing its information.

Feature Extraction: Feature Extraction is the process of transforming raw data into numerical features that can be processed while preserving the information in the original data set. It yields better results than applying machine learning directly to the raw data. Here, in this system, we will be extracting Bone Mineral Density (BMD) and bone texture using feature extraction process.

Decision making: Here, in this phase we will be deciding on whether or not the Osteoporosis condition would exist in the bones of the patient in the future. After this phase, the deep learning model will be created suiting our purpose of Osteoporosis diagnosis for future.

Now that all these three models are developed and tested, we will use Ensemble Learning Techniques bascially Static method Ensemble Learning Technique to create a hybrid of all these models and create a single model which will provide a better accuracy, precision in detection of Osteoporosis. By this, we will be deploying this model on real-time applications for predicting the Osteoporosis condition on the present condition of the bone to avoid future fracture risks of the patients.

## 4.4 Data Partitioning

The dataset is divided into three subsets to facilitate model training and evaluation

**Training Dataset:** 70% of the dataset is used for training the model to learn patterns and relationships within the data.

**Validation Dataset:** 10% of the dataset is reserved for tuning model parameters and assessing its performance on unseen data during training.

**Testing Dataset:** 20% of the dataset is used for final model evaluation, ensuring the model generalizes well to new data.

## 4.5 Feature Extraction

Feature extraction is a critical step in deep learning workflows for medical imaging. It involves isolating and quantifying relevant attributes from DXA scans to facilitate accurate predictions by the classification model. For this project, the following aspects of feature extraction are considered:

### 4.5.1 Bone Mineral Density (BMD)

Bone Mineral Density is one of the primary indicators of bone strength and osteoporosis. The system computes BMD values from DXA images using advanced image analysis techniques. High-resolution imaging is leveraged to ensure accurate representation of bone density.

### 4.5.2 Bone Texture Analysis

Bone texture features, such as trabecular patterns, are extracted using convolutional layers within the CNN models. These patterns help distinguish between healthy and osteoporotic bones.

### 4.5.3 Region of Interest (ROI) Selection

Automated or manual segmentation techniques are employed to isolate regions of interest in DXA scans, such as vertebrae or other critical bone structures. This prevents irrelevant data from influencing model predictions.

### 4.5.4 Dimensionality Reduction

Once features are extracted, dimensionality reduction techniques like Principal Component Analysis (PCA) are applied to eliminate redundant features. This ensures the model focuses on the most informative aspects of the data, improving efficiency and accuracy.

### 4.5.5 Edge and Gradient Features

Using edge detection filters and gradient operations, the system captures fine structural details of the bones. These features contribute to identifying micro-fractures or early signs of osteoporosis.

## 4.6 Model Evaluation

The evaluation phase ensures that the developed model is robust, generalizes well, and performs effectively in real-world applications. The following evaluation methods and considerations are employed:

### 4.6.1 Evaluation Metrics

The following metrics are used to assess the model's performance:

**Accuracy:** Measures the overall correctness of the predictions.

**Precision:** Evaluates the proportion of true positive results among all positive predictions.

**Recall (Sensitivity):** Measures the ability of the model to correctly identify positive cases (e.g., detecting osteoporosis when it exists).

**Specificity:** Evaluates the model's capability to identify true negatives.

**F1-Score:** The harmonic mean of precision and recall, providing a balance between the two.

Area Under the Curve (AUC): Evaluates the performance of the model over different classification thresholds.

### 4.6.2 Cross-Validation

A k-fold cross-validation approach is used, where the dataset is split into k folds. The model is trained on k-1 folds and validated on the remaining fold, iteratively. This ensures the model's robustness and avoids overfitting.

### 4.6.3 Hyperparameter Tuning

Grid search or randomized search techniques are employed to optimize hyper parameters such as learning rate, batch size, and the number of filters in convolutional layers.

### 4.6.4 Baseline Model Comparison

The hybrid RESNET-SVM model's performance is compared against baseline models (e.g., CNN, VGG16) to validate its superiority in feature extraction and classification accuracy.

### 4.6.5 Confusion Matrix Analysis

The confusion matrix is analyzed to identify the distribution of true positives, true negatives, false positives, and false negatives. This helps identify areas where the model needs improvement.

### 4.6.6 Real-World Testing and Validation

The trained model is deployed in a simulated real-world environment using DXA images from external sources. Its predictions are evaluated by comparing them with expert diagnoses to ensure clinical reliability.

### 4.6.7 Error Analysis

Miss-classified samples are analyzed to understand the reasons for incorrect predictions. This analysis informs further improvements to the model, such as refining feature extraction methods or adjusting hyper parameters.

### 4.6.8 Robustness Testing

The model's robustness is tested by introducing variations such as noisy inputs, low-resolution images, and unseen data distributions to evaluate its adaptability and reliability in diverse scenarios.

### 4.6.9 Deployment Readiness

Post-evaluation, the model is tested for real-time performance metrics such as inference time, scalability, and resource efficiency to ensure readiness for deployment in clinical settings.

# SYSTEM IMPLEMENTATION

# CHAPTER 5

## SYSTEM IMPLEMENTATION

Model training involves leveraging pre-trained VGG16, CNN, and ResNet architectures, fine-tuning them with our dataset over multiple epochs to optimize their performance in identifying osteoporotic conditions. Then these models are ensembled together to generate one hybrid model used to detect Osteoporosis. The integration phase employs FastAPI as the backend framework to manage model inference and interactions efficiently, while the frontend interface is developed using HTML and CSS, ensuring a user-friendly experience for medical professionals.

The validation process includes rigorous testing of the models using metrics such as accuracy, precision, recall, and F1 score, alongside cross-validation to verify the models' generalization capabilities. Following validation, the system is deployed in a clinical setting. Continuous performance monitoring and periodic updates with new data ensure the system remains accurate and effective over time, providing a reliable tool for the early detection and timely treatment of osteoporosis.

The project focuses on developing a system for detecting spine osteoporosis using deep learning models (VGG16, CNN, ResNet) and Random Forest (RF) models integrated together to form one hybrid model using knowledge of Ensemble Learning. The implementation involves collecting and preprocessing spine images, fine-tuning pre-trained deep learning models through transfer learning, and training RF classifiers on extracted features. The system's performance is validated using metrics like accuracy and F1 score, ensuring reliable detection.

Then all these models are integrated together to a hybrid model which will be used for detecting the osteoporosis condition, hence enhancing the accuracy of the model.

The final step involves deploying the integrated system in a clinical setting, making it user-friendly for medical professionals. Continuous monitoring and periodic updates with new data will maintain and enhance the detection accuracy, providing an effective tool for early osteoporosis diagnosis.

# 5.1 CNN MODEL IMPLEMENTATION

The implementation of the CNN model for detecting spine osteoporosis involves several key steps. Initially, a comprehensive dataset of spine images is collected and preprocessed through normalization, augmentation, and segmentation to ensure high- quality inputs. The CNN model architecture is then designed or selected, with a focus on layers such as convolutional layers, pooling layers, and fully connected layers to effectively extract and learn features from the images. The model undergoes training over multiple epochs, utilizing techniques like backpropagation and optimization algorithms to minimize the loss function and enhance accuracy. During training, the model's performance is continuously monitored using validation metrics such as accuracy, precision, recall, and F1 score, ensuring its reliability in detecting osteoporosis.

