Acta Informatica Pragensia
2026, Volume 15, Issue 1, 135–156
https://doi.org/10.18267/j.aip.294

PRAGUE UNIVERSITY OF ECONOMICS AND BUSINESS

Article Open Access

Optimizing Osteoporosis Detection with Cascaded Convolutional Neural Network and Real-Coded Genetic Algorithm

Hemalatha Balan[1], **Madhavi Latha Pandala**[2], **Venkatasubramanian Srinivasan**[3], **Venkatachalam Kandasamy**[4]

[1] Department of Information Technology, Dr. N. G. P. Institute of Technology, Coimbatore, India
[2] Department of Information Technology, Siddhartha Academy of Higher Education Deemed to be University, Vijayawada, India
[3] Department of Computer Science and Business Systems, Saranathan College of Engineering, Trichy, India
[4] Department of Computer Science and Engineering, Karunya Institute of Technology and Sciences, Karunya Nagar, Coimbatore, India

Corresponding author: Hemalatha Balan (balanhemalatha83@gmail.com)

Editorial Record

First submission received:
June 20, 2025

Revisions received:
July 31, 2025
October 3, 2025
October 21, 2025

Accepted for publication:
October 22, 2025

Academic Editor:
Zdenek Smutny
Prague University of Economics and Business, Czech Republic

This article was accepted for publication by the Academic Editor upon evaluation of the reviewers' comments.

How to cite this article:
Balan, H., Pandala, M. L., Srinivasan, V., & Kandasamy, V. (2026). Optimizing Osteoporosis Detection with Cascaded Convolutional Neural Network and Real-Coded Genetic Algorithm. *Acta Informatica Pragensia*, 15(1), 135–156. https://doi.org/10.18267/j.aip.294

Abstract

Background: Osteoporosis is a condition characterized by bones that are porous and brittle, increasing the risk of fractures. It is often asymptomatic until substantial harm develops, making it crucial to treat at the onset of the disorder.

Objective: This study aims to develop a practical, in-depth and adaptable framework utilizing constructed clinical and demographic datasets for the early detection of osteoporosis.

Methods: We design a cascade convolutional neural network with adaptive weight fusion and fine-tune it with a real-coded genetic algorithm. An anonymized clinical and demographic record publicly available bone mineral density dataset was used. Missing data identification, normalization and encoding of categorical variables were key diagnostic steps. The training subset consisted of 70% of the dataset, while the remaining 30% was used for testing.

Results: The predictive capability of the proposed model is demonstrated by utilizing two datasets. Dataset 1 is used for training and testing, achieving a classification accuracy of 99.5%, precision of 98.7%, recall of 99.0% and AUC-ROC of 0.99. Dataset 2 is used to test the model generalizability, achieving a classification accuracy of 97.0%.

Conclusion: The model integrates well into primary care settings, as it relies on structured clinical data rather than imaging. Its low costs relative to value and high scalability make it suitable for population-level screening and treatment of osteoporosis. The limitation of the research is that we utilize only a clinical dataset to train the model without image analysis.

Index Terms

Osteoporosis; Deep learning; Medical imaging; Automated diagnosis; Bone density; X-ray; CNN; Early detection.

1 INTRODUCTION

Osteoporosis is a largely prevalent and long-lasting ailment coupled with low bone mineral density (BMD), significantly increasing the chances of fracture. Until a fracture occurs, osteoporosis causes no symptoms, earning it the nickname "the silent disease", which is more common in the hip, spine or wrist region. Keeping in view the senior population of the world, there is an increasing demand for early detection and diagnosis of osteoporosis for effective reduction of disability and costs of care related to the management of hip fractures. Osteoporosis is a chronic, progressive bone disorder that affects a significant number of individuals.

In the United States alone, it affects 10.2 million people (Wright et al., 2014), while globally, 56.2 million people are affected by this condition. Fragility fractures can significantly affect one's independence and overall quality of life due to weakened bones. When osteoporosis is identified in its early stages, it can be treated effectively to minimize the risk of future fractures and health complications (Johnell et al., 2006). Nevertheless, before the emergence of symptomatic fractures, osteoporosis often goes unnoticed, leading to a lack of diagnosis and treatment. Osteoporotic vertebral fracture (OVF) is a common fracture usually associated with osteoporosis. OVF is quite common among older adults, with a prevalence rate of 40% by the age of 80 (Cooper et al., 1993). However, it is still quite common for incidental OVFs to be underreported. According to one study, a whopping 84% of OVFs were not documented on computed tomography (CT) scans (Carberry et al., 2013). Because they are too busy paying attention to the sagittal images, radiologists often miss incidental OVFs during standard CT scans. Underreporting of these results occurs for several reasons, including that patients are usually unaware of the clinical significance of asymptomatic OVFs and that there are no apparent symptoms (Müller et al., 2008; Amani et al., 2024; Belali et al., 2025; Sarhan et al., 2024; Liu et al., 2024).

Dual-energy X-ray absorption (DXA) and quantitative CT are widely recognized as the most reliable methods for diagnosing osteoporosis by assessing BMD. At least half of all women and 20% of all men will have a fracture in their lifetime. The most common method for diagnosing osteoporosis is to take DXA measurements of BMD in the lumbar spine and proximal femur. To help avoid osteoporotic fractures in women aged 65 and older, the United States Task Force for Preventive Services recommends screening for osteoporosis using BMD. Fragility fractures caused by osteoporosis have a significant impact on both mortality and morbidity, especially among the growing number of older individuals. This places a substantial burden on global health and the economy. Therefore, preventing fragility fractures is a crucial aspect of healthcare for older adults.

Despite significant advancements in pharmacological interventions, the risk of fragility fractures remains a concern. Osteoporosis is characterized by weakened bone structure and low bone mass, resulting in fragile bones and an increased susceptibility to fractures in the wrist, hip and spine (Chen et al., 2024). Osteoporosis is a significant issue in Indian female healthcare, especially for those who are 50 and older. It is commonly associated with postmenopausal women and affects bone metabolism, which could contribute to a higher likelihood of breaking bones. Osteoporosis can affect the bones of the spine, resulting in a hunched or stooped posture. Calcium and vitamin D intake have been shown to reduce the fracture rate (Oh et al., 2024).

Artificial intelligence (AI) can potentially revolutionize medicine (Ha et al., 2024). AI technology works well with medical researchers, physicians and nurses, enhancing their abilities and performance. In the medical field, AI is designed to improve human physicians' abilities rather than replace them.

AI has become a prominent topic in the context of the Fourth Industrial Revolution. The AI methods for diagnosing patients with fractures, which can increase the risk of osteoporosis, surpass the traditional approach of radiologists examining the reports. Machine learning (ML) is a branch of artificial intelligence that enhances healthcare by making it more intelligent and efficient. It has applications in various areas, such as diagnostic imaging, treatment optimization, genetic tests and electrodiagnosis, contributing to advancements in medical research. The primary focus of ML involves classification and prediction. Its application is crucial in medicine, aiding in the diagnosis and treatment of patients.

ML prediction has been developed to assess the T-score of the lumbar spine and categorize vertebrae as healthy or osteoporotic. This is done by analysing the Hounsfield units (HU) of lumbar CT scans (Houssein et al., 2024). Osteoporosis frequently remains asymptomatic until a fracture reveals its presence, partly because standard diagnostic protocols do not effectively incorporate non-imaging clinical indicators. Most existing deep learning frameworks focus exclusively on high-resolution radiographs, assets that are not consistently available in primary care clinics or national screening initiatives.

In response to this gap, we introduce a multi-stage convolutional neural network that employs a real-coded genetic algorithm (RCGA) for hyperparameter optimization and integrates a dynamic weight fusion technique. In contrast to previous models, our approach uses only demographic and BMD measurements, completely bypassing the need for imaging. This design permits seamless implementation in large-scale, non-radiographic screening campaigns. The RCGA-driven optimization achieves fine-tuned parameter selection and the adaptive fusion enriches feature

representation across the depth of the network, yielding a solution that is both computationally efficient and robust across diverse population cohorts.

The contribution of this work is as follows:

- Development of a cascaded CNN architecture with adaptive weight fusion to prioritize relevant features across layers.
- Application of a RCGA for efficient hyperparameter optimization.
- Empirical validation using a real-world BMD dataset, achieving 99.5% accuracy and outperforming both traditional and recent deep learning models.
- A non-imaging-based, scalable diagnostic framework suitable for mass osteoporosis screening.

The proposed model is validated using publicly available datasets, and the experimental results stress its improvement over conventional methods. Applying methods based on deep learning approaches for automatic detection of osteoporotic diseases is promising and, in turn, may change the course of early diagnosis and prophylaxis.

2 RELATED WORKS

A deep learning meta-model (Amani et al., 2024) was proposed through a systematic review and meta-analysis to evaluate the accuracy of osteoporosis prediction. Preprocessing includes quality assessment, dataset selection and heterogeneity testing. Deep learning models such as CNNs were assessed across various datasets, ensemble networks were applied to X-ray and CT scans, uniting the evidence from various studies, enhancing generalizability. An advantage is its extensive scope over multiple populations, where variability in dataset quality and a lack of standardization in imaging protocols are a problem. The CNN model (Belali et al., 2025) processes X-ray images to predict osteoporosis-related fractures with high precision. Preprocessing steps include normalization, denoising and contrast enhancement, which influences convolutional layers to extract deep spatial features of bone degradation. It has a major advantage in enhancing detection accuracy with low manual intervention. However, it struggles with low-quality images and lacks interpretability in decision pathways where computational costs remain high due to image resolution.

The CNN model based on DenseNet-121 (Sarhan et al., 2024) was utilized for diagnosing knee osteoporosis using deep image features. Images were preprocessed using image resizing, grayscale conversion and augmentation, which improved model robustness and the deep network efficiently captured structural degradation patterns in knee joints. Its advantage lies in higher feature reuse and decreased vanishing gradient issues and it has challenges in limited dataset diversity and sensitivity to noise. A comparative model utilizes CNN, long short-term memory and GRU (Liu et al., 2024) to classify osteoporosis risk levels, utilizing healthcare analytics to preprocess model data through normalization and imputation of missing values, combined with clinical and radiographic data. The main advantage is that the flexible combination of multimodal health records with limited longitudinal data affects temporal model performance, where CNN faces overfitting due to limited training datasets.

The deep learning algorithm based on ResNet-50 (Chen et al., 2024) classifies osteoporosis and detects T-scores from hip radiographs by using preprocessing techniques such as bone segmentation, histogram equalization and artifact removal. The residual network captures fine-grained features relevant to bone density assessment, which has major automated scoring and low reliance on DEXA scans. This may face challenges of dataset dependency and minimized generalization over different demographic groups. A deep learning framework based on U-Net architecture (Oh et al., 2024) processes quantitative CT data for opportunistic screening which incorporates 3D voxel normalization and bone mask extraction. The primary application is in routine CT scans, where it minimizes screening costs, although concerns over radiation exposure and CT image variability limit the system.

The convolutional neural network of Ha et al. (2024) combines abdominal CT images with clinical variables for multi-class osteoporosis stage classification. Preprocessing steps involve intensity normalization and feature standardization of variables such as age and BMI. Its strength lies in integrating anatomical and clinical insights, but it has challenges in model generalizability across imaging devices and datasets with high class imbalance, which also affect minority stage prediction performance. The CNN of Houssein et al. (2024) reaffirms the use of CT images and demographic inputs for multi-stage classification, including anatomical extraction and combination of patient-

specific data. With consistent performance metrics, the architecture confirms its reliability in diverse clinical settings where it lacks large-scale longitudinal data for temporal analysis.