- **Convolutional Layers**: Extract spatial features from the images by applying filters that detect edges, textures, and patterns indicative of osteoporosis.
- **Activation Functions**: Apply non-linear transformations like ReLU (Rectified Linear Unit) to introduce non-linearity, enabling the network to learn complex patterns.
- **Pooling Layers**: Down sample the feature maps using max pooling or average pooling to reduce dimensionality and computation while retaining critical features.
- **Fully Connected Layers**: These layers connect all neurons, aggregating features learned by convolutional layers to make predictions about osteoporosis.
- **Output Layer**: Use an output layer with a softmax or sigmoid activation function, depending on whether the model performs binary classification (osteoporosis vs. no osteoporosis) or multi-class classification.

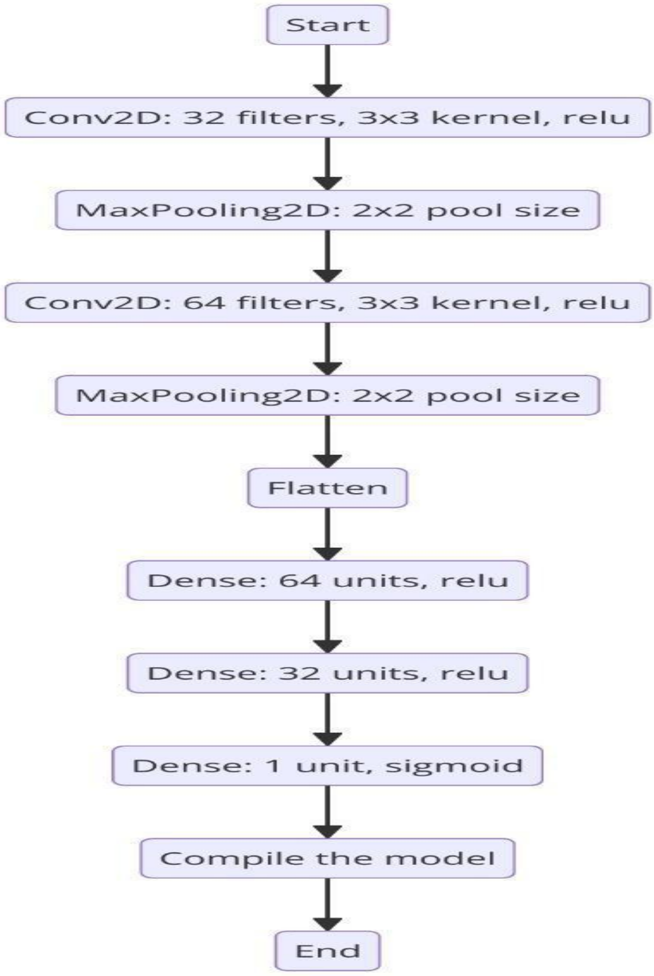


Fig 5.11 CNN Model Architecture

| Layer (type) | Output Shape | Param # |
|---|---|---|
| conv2d_4 (Conv2D) | (None, 222, 222, 32) | 320 |
| max_pooling2d_4 (MaxPooling2D) | (None, 111, 111, 32) | 0 |
| conv2d_5 (Conv2D) | (None, 109, 109, 32) | 9,248 |
| max_pooling2d_5 (MaxPooling2D) | (None, 54, 54, 32) | 0 |
| conv2d_6 (Conv2D) | (None, 52, 52, 32) | 9,248 |
| max_pooling2d_6 (MaxPooling2D) | (None, 26, 26, 32) | 0 |
| conv2d_7 (Conv2D) | (None, 24, 24, 32) | 9,248 |
| max_pooling2d_7 (MaxPooling2D) | (None, 12, 12, 32) | 0 |
| flatten_1 (Flatten) | (None, 4608) | 0 |
| dense_3 (Dense) | (None, 64) | 294,976 |
| dense_4 (Dense) | (None, 32) | 2,080 |
| dense_5 (Dense) | (None, 1) | 33 |

Fig 5.12 CNN Model Trained Summary

# 5.2 VGG16 MODEL IMPLEMENTATION

The implementation of the VGG16 model for detecting spine osteoporosis begins with collecting and preprocessing a large dataset of spine images, including steps like normalization, augmentation, and segmentation to enhance image quality. Utilizing the pre-trained VGG16 architecture, known for its depth and performance in image recognition tasks, we fine-tune the model with our specific dataset. This process involves replacing and retraining the top layers of VGG16 while preserving its learned weights for lower layers, thus leveraging its feature extraction capabilities. The model is trained over multiple epochs, employing optimization techniques to minimize the loss function and improve detection accuracy. Performance is evaluated using validation metrics such as accuracy, precision, recall, and F1 score.

**Leveraging the VGG16 Architecture**

**Transfer Learning with VGG16:**

VGG16 is pre-trained on ImageNet, a dataset with millions of images across 1,000 categories. This pre-training allows the model to capture generic visual features like edges and textures.

**Fine-Tuning Strategy:**

- The lower layers (closer to the input) of VGG16, which learn general features, are frozen to retain their pre-trained weights.
- The top layers (closer to the output) are replaced with new fully connected layers designed for binary classification (osteoporotic vs. non-osteoporotic).
- Dropout layers are added to mitigate overfitting.
- The new architecture might include layers such as:
- A global average pooling (GAP) layer to reduce feature maps' dimensionality.
- Dense layers with activation functions like ReLU for non-linearity.
- A final dense layer with a sigmoid activation function for binary classification.

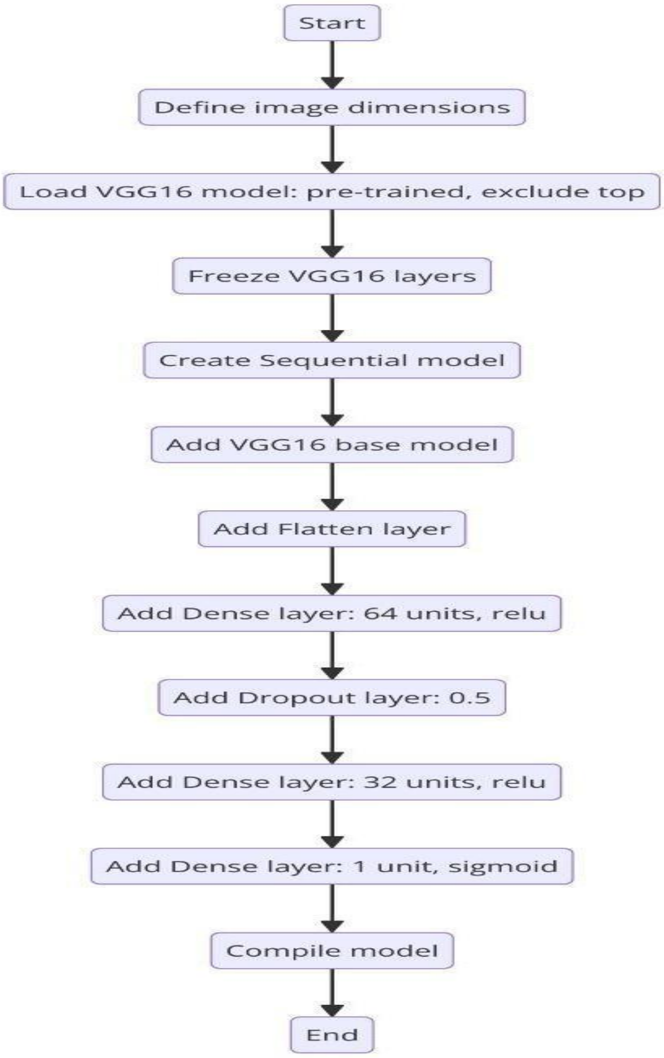


Fig 5.2.1 VGG16 Model Architecture

## 5.3 ResNet WITH RF CLASSIFIER IMPLEMENTATION

The implementation of the ResNet model combined with a Random Forest (RF) classifier for detecting spine osteoporosis begins with an extensive preprocessing phase, involving normalization, augmentation, and segmentation of spine images to ensure high-quality inputs. ResNet, known for its residual learning capabilities and deep architecture, is fine-tuned on the preprocessed dataset, with multiple epochs run to optimize performance. The deep features extracted by the ResNet model are then used to train the RF classifier, which excels in handling complex feature sets and providing robust classification results. This hybrid approach leverages ResNet's deep learning strengths and the RF model's decision tree ensemble capabilities. Validation is performed using metrics like accuracy, precision ,recall, and F1 score to ensure the system's effectiveness