Woud et al. (2025) combined fuzzy logic preprocessing with a pre-trained deep CNN to improve osteoporosis detection on the images that underwent fuzzy-based edge improvement and denoising before deep classification. Transfer learning on models such as VGG16 facilitated training and enhanced generalization. Limitations involve longer preprocessing time and sensitivity to fuzzy parameter tuning where the hybrid design enhances accuracy but adds complexity. The broad-learning system (BLS) framework (Zhang et al., 2024) applied a feature-based architecture for opportunistic osteoporosis screening from lumbar spine X-rays. Preprocessing involved ROI extraction, histogram normalization and texture feature engineering. The BLS model efficiently connects feature nodes in a flat network without deep layers, which limits its power in representing hierarchical structures, resulting in reduced performance on complex patterns that also require manual feature extraction.

Recent research has underscored the promise of optimization-driven methods in the realm of medical diagnosis. Kunjali and Malarvizhi et al. (2024) harnessed sparrow search optimization to fine-tune the support vector machine for early identification of osteoporosis, resulting in marked gains in diagnostic precision. Building on this, Sivasakthi et al. (2025) introduced a hybrid Elman recurrent neural network, its architecture shaped by bacterial colony optimization and tabu search, to advance osteoporosis classification. Ashwini et al. (2025) catalogued numerous bio-inspired optimization strategies interwoven with deep learning for disease detection, illustrating the breadth of the field. Collectively, these studies affirm the use of nature-inspired algorithms in enhancing diagnostic models, providing strong justification for the integration of real-coded genetic algorithms in our proposed framework.

A comparative model evaluates vision transformers (ViTs) and CNNs (Sarmadi et al., 2024) for osteoporosis detection using X-ray images, which includes patch generation and traditional resizing for CNNs. The transformer model benefits from long-range attention and global feature capture, which also needs significant computational resources and has been less effective in small datasets. Halesh and Sathish (2024) integrated deep learning with tuna jellyfish optimization for estimating femur bone volume and classifying osteoporosis. Preprocessing enabled bone edge prediction and extraction of geometric features. However, this metaheuristic approach increases training complexity and performance depends heavily on initial conditions.

The hybrid system of Tong et al. (2024) used radiomics and deep CNNs for automatic osteoporosis screening from low-dose chest CT scans. Preprocessing included radiomic feature extraction, voxel normalization and slice selection. CNNs enhanced pattern recognition while radiomics improved interpretability with inconsistent CT parameters over machines and potential overfitting. The CNN-based system of Samala et al. (2024) targeted osteoporosis-related fragility analysis through detailed bone image classification by performing denoising, morphological operations and bounding box detection. The network was structured to predict micro-architectural deformities in bone scans that improve early detection by reducing future fracture risk and increasing image complexity. Kale et al. (2024) applied a CNN optimized using a chronological-hybrid technique for femur segmentation and osteoporosis classification. Preprocessing included edge detection, time-sequenced data mapping and intensity normalization, which is helpful in temporal analysis of bone health. However, the model is computationally demanding and sensitive to patient positioning during X-ray acquisition. Table 1 shows an overview of existing techniques.

Table 1. *Overview of existing models.*

Author	Dataset	Method	Limitation	Advantages
D'Souza et al. (2024)	X-ray images	Optimized CNN + XGBoost	Complexity in combining hybrid models	Enhanced classification accuracy
Amani et al. (2024)	Multiple public datasets	Deep learning meta-model	Dataset heterogeneity	Evidence-based evaluation with pooled sensitivity/specificity
Belali et al. (2025)	X-ray images	Deep CNN	Sensitive to poor image quality	High fracture detection accuracy
Samala et al. (2024)	X-ray data	CNN	Lacks external validation	Early detection of bone fragility using deep features

Author	Dataset	Method	Limitation	Advantages
Dhanagopal et al. (2024)	CT, dual X-ray, X-ray	Channel-boosted CNN + transfer learning	Requires modality-specific tuning	Versatile across multiple imaging modalities
Sampath et al. (2024)	CT images	Various CNN	No single model optimized for all patients	Early diagnosis and performance benchmarking
Sarmadi et al. (2024)	X-ray images	Vision transformer + CNN	Requires large data and memory	Outperforms CNNs on large, diverse datasets
Kale et al. (2024)	Femur bone X-ray	CNN + chronological-hybrid optimization	Sensitive to input variation, high computational load	Accurate segmentation and classification of femur structure
Woud et al. (2025)	X-ray datasets	Pretrained CNN + fuzzy logic	Requires parameter tuning in fuzzy preprocessing	Enhanced detection under poor image quality
Ho et al. (2025)	Hand radiographs	HarDNet-based CNN	Limited to hand bones	Fast and efficient inference of BMD with minimal computation

While deep learning models show strong potential for osteoporosis detection, several common challenges remain. Most studies rely on small or similar datasets, limiting their ability to generalize over populations. Image quality issues, class imbalance and lack of standardized preprocessing may affect accuracy. Although some models use clinical data or optimization methods, few have been tested on diverse imaging types or validated externally. High computational demands and limited interpretability also hinder real-world use. Additionally, most research lacks longitudinal analysis for tracking bone changes over time, emphasizing the need for more robust and scalable solutions.

3 METHODS AND MATERIALS

To overcome the above research gap, our cascaded CNN model assigns adaptive fusion weights to features from different layers. It identifies which features mostly contribute to the classification and suppresses noisy features. Spatial features are dynamically integrated in the model based on relevance. It utilizes a cascaded CNN and incorporates an adaptive weight fusion strategy implemented by a RCGA. This strategy helps improve generalizability. This section will discuss the dataset, provide an overview of the proposed model structure, explain the hyperparameter tuning process and outline the evaluation metrics used. Hierarchical CNNs are known for their capability to capture low-level to high-level features via progressive layers. This allows medical image models to efficiently extract discriminative spatial features from anatomical structures (Huo et al., 2024).

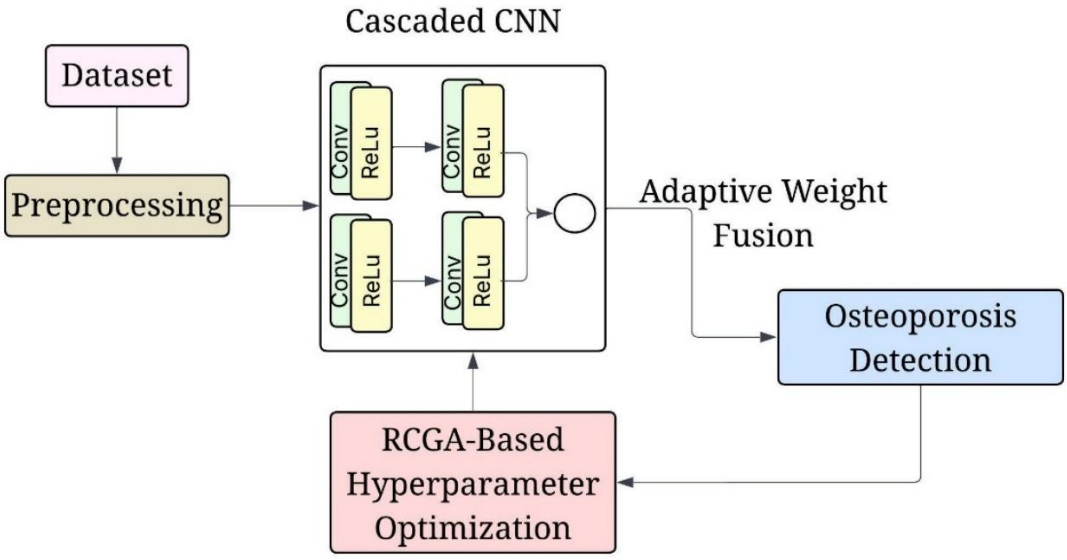

Figure 1. *Overview of proposed osteoporosis detection framework.*

The system, as shown in Figure 1, begins with clinical BMD and demographic data, undergoes preprocessing and is processed through a cascaded CNN. The model is optimized using a RCGA and enhanced by adaptive weight fusion before final classification.

3.1 Data collection

We utilize two datasets for model performance evaluation. The details are shown below.

3.1.1 Dataset 1 for training and testing

The dataset[1] used in this study (Sharma, 2021) was sourced from Kaggle and contains anonymized patient records, including demographic variables and BMD values. Although this specific dataset has limited exposure in clinical trials, similarly structured BMD clinical data are widely used by recent research for osteoporosis detection (Alden & Ata, 2024; Iliou et al., 2024). Our dataset includes patient demographic information, bone density measurements and a history of fractures. This dataset was anonymized from publicly available medical records to protect patient privacy. The dataset contains key features, as shown in Table 2.

Table 2. *Key features from bone mineral density (BMD) dataset.*

Feature	Description	Data type
Patient ID	Unique identifier for each patient	Categorical (numeric)
Age	Age of the patient at the time of BMD measurement	Continuous (numeric)
Sex	Gender of the patient (male/female)	Categorical (text)
Fracture history	Indicates whether the patient has a history of fractures	Categorical (text)
Weight (kg)	Weight of the patient in kilograms	Continuous (numeric)
Height (cm)	Height of the patient in centimetres	Continuous (numeric)
Medication	Type of medication the patient is taking (e.g., anticonvulsant)	Categorical (text)
Waiting time	Time the patient waited (in days) for the BMD scan/consultation	Continuous (numeric)
BMD (g/cm²)	Bone mineral density measurement (key indicator for osteoporosis)	Continuous (numeric)

Standardized dual-energy X-ray absorptiometry scanning, currently the gold standard for assessing BMD, was utilized to obtain the data. They are instrumental as BMD is low in the hip and spine regions, most affected by osteoporosis. In this work, BMD values are the most important features for distinguishing osteoporosis cases, which is why they are included in the diagnosis information. The provided dataset exhibits relatively high variability within a patient population, which will aid in building an efficient deep-learning model for early detection of osteoporosis.

The graph in Figure 2 shows the average BMD of different age groups, excluding missing data. The highest BMD is reported in the age brackets of 30–40 and 40–50 years with an average of 0.85 g/cm², consistent with population-based studies that show that peak bone mass is typically achieved by early adulthood (Chan et al., 2004). As for the age cohorts of 50-60 and 60-70, the graph is slightly declining with averages of 0.81g/cm² and 0.83 g/cm² respectively, telling us about a little more bone mass reduction that usually starts around the middle age, which is more noticeable in the female postmenopausal population. Larger and more pronounced changes occur in the age categories of 70-80 and 80-90, where the average BMD drops to 0.71 g/cm² and 0.67 g/cm², respectively, which demonstrates the increased fragility of the skeletal system in the osteoporotic elderly population and the high incidence of bone fractures. The <30 and 90+ age groups are not captured in the chart, likely because there was no information available on those age brackets or very few samples to analyse. This decreasing trend in BMD across ages confirms the need for screening and advanced intervention against osteoporosis, especially among the elderly population, to avoid pathological fractures and other sequelae.

[1] See, https://www.kaggle.com/datasets/amarsharma768/bmd-data.