- **Transfer Learning with ResNet:**
  - A pre-trained ResNet model (e.g., ResNet50 or ResNet101) is fine-tuned for the dataset.
  - The lower layers of ResNet, responsible for learning basic image features like edges and textures, are frozen to retain pre-trained weights.
  - The final layers are replaced with new layers designed for feature extraction. This ensures the extracted features are specific to spine osteoporosis detection.
- **Residual Connections:**
  These skip connections in ResNet help learn residual mappings, enabling efficient gradient flow and preventing degradation in deep networks.
- **Feature Extraction:**
  instead of directly classifying outputs, the model’s penultimate layer (feature vector) is used as input for the Random Forest classifier. These vectors are high-dimensional representations of the image's key features.

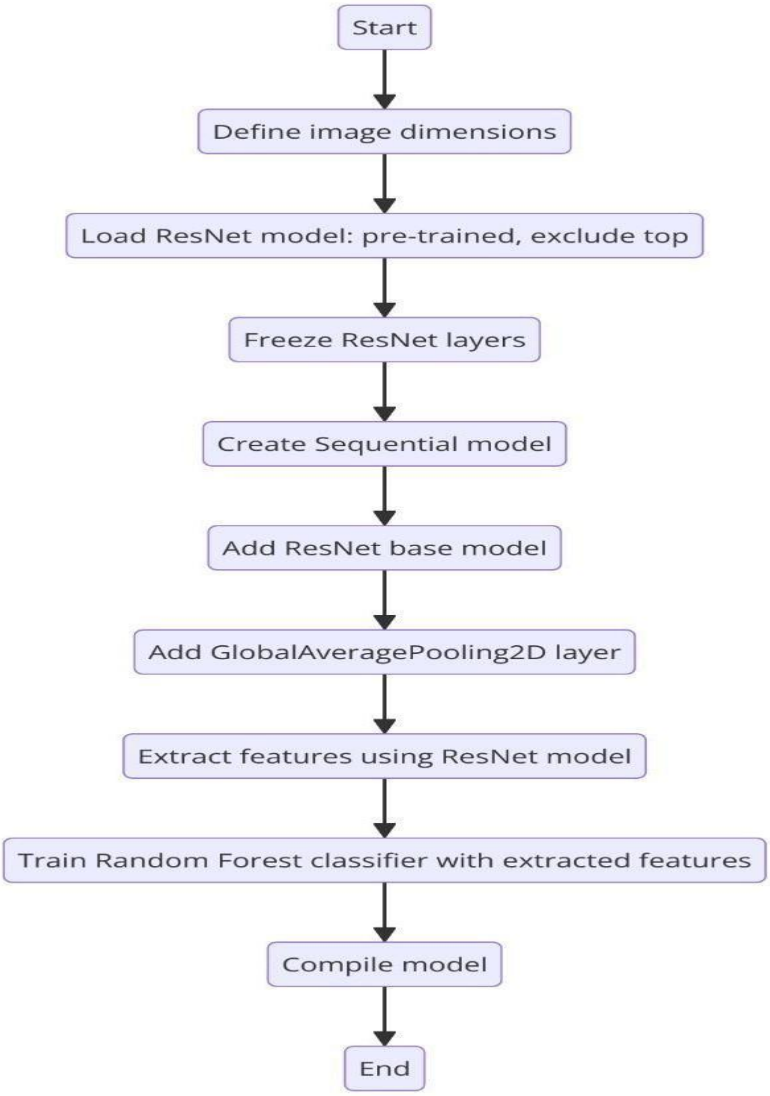


Fig 5.3.1 ResNet with RF classifier Model Architecture

## 5.4 Ensemble Learning Implementation (Stacking):

Ensemble means ‘a collection of things’ and in Machine Learning terminology, Ensemble learning refers to the approach of combining multiple ML models to produce a more accurate and robust prediction compared to any individual model. It implements an ensemble of fast algorithms (classifiers) such as decision trees for learning and allows them to vote.

Now all the models created such as CNN, ResNet with RF Classifier and VGG 16 are combined using a method of ensemble learning called stacking.

**Stacking**: It combines the output of multiple base models by training a combiner(an algorithm that takes predictions of base models) and generates more accurate predictions. Stacking allows for more flexibility in combining diverse models, and the combiner can be any machine learning algorithm.

Now with this model created, outputs from various models are taken and based on voting, a final output is provided which is more accurate than the individual model’s output.

- **Base Models (Level-0 Models):**

**CNN:** Trained to extract spatial and local features from the input spine images, providing predictions based on learned patterns.

**ResNet with RF Classifier:** ResNet extracts deep features from the images, which are then passed to the Random Forest Classifier for robust decision-making.

**VGG16:** Fine-tuned on the dataset to leverage its pre-trained weights and depth, generating predictions based on its advanced feature extraction capabilities.

Each model processes the spine image dataset and outputs its individual predictions, such as probabilities for osteoporosis detection.

- **Combiner (Level-1 Model):**

  - **Stacking** involves training a **meta-model** or **combiner** to integrate the outputs of the base models.
  - This combiner could be a lightweight classifier, such as a **logistic regression model**, **SVM**, or even another deep learning model.
  - The meta-model takes the predictions (e.g., probabilities or class labels) from the CNN, ResNet with RF Classifier, and VGG16 as inputs, learning how to weigh and combine them optimally.

- **Voting Mechanism:**

  - To ensure robust and accurate final predictions:
  - **Hard Voting:** If each model outputs a class label (e.g., "osteoporosis" or "healthy"), the final prediction is made based on the majority class.
  - **Soft Voting:** If each model outputs probabilities, the combiner aggregates these probabilities and predicts the class with the highest combined probability.
  - Voting helps balance the strengths and weaknesses of individual models, ensuring that no single model dominates the ensemble's decision-making.