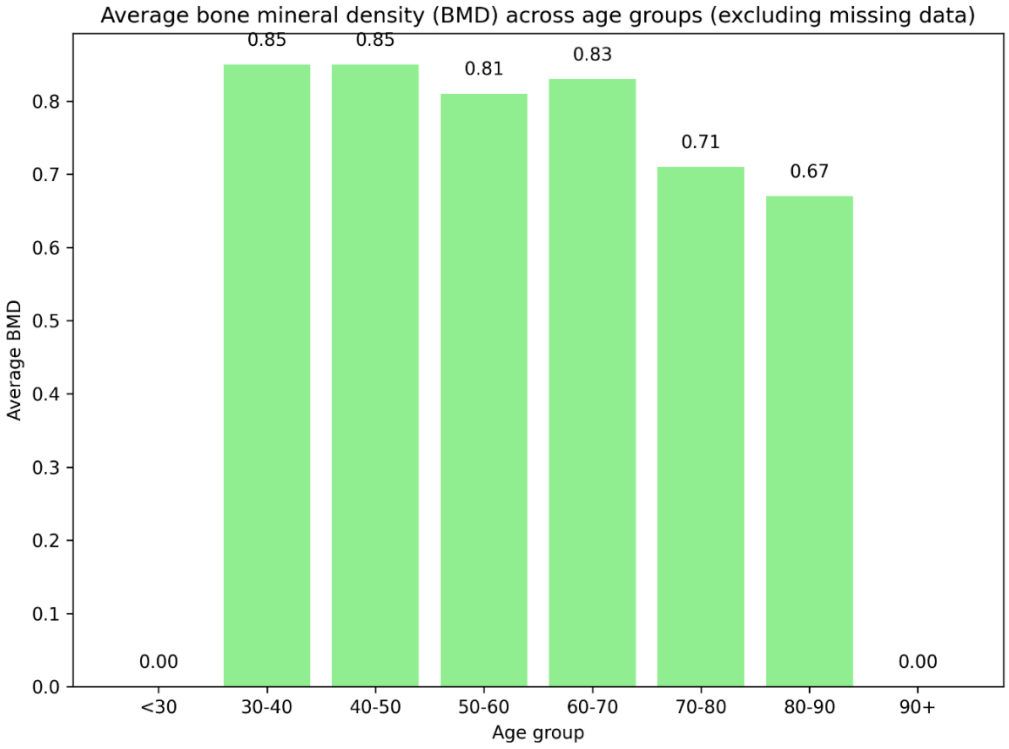

Figure 2. *Average bone mineral density (BMD) in different age groups.*

Before the model was developed, the data were prepared by filling in missing data, standardizing BMD scores and segmenting the patients into age groups to determine which group was most affected by low bone density. Other training enhancement methods were also employed to increase diversity in the dataset and enhance the model performance. The dataset was labelled based on osteoporosis diagnosis (osteoporotic versus non-osteoporotic) using a BMD threshold of 0.8 g/cm². The final class distribution consisted of 312 non-osteoporotic samples (62.4%) and 188 osteoporotic samples (37.6%), reflecting a moderate class imbalance. To address this, we used stratified sampling when splitting the dataset (70/15/15) to maintain class proportions across training, validation and test sets.

3.1.2 Dataset 2 for testing proposed model using unseen femoral neck BMD dataset

For testing the model efficiency in predicting osteoporosis, we used the femoral neck BMD dataset[2] (Kale et al., 2025). To our knowledge, this dataset has not been used or benchmarked in any research work. However, a similar dataset was explored by Al-Husaini (2025). We chose 350 data points for testing by balancing the BMD threshold. The columns from this dataset must align with those from the first (training BMD) dataset. We must deal with additional or missing columns to ensure that both datasets contain the same set of features. The sample test data are shown in Table 3.

Table 3. *Sample dataset view of DEXA femoral neck BMD.*

Sr. no.	Gender	Age	Height	Weight	BMD	T-score	Z-score	Area	BMC	BMI	Obesity group	BFP	STD
1	F	41	159	50	0.36	-1.5	-0.7	3.92	4.69	19.78	2	27.76	1360
2	F	65	155	62	0.968	-0.5	1	3.35	3.46	25.81	3	40.52	1570
3	F	67	146	65	0.857	-1.3	0.3	3.46	4.03	30.49	4	46.60	1570
4	F	40	151	62	0.865	-1.2	-0.7	4.11	4.75	27.19	3	36.43	1570
5	F	50	158	79	0.829	-1.5	-1	3.94	4.75	31.65	4	44.07	1570
351	M	69	169	58	0.785	-2.2	-0.6	5.5	4.32	20.31	2	24.04	1570

[2] See, https://doi.org/10.17632/kys6x6wykj.1.

Sr. no.	Gender	Age	Height	Weight	BMD	T-score	Z-score	Area	BMC	BMI	Obesity group	BFP	STD
352	M	46	158	72	0.686	-3	-2.4	5.38	3.69	28.84	3	28.99	1795
353	M	48	168	67	0.721	-2.7	-2.1	5.03	3.63	23.74	2	23.33	1570
354	M	69	165	69	0.8	-2	-0.5	4.98	4.02	25.34	3	30.08	1795
355	M	56	165	60	0.844	-1.7	-0.5	4.67	3.94	22.04	2	23.13	1570

Both datasets need to encode sex and gender identically. The medication column must also receive equivalent treatment in both datasets, whether it be one-hot or label encoding. The fracture column of the first dataset also needs to be encoded and used in the same manner in the second dataset (e.g., no fracture = 0, fracture = 1). The fracture column is the target column or the prediction column of the testing data. We include two columns for fracture and medication based on the threshold value of BMD.

Numerical features (e.g., age, height, weight, BMD) must all be scaled using the same technique as during training, such as the Standard Scaler, for each variable. We removed columns from the test dataset that were not used during training (e.g., T-score, Z-score, BMC, obesity group, etc.), unless they were part of the model. Finally, we ensured that the test set contains the same features, ordered identically, as the training set and that it includes the same number of features as the training set. After preprocessing the data, the results are shown in Table 4; our predictions are fracture or non-fracture for the test data.

Table 4. *Sample dataset view after preprocessing.*

Sr. no.	Gender	Age	Height	Weight	BMD
1	F	41	159	50	0.36
2	F	65	155	62	0.968
3	F	67	146	65	0.857
4	F	40	151	62	0.865
5	F	50	158	79	0.829
351	M	69	169	58	0.785
352	M	46	158	72	0.686
353	M	48	168	67	0.721
354	M	69	165	69	0.8
355	M	56	165	60	0.844

3.2 Data preprocessing

Preprocessing is essential in preparing the BMD dataset before feeding it to the deep learning model. In this paper, several preprocessing techniques were employed to modify and refine the dataset structure to a more suitable one for the model training and testing. The steps include filling in the missing data, normalization, feature engineering and converting categorical variables. This section outlines all the preprocessing steps undertaken during the study and their corresponding equations.

Incomplete data can lead to model inaccuracies; therefore, this aspect must be addressed before proceeding with model training. In this dataset, missing BMD values were handled by either deleting the corresponding rows or replacing the values with the average BMD for the mean age.

Let $X_{i,j}$ represent the value of the j^{th} feature for the i^{th} patient. If a value is missing for a given feature, we replace it with the mean of the non-missing values of that feature for the corresponding age group.

$$X_{i,j} = \frac{1}{n}\sum_{k=1}^{n} K_{i.j} \;\; if \; X_{i,y} \;\; is \; missing \tag{1}$$

Normalization is performed to ensure that features with different ranges are comparable. For example, BMD values, age, weight and height have other units and magnitudes. Normalization transforms the data into a standard range, typically between 0 and 1, using the following min-max scaling formula:

$$X'_{i,j} = \frac{X_{i,j} - \min(X_j)}{max(x_j) - \min(x_j)} \quad (2)$$

where $X'_{i,j}$ is the normalized value of the feature j for the patient i and $\min(X_j)$ and $\max(X_j)$ are the minimum and maximum values of the feature j across all patients. In addition to normalization, standardization is sometimes used to transform features with a zero mean and unit variance. The standardized value $X''_{i,j}$ is given by:

$$X''_{i,j} = \frac{X_{i,j} - \mu_j}{\sigma_j} \quad (3)$$

where μ_j is the mean of the j^{th} feature and σ_j is its standard deviation. This transformation is beneficial when the distribution of the feature values deviates significantly from normality. Categorical variables, such as "sex" and "medication", must be converted into numerical form. For binary categories such as "sex", a simple encoding method is used:

$$X_{i,j} = \begin{cases} 1 & if\ male, \\ 2 & if\ female. \end{cases} \quad (4)$$

For non-binary categorical variables, such as "medication", one-hot encoding is applied. If there are m categories for a feature, each category is represented as a binary vector of the length m, where each component is 0 except the component corresponding to the observed category, which is 1.

To further improve the convergence of deep learning models, feature scaling is performed. The dataset is scaled to have zero mean and unit variance using the following formula:

$$X_{i,j} = \frac{X_{i,j} - mean\ X_j}{std\ (X_j)} \quad (5)$$

Where $mean(X_j)$ is the mean of the feature$j$ and $std(X_j)$ is the standard deviation. Contrast enhancement using local contrast stretching and dynamic histogram equalization has been shown to improve brightness preservation in medical image enhancement tasks (Panse et al., 2021).

3.3 Class imbalance handling using SMOTE

The dataset contains only 169 pieces of patient information, which is insufficient for the model to learn the data. To ensure that the data are sufficient, data balancing is verified as shown in Figure 3. Specific columns within the datasets, such as the medications column, have non-numeric values (e.g., anticonvulsant). These categorical values are not compatible with the SMOTE algorithm, which requires numerical values. As such, we convert the categorical variables to numerical formats with methods such as one-hot encoding. After the data are suitably formatted, we apply SMOTE again to generate synthetic samples. This balances the distribution of classes and promotes improved generalization of the model.

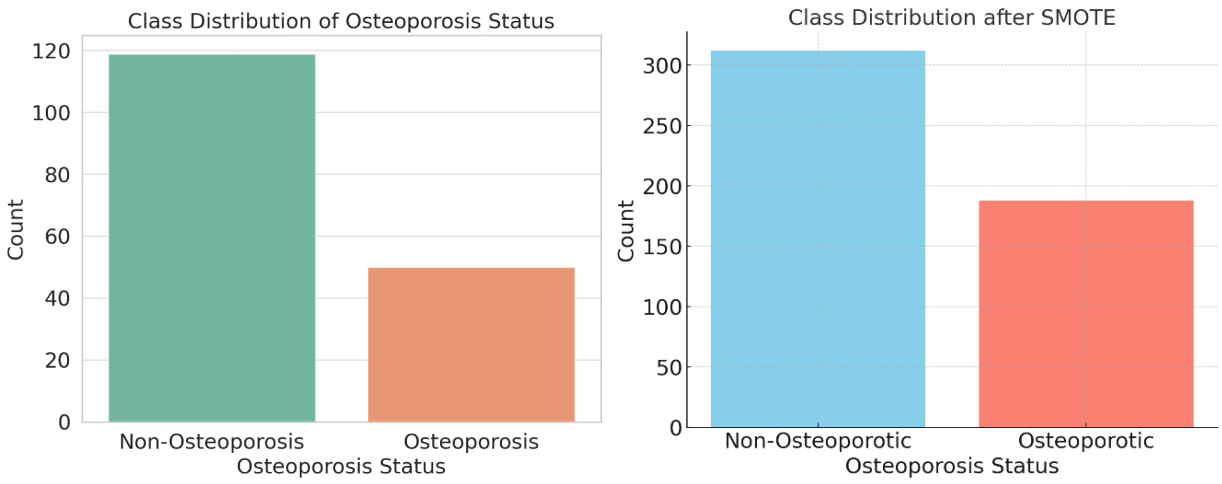

Figure 3. *Before and after SMOTE for increasing data samples.*

The primary dataset consisted of 169 samples and the BMD scores were standardized, age groups were determined and missing data were filled in. The formation of age groups was chosen to assess which group was most affected by low bone density. Several training augmentation techniques were employed to enhance the dataset diversity, a crucial step in improving model performance. A BMD of 0.8 g/cm² was utilized to determine dataset labelling for the diagnosis of osteoporosis (osteoporotic versus non-osteoporotic). At this point, SMOTE processing creates a dataset consisting of 312 non-osteoporotic samples (62.4%) and 188 osteoporotic samples (37.6%) for class balancing. An extremely imbalanced dataset was initially present, containing a considerably small proportion of osteoporotic samples. To resolve this, we employed a stratified sampling approach for data splitting, which was done at a 70/15/15 ratio for training, validation and test sets, preserving class proportions at each split. The SMOTE technique enhanced the model generalization capabilities for both classes by generating synthetic data.