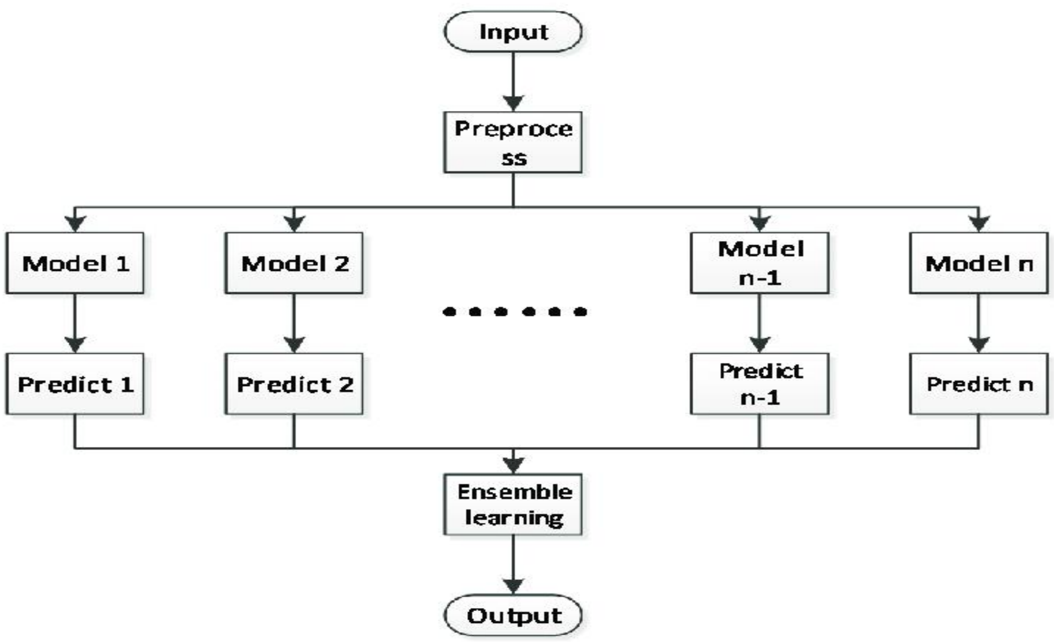


Fig 5.4.1 Ensemble Learning Flow Diagram

# SYSTEM TESTING

# CHAPTER 6

# SYSTEM TESTING

System testing is a crucial phase in software development or model deployment, focusing on evaluating the system's performance against predefined requirements. It involves rigorous testing to detect defects, assess functionality, and ensure reliability across various conditions. In the context of the osteoporosis detection project, system testing is essential for validating model accuracy and readiness for real-world deployment.

## 6.1 Accuracy testing

Accuracy testing is a pivotal phase in system testing, focusing on evaluating the predictive models' ability to correctly classify osteoporosis and non-osteoporosis cases. It involves measuring metrics like precision, recall, F1-score, and overall accuracy to assess the models' performance against predefined criteria. By systematically analyzing the models' performance on a designated test dataset, accuracy testing provides valuable insights into their reliability and suitability for real-world deployment, facilitating comparisons between different models and identifying areas for improvements.

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| NORMAL | 0.97 | 0.96 | 0.97 | 108 |
| Osteoporosis | 0.96 | 0.97 | 0.96 | 89 |
| accuracy | | | 0.96 | 197 |
| macro avg | 0.96 | 0.96 | 0.96 | 197 |
| weighted avg | 0.96 | 0.96 | 0.96 | 197 |

Fig 6.11 Classification Report of CNN Model

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| NORMAL | 0.94 | 0.92 | 0.93 | 108 |
| Osteoporosis | 0.90 | 0.93 | 0.92 | 89 |
| accuracy | | | 0.92 | 197 |
| macro avg | 0.92 | 0.92 | 0.92 | 197 |
| weighted avg | 0.92 | 0.92 | 0.92 | 197 |

Fig 6.12 Classification Report of VGG16 Model

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| NORMAL | 0.87 | 0.90 | 0.88 | 108 |
| Osteoporosis | 0.87 | 0.83 | 0.85 | 89 |
| accuracy | | | 0.87 | 197 |
| macro avg | 0.87 | 0.86 | 0.87 | 197 |
| weighted avg | 0.87 | 0.87 | 0.87 | 197 |

Fig 6.13Classification Report of ResNet Model

## 6.2 Performance Testing

The ROC curve evaluates the performance of a classification model by plotting the true positive rate (sensitivity) against the false positive rate (1-specificity). It helps in understanding how well the model distinguishes between classes.