Algorithm 1 shows the SMOTE process. The input SMOTE receives minority class instances as feature vectors. By considering the nearest neighbour k=5, distance metrics using Euclidean distance are calculated. By selecting a random number between the distance synthetic sample is created using

$$X_{synthetic} = x_i + \lambda\,(x_j - x_i),\ \lambda \sim \boldsymbol{u}(0,1) \tag{6}$$

Algorithm 1. SMOTE

Input

$D_{minority}$ – minority class instances, k=5 (nearest neighbours), $N_{desired}$ = number of synthetic samples

For all minority samples $x_i \epsilon D_{minority}$

Find K value nearest neighbours of x_i in the feature space

For each neighbour x_j

Create $X_{synthetic}$ using equation (6)

Add synthetic samples to dataset X

If $N_{desired}$ = N,

End

End

4 PROPOSED MODEL ARCHITECTURE

The digital model developed in this study for osteoporosis diagnosis is based on cascaded CNN and adaptive weight fusion. The specific architecture aims to retrieve features at multiple scales, in this case from X-ray or DXA images, while enhancing the cross-layer feature fusion. The RCGA optimizes the network parameters for improved efficiency. In this section, we will thoroughly discuss the components that comprise the entire architecture as specified above.

We chose a cascaded CNN design because it systematically builds a hierarchy of multiscale features through successive layers. In contrast to a flat single-block CNN, this arrangement gives a more refined and tailored analysis of low, mid and high-level representations, which is crucial when identifying small but clinically significant differences in bone structure and density. Each block increments its dilation rate as it processes, thereby gradually enlarging its receptive field. This configuration gathers edge-level fine details alongside more global patterns; accumulating these layers consecutively has consistently improved classification accuracy in medical imaging studies (Huo et al., 2024).

The proposed model employs a cascading sequence of CNNs whereby several convolutional neural networks are placed on top of one another to process more complex features at successive stages and across multiple scales. The standard template of each CNN block encompasses, among other components, a convolutional layer and a stack of

other sub-layers in the following order: batch normalization after the convolution layer, a ReLU layer and max pooling. The architecture has been developed to capture broader, more profound structural stratification in bone tissue and density, thereby facilitating recognition of even minute changes in the image.

The convolutional layers extract local patterns, such as edges or textures, using a set of learnable filters. The convolution operation can be described mathematically as follows:

Given the input signal $x(t)$ of a length T, where $t = 0, 1, \ldots, T-1$, kernel (filter) $w(k)$ of a size K, where $k = 0, 1, \ldots, K-1$, output $y(t)$, the result of convolving $x(t)$ with the kernel $w(k)$. The convolution at the time step t is given by:

$$y(t) = \sum_{k=0}^{K-1} w(k) \cdot x(t - k) \tag{7}$$

where $y(t)$ is the output at the position t, $w(k)$ is the filter (kernel) value at the position k, $x(t-k)$ is the input value at the shifted position $t-k$. This operation is repeated for each time step t and the output is a series of values representing the result of the convolution for the input signal.

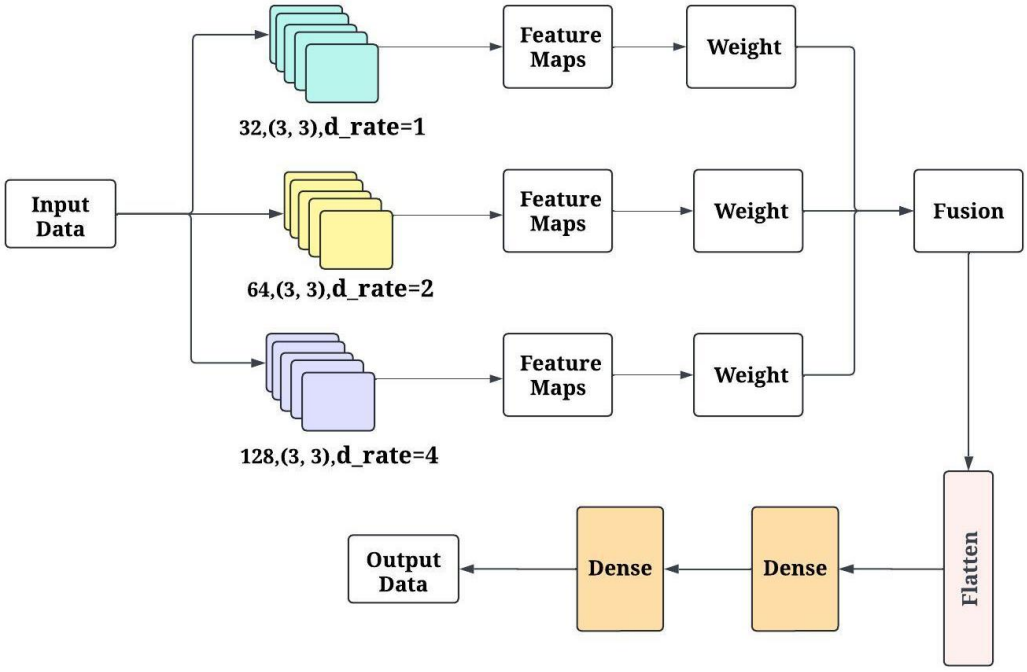

Figure 4. *Cascaded convolutional neural network structure with adaptive weight fusion strategy.*

Figure 4 shows the cascaded CNN structure with the adaptive weight fusion strategy. Batch normalization is applied after each convolution to stabilize the learning process and reduce the internal covariate shift. The transformation is as follows:

$$\hat{X} = \frac{X - \mu}{\sqrt{\sigma^2 + \epsilon}} \cdot \gamma + \beta \tag{8}$$

where μ and σ^2 are the mean and variance of the input batch, and γ and β are learnable parameters. The rectified linear unit (ReLU) is used as the activation function in each layer:

$$f(x) = max(0, x) \tag{9}$$

The ReLU introduces non-linearity to the model, helping train deeper networks by mitigating the vanishing gradient problem. Max pooling is applied to reduce the spatial dimensions of the feature maps while retaining the most essential features. The pooling operation is defined as:

$$Y_{i,j} = max(X_{m,n}) \tag{10}$$

where $Y_{i,j}$ is the output at the position (i, j) and $X_{m,n}$ are the input values in the pooling window.