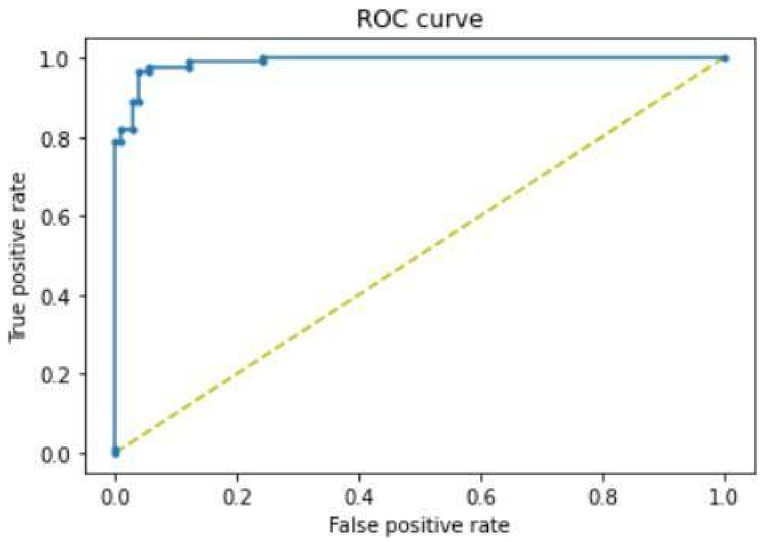


Fig 6.2.1 ROC Curve of CNN model

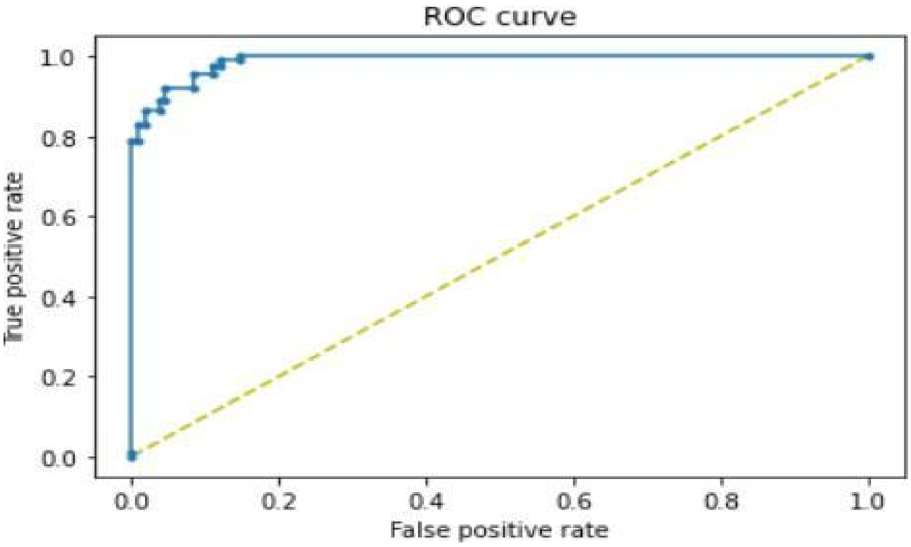


Fig 6.2.2 ROC Curve of Vgg16 model

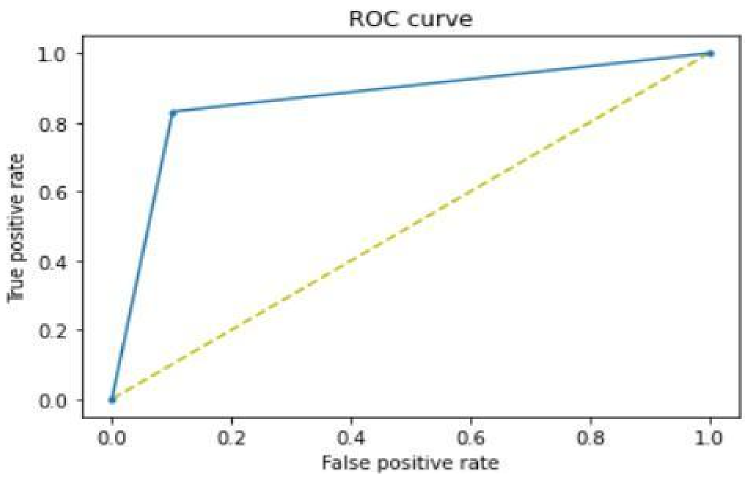


Fig 6.2.3 ROC Curve of ResNet model

## 6.3 Comparative Analysis

- **CNN:** Boasts the highest accuracy among the models, suitable for tasks requiring intricatepattern recognition, yet demands substantial computational resources and data for training.
- **VGG16**: Offers a simpler architecture than CNN, making it easier to implement and understand, thus ideal for scenarios with limited computational resources or where interpretability is key.
- **ResNet with RF Classifier:** Strikes a balance between depth and interpretability, leveraging ResNet's capacity for capturing complex features and RF Classifier's robustness, making it a viable option for diverse datasets with varying computational constraints and interpretative needs.

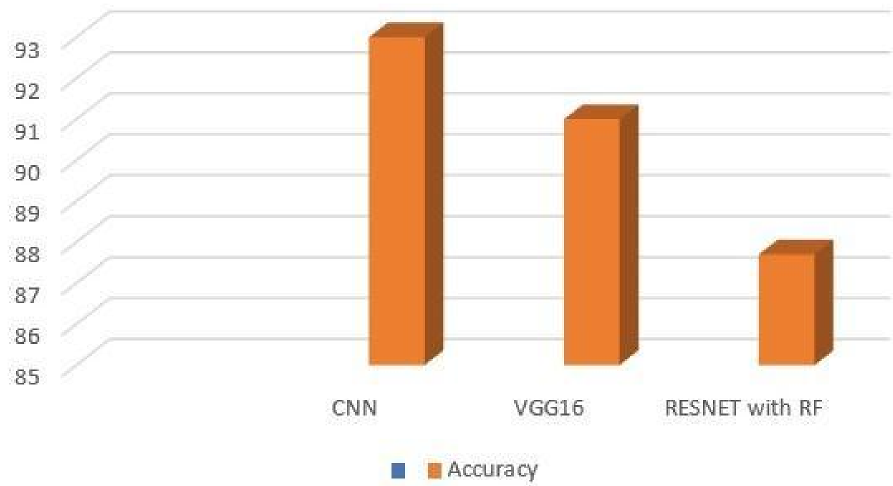


Fig 6.3.1 Comparative Analysis of the Trained models

The analysis of the accuracy bar chart reveals that the CNN model significantly outperforms the other models, achieving an accuracy of approximately 97%. VGG16 follows with an accuracy of around 90%, demonstrating solid performance but still falling short of CNN. RESNET with RF, however, shows the lowest accuracy at about 87%, indicating that it is the least effective model among the three in this context. These results suggest that for the given dataset and task, CNN is the most reliable model, providing the highest accuracy, while VGG16 and RESNET with RF may require further optimization or might be less suitable for this specific application. Further implementing the concept of Ensemble Learning we will integrate all these models together to produce a single robust hybrid model which produces greater accuracy, precision and will be used to detect osteoporosis.

# SYSTEM STUDY

# CHAPTER 7

# SYSTEM STUDY

## 7.1 Feasibility Study

**Objective:** Determine the overall viability of predicting or assessing osteoporosis risk based on DXA scan images and clinical data.

**Market Demand**

**Analyze Demand:**

Examine the demand for osteoporosis risk prediction systems, considering the rising prevalence of osteoporosis, the aging population, and the growing focus on preventative healthcare.

Highlight the need for automated systems to aid clinicians in early diagnosis and prediction of osteoporosis to reduce fracture risks and improve patient outcomes.

**Health and Wellness Trends:**

Consider the adoption of AI-based healthcare solutions, increased patient engagement with diagnostic tools, and integration of machine learning models into radiological workflows for enhanced accuracy and efficiency.

**Technical Complexity**

**Data Integration:**

Evaluate the feasibility of integrating DXA scan images with clinical measurements (e.g., bone mineral density, bone texture) to predict osteoporosis risk.

Address challenges in handling diverse datasets, including image preprocessing, feature extraction, and annotations for training the models.

**Data Processing:**

Assess the preprocessing requirements for DXA scans, such as augmentation, normalization, standardization, and denoising.

Explore feature engineering to derive clinically relevant features like bone mineral density (BMD) and structural details.

Investigate the selection of hybrid deep learning models such as CNN, RESNET or VGG16 for optimal performance.

**Resource Availability**

**Data Sources:**

Ensure the availability of annotated DXA datasets from hospitals or research institutions, along with supplementary clinical data.

**Tools and Expertise:**

Confirm the availability of tools such as TensorFlow, PyTorch, and image preprocessing libraries.

Identify required expertise, including data scientists, radiologists, and machine learning engineers, for system development.

**Risk Analysis**

**Identify Potential Risks:**

Address challenges related to data privacy, regulatory compliance (e.g., HIPAA), and handling sensitive medical data.

Analyze the risk of bias in prediction models due to limited or unbalanced datasets.

**Integration Challenges:**

Examine potential obstacles in integrating the prediction system with existing hospital systems or electronic health record (EHR) platforms.

## 7.2 Economic Feasibility

**Objective:** Assess the economic viability of developing an DXA-based osteoporosis risk assessment system.

**Cost-Benefit Analysis**

**Estimate Costs:**

Identify expenses for acquiring DXA datasets, data annotation, software development, and AI infrastructure.

Include the cost of model training, validation, and deployment.

**Potential Benefits:**

Highlight benefits such as reduced long-term treatment costs for fractures, improved diagnostic accuracy, and enhanced clinical workflows.

Consider potential revenue streams from partnerships with healthcare providers and licensing AI solutions.

**ROI Calculation**

**Return on Investment:**

Calculate ROI based on expected adoption by hospitals, clinics, and diagnostic centers.

Assess patient engagement, healthcare cost savings, and potential scalability to other diagnostic applications.

## 7.3 Technical Feasibility

**Objective:** Evaluate the technical feasibility of predicting osteoporosis risk using DXA scans and deep learning techniques.

**Data Integration:**

**Compatibility and Availability:**

Ensure compatibility of DXA scan data with image processing and machine learning frameworks.

Validate the availability of high-quality annotated datasets for training and testing.

**Data Quality:**

Develop strategies to handle noise in DXA images, remove outliers, and standardize data for consistent analysis.

**Algorithm Development:**

**Machine Learning Models:**

Investigate the use of hybrid algorithms, such as VGG 16, RESNET, CNN and ensemble all these model to produce a hybrid single model to predict osteoporosis risk.

**Performance Metrics:**

Focus on metrics such as accuracy, sensitivity, specificity, and F1-score for clinically relevant performance evaluation.

## 7.4 Social Study

**Objective:** Analyze the social implications and acceptance of an DXA-based osteoporosis risk assessment system.

**Health Impact**

**Public Health Benefits:**

Highlight how the system can enable early detection and proactive management of osteoporosis, reducing fracture risks and healthcare costs.

**Patient Motivation:**

Emphasize the role of personalized risk assessments in encouraging healthier lifestyles, such as improved diet and physical activity.

**Privacy Concerns**

**Data Sensitivity:**

Address public concerns regarding the storage and analysis of sensitive medical data, ensuring adherence to GDPR and HIPAA regulations.

**Transparency and Consent:**

Develop clear communication strategies about data usage, offering opt-in/out mechanisms for patients.

## Conclusion

This feasibility study provides a detailed assessment of the viability of developing an DXA-based osteoporosis risk assessment platform. By examining market demand, technical complexity, economic viability, and social considerations, it outlines the potential advantages and challenges of deploying such a system. These findings serve as a roadmap for decision-makers in healthcare and technology sectors to advance system development and adoption strategies. The proposed system has the potential to significantly improve bone health outcomes, reduce fracture risks, and alleviate the burden on healthcare systems worldwide.

# OUTCOME SNAPSHOTS

# CHAPTER 8

# OUTCOME SNAPSHOTS

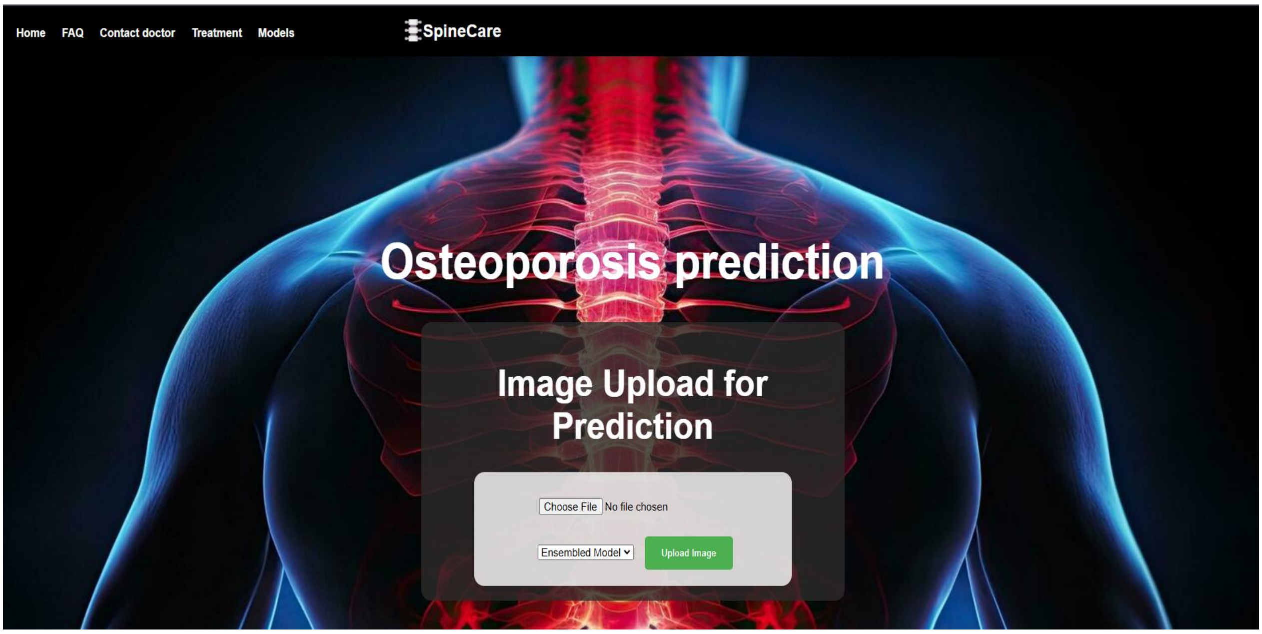


Fig 8.1 Home page

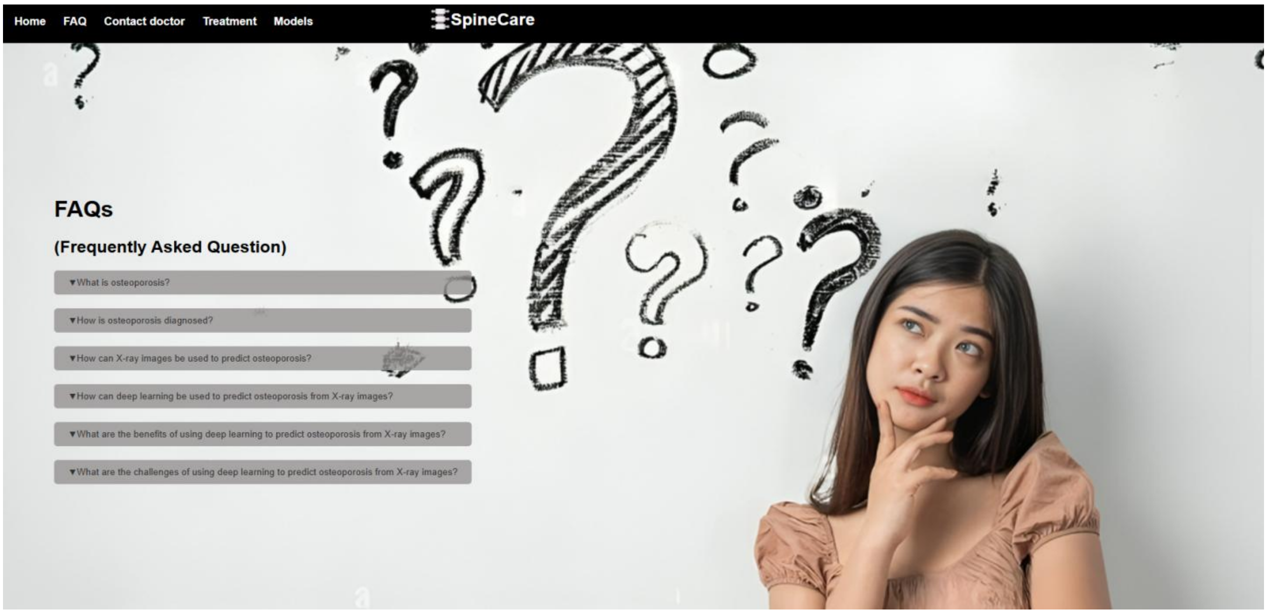


Fig 8.2 FAQ page

# OUTCOME SNAPSHOTS

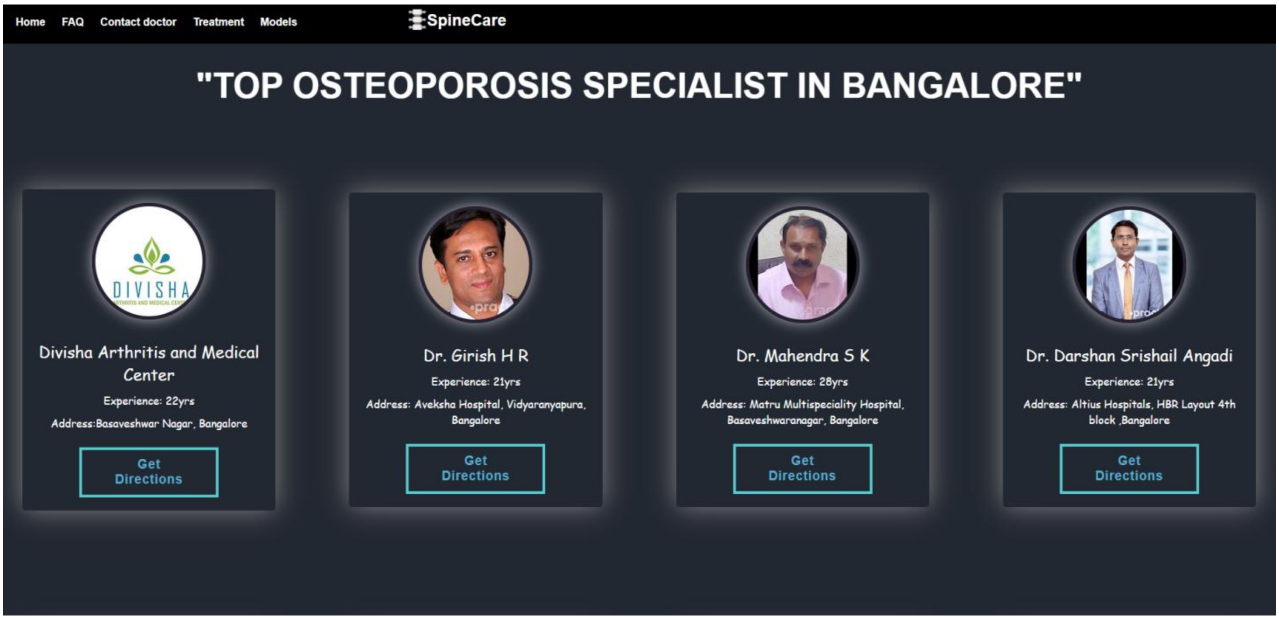


Fig 8.3 Doctors Information


Fig 8.4 Insights of Osteoporosis

# OUTCOME SNAPSHOTS

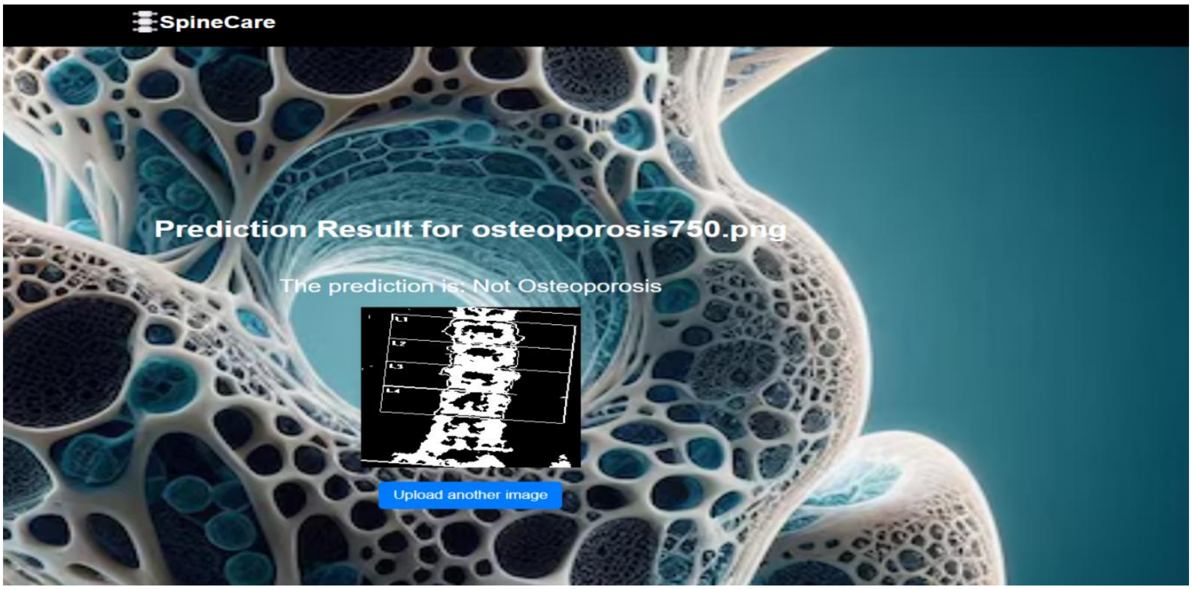


Fig 8.5 Result page

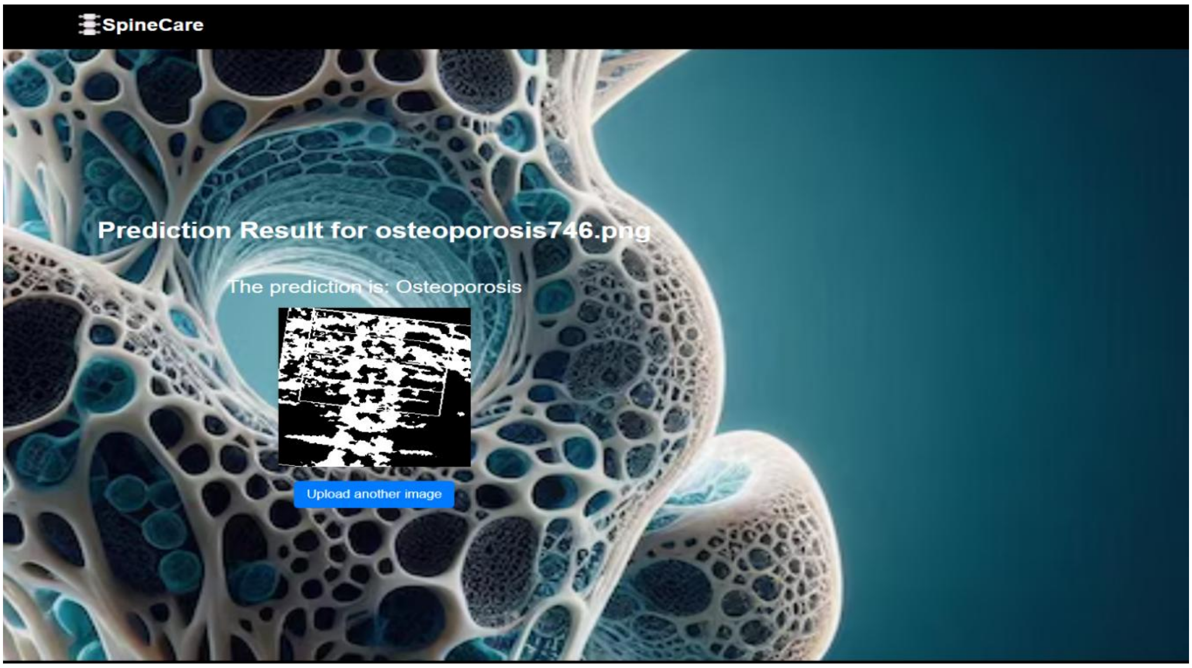


Fig 8.6 Result page

# CONCLUSION

# CHAPTER 9

## CONCLUSION

In the pursuit of advancing the field of medical image analysis, this project focused on "Osteoporosis Detection". Through a meticulous exploration of CNN, VGG16, ResNet combined with a Random Forest (RF) Classifier and a hybrid ensemble model, we gained valuable insights into their respective strengths and applicability. The Convolutional Neural Network (CNN) model demonstrated the highest accuracy among the models, highlighting its effectiveness in feature extraction and classification tasks. However, this accuracy came at the cost of significant computational resources and extensive data requirements for training. This makes CNNs particularly suited for environments where these resources are readily available. The VGG16 model, despite being less computationally intensive, offers simplicity and ease of implementation, which can be advantageous in scenarios with limited computational resources or where model interpretability is a priority. VGG16's straightforward architecture allows for easier debugging and tuning, making it a practical choice for rapid deployment and prototyping.

Leveraging ResNet in combination with a Random Forest (RF) Classifier provided a nuanced balance between model depth and interpretability, catering to the complex nature of medical imaging datasets. ResNet's deep architecture effectively captures intricate patterns in the data through residual learning, while the RF classifier enhances decision-making by utilizing an ensemble of decision trees, which improves robustness and generalization.

Stacking, a type of Ensemble Learning Is a different paradigm. The point of stacking is to explore a space of different models for the same problem. So, we built multiple different learners and used them to build an intermediate prediction, one prediction for each learned model. Then you add a new model which learns from the intermediate predictions the same target.

This final model is said to be stacked on the top of the others. Thus, improving overall performance, and end up with a model which is better than any individual intermediate model.

Moving forward, this research underscores the importance of informed model selection tailored to the specific requirements and constraints of medical image analysis tasks. By understanding the trade-offs between accuracy, computational demands, and interpretability, we can develop more accurate and accessible predictive tools for osteoporosis diagnosis and beyond, ultimately improving patient outcomes and advancing the capabilities of medical image analysis.

# FUTURE ENHANCEMENT

# CHAPTER 10

# FUTURE ENHANCEMENT

## 1. Integration with Clinical Decision Support Systems (CDSS)

- Develop an interface that integrates the model into clinical decision-making software, providing doctors with automated risk assessments and actionable insights.
- Include features like patient history analysis, BMD data integration, and personalized treatment recommendations.

## 2. Multi-Modality Data Fusion

- Extend the system to combine information from other medical imaging modalities such as CT scans, DXA scans, or ultrasound.
- Use multi-modal ensemble learning to improve the system's diagnostic accuracy by leveraging complementary information from different imaging sources.

## 3. Explainable AI (XAI)

- Implement explainability techniques like Grad-CAM or SHAP to visualize which areas of spine images contribute most to the model's predictions.
- This will help build trust among clinicians and allow them to validate the AI system's decisions.

## 4. Real-Time Processing Capability

- Optimize the system for real-time processing of spine images, enabling immediate results during clinical scans.
- Use model compression techniques like quantization or pruning to make the models lighter and deployable on edge devices.

## 5. Incorporation of Longitudinal Data

- Enable the system to analyze sequential spine images from the same patient over time to track disease progression.
- Incorporate temporal models such as Recurrent Neural Networks (RNNs) or Long Short-Term Memory (LSTM) networks for longitudinal analysis.

## 6. Personalized Risk Prediction

- Extend the model to integrate non-imaging factors such as age, gender, BMI, genetic predisposition, and lifestyle factors.
- Create a multi-input architecture that combines image-based predictions with statistical risk models for personalized risk assessments.

## 7. Generalization to Other Skeletal Disorders

- Adapt the model to detect and diagnose other skeletal conditions like scoliosis, vertebral fractures, or arthritis.
- Expand the dataset to include images of different skeletal regions, enabling the system to generalize its applications.

## 8. Cloud-Based Deployment

- Deploy the system on a cloud platform to provide access to hospitals and diagnostic centers globally.
- Enable secure data storage and HIPAA-compliant data processing for sensitive patient information.

## 9. Dataset Expansion and Diversity

- Collect a larger and more diverse dataset from multiple sources, including different ethnicities, age groups, and clinical conditions.
- Train the model on datasets from multiple imaging centers to improve generalizability and reduce bias.

## 10. Self-Learning and Continuous Improvement

- Implement a semi-supervised learning approach where the system learns from new, unlabeled data during deployment.
- Use active learning techniques to flag ambiguous cases for human review, improving the model's performance over time.

## 11. Integration with Wearable Technology

- Explore integration with wearable devices that monitor skeletal health indicators such as vitamin D levels, physical activity, or posture.
- Use this additional data to improve risk prediction and early detection capabilities.

## 12. Mobile Application Development

- Create a user-friendly mobile application for remote diagnosis and patient engagement.
- Include features like uploading spine images, accessing reports, and receiving alerts for follow-up consultations.

# REFERENCES

# CHAPTER 10

## REFERENCES

[1] Marko Rak(B) and Klaus-Dietz T¨onnies, A Learning- Free Approach to Whole Spine Vertebra Localization in DXA, Department of Simulation and Graphics, Otto von Guericke University, Magdeburg, Germany, 2016.

[2] Zhiwei Chen†, Wenqiang Luo†, Qi Zhang, Baiying Lei, Tianfu Wang, Zhong Chen, Osteoporosis Diagnosis Based on Ultrasound Radio Frequency Signal via Multi- channel Convolutional Neural Network, Sun Yat-sen Memorial Hospital Ethics Board (No: SYSEC-KY-KS-2019- 159), 2021.

[3] Jonathan S. Ramos∗, Jamilly G. Maciel†, Mirela T. Cazzolato∗ , Caetano Traina Jr.∗ , BEAUT: a radiomic approach to identify potential lumbar fractures in magnetic resonance imaging, Institute of Mathematics and Computer Science (ICMC), University of Sao Paulo (USP). S ˜ ao Carlos, SP, BR 13566-590 ˜ †Ribeirao Preto Medical School (FMRP), University of S ˜ ao Paulo (USP). Ribeir ˜ ao Preto, SP, BR 14048-900, 2021.

[4] Jamilly Gomes Maciel , I,* Iana Mizumukai de Arau´ jo, II Lucio C. Trazzi, I Paulo Mazzoncini de Azevedo-Marques, Association of bone mineral density with bone texture attributes extracted using routine magnetic resonance imaging, Financial Support: JGM: CAPES CODE 1; FJAP: FAPESP 2018/14060-9; MHN-B: FAPESP 2018/04266-9, CNPq 307138/2017-8 and 404795/2018-8, 2020.

[5] Ahmed A. Ewees 1 , Mohamed Abd Elaziz 2 , Rasha M. Arafa3 , And Rania M.Ghoniem, Improved Approach Based on Fuzzy Rough Set and Sine-Cosine Algorithm: A Case Study on Prediction of Osteoporosis, his work was supported by the Deanship of Scientific Research at Princess Nourah bint Abdulrahman University through the Fasttrack Research Funding Program, 2020.

[6] Correlation of bone mineral density with DXA T2 values in quantitative analysis of lumbar osteoporosis by Jin Liu, Jian-Wei Liao, Jian-Wei Liao, Wei Li, Xiao-Jun Chen, Jia-Xin Feng, Lin Yao, Pan-Hui Huang, Zhi-Hai Su, Hai Lu, Yu-Ting Liao, Shao-Lin Li, Ya-Jun 1 Ma, Ya-Jun Ma.

[7] Q. Liu, X. Cui, Y.-C. Chou, M. F. Abbod, J. Lin, and J.-S. Shieh, ‘‘Ensem-ble artificial neural

networks applied to predict the key risk factors of hip bone fracture for elders,'' Biomed. Signal Process. Control, vol. 21,pp. 146–156, Aug. 2015.

[8] Artificial intelligence on the identification of risk groups for osteoporosis, a general review Agnaldo S Cruz 1,✉, Hertz C Lins 2, Ricardo V A Medeiros 2, José M F Filho 2, Sandro G da Silva 1