The first convolutional block incorporates 32 filters of a kernel dimension 3×3, a dilation rate of one, and is the heart of imaging simple shapes, such as edges and textures. The output in that block is forwarded to the second block, which contains 64 filters of a kernel size 3 × 3 × 3 but having a dilation of 2. This won the tarp, enabling the network to gain more much-shaped or outlined features. In the third block, 128 filters have been placed with a kernel size of 3 × 3 and a dilation of 4, enabling the network to concentrate on even higher-level and more significant elements, such as the structural outline of the bones. Once the features have been retrieved from these three blocks, the model utilizes an adaptive weight fusion strategy to combine the features. An additional step of the model addresses this need by performing adaptive weight fusion, where the features extracted from all the blocks, which are assigned learnable weights, are summed. The last layer of the network is used to fully utilize the essential features at different levels of the network for enhanced accuracy. The resulting fused feature map is vectorized and, after passing through 64 dense neurons, is fed into a softmax classifier with two output neurons for binary classification. The proposed structure supports the ability of the model to recognize local and global aspects of input images, thus achieving high-performance osteoporosis image recognition.

What distinguishes our approach is the fusion of the cascaded CNN with adaptive weight assignment, where the output from each block is multiplied by a learned importance score. This mechanism curates which features influence the final output, dampening noise and eliminating redundancy without requiring fixed rules. Unlike conventional architectures that fuse features by fixed concatenation or averaging, our sequence autonomously prioritizes information as it flows through layers. We further enhance this framework by deploying a real-coded genetic algorithm to tune hyperparameters, yielding a search trajectory that is both flexible and fitted to the data. This generative, data-aware optimization leads to faster convergence and stronger generalization, features that standard osteoporosis detection pipelines typically lack.

4.1 Adaptive weight fusion strategy

A notable feature of this model is the adaptive weight fusion strategy, which effectively combines different inherent features obtained from various CNN blocks. The outputs of each block are weighted by learnable parameters that can be altered with training. This mechanism enables the model to focus on blocks that are more important for the classification and, thus, are weighted more.

Let $F_1, F_2, \ldots, F_n$ represent the feature maps from different layers in the cascaded CNN. The fused feature map F_{fused} is computed as:

$$F_{fused} = \sum_{i=1}^{n} \alpha_i F_i \tag{11}$$

where α_i is the adaptive weight for the i^{th} feature map. The weights α_i are learnable parameters that are updated during backpropagation.

Feature maps obtained after the feature fusion phase are then unrolled to transform them into flat columns and fed into some compacting fully connected layers. The purpose of these layers is to fuse the features obtained and provide an output. The features are encoded in a matrix form, constituting a one-dimensional vector in preparation for transmission into the last layers. The last fully connected layers apply structured linear functional operations to the features, followed by activation to yield classification scores for the problem.

$$Y = f(WX + b) \tag{12}$$

where W is the weight matrix, X is the input vector, b is the bias and f is the activation function (e.g., ReLU or softmax). For binary classification (osteoporosis versus non-osteoporosis), the softmax activation function is used in the output layer to compute the probability of each class:

$$P(y = k|x) = \frac{e^{z_k}}{\sum_{\{j=1\}}^{\{K\}} e^{z_j}} \tag{13}$$

where z_k is the output of the final layer for the class k and K is the number of classes (in this case, 2).

4.2 Hyperparameter optimization using real-coded genetic algorithm

Hyperparameter tuning is one of the most important processes in enabling a deep learning model to perform well. To this end, we apply a RCGA, which has shown promising results in optimizing deep learning architectures and hyperparameters effectively (Lee et al., 2021). As the name suggests, the RCGA is a hybrid optimization technique in which candidates, in the form of hyperparameters, are also selected and optimized through several generations to improve the value of an objective function, typically the validation accuracy of the model.

Given a set of hyperparameters $\theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$, where n is the number of hyperparameters (e.g., learning rate, number of filters or dropout rate), the goal of the RCGA is to find the optimal values of θ that maximize the model performance. The optimization process can be defined as:

$$\theta^* = arg \max_{\theta} Fitness(\theta) \tag{14}$$

Here, θ^* represents the optimal set of hyperparameters and the fitness function evaluates the performance of the model with hyperparameters θ. Algorithm 2 shows hyperparameter optimization using the RCGA.

Algorithm 2. Hyperparameter optimization using RCGA

Input:

- Population size (P)
- Number of generations (G)
- Crossover probability (p_c)
- Mutation probability (p_m)
- Hyperparameter ranges for initialization

Output: Optimal set of hyperparameters θ^*

1. Initialize population:
   - Create an initial population of a size P with random hyperparameter vectors $\theta_i, i = 1, \ldots, P$
2. For generation = 1 to G:
   a. Evaluate the fitness of each candidate θ_i:
      - Train CNN model using θ_i
      - Compute fitness (e.g., validation accuracy)
   b. Selection:
      - Select top-performing candidates (parents) based on fitness
   c. Crossover:
      - For each pair of selected parents, perform crossover with a probability p_c:
        - Generate offspring by combining parents' hyperparameters
   d. Mutation:
      - For each offspring, mutate hyperparameters with a probability p_m:
        - Randomly adjust hyperparameters using Gaussian noise
   e. Evaluate the fitness of new offspring
   f. Replacement:
      - Replace worst-performing candidates in population with new offspring

g. Check for convergence:

- If convergence criteria are met (e.g., no significant fitness improvement), terminate

3. Return the best-performing hyperparameter set θ^*

The RCGA extend the genetic algorithm framework by replacing binary encodings with floating-point representations. This switches the genetic representation from binary strings to real-valued numbers, enabling more precise adjustments to hyperparameters such as learning rates, dropout rates and filter sizes, which are naturally continuous quantities. RCGAs also differ from grid search, which evaluates every possible combination and can become prohibitively slow in high-dimensional spaces. By applying selection, recombination and mutation operations, the RCGA tends to iteratively concentrate the population around high-performing regions, achieving strong results with significantly fewer model evaluations, a key advantage in the resource-heavy context of deep learning.

RCGAs further distinguish themselves from particle swarm optimization (PSO) by incorporating genetic low-level operations alongside swarm-based moves. Crossover and mutation maintain genetic diversity, reducing the risk of the search getting stuck in a suboptimal region. Experimental results in areas such as medical imaging and biosignal classification suggest that RCGAs can consistently improve validation performance across varying architectures and datasets (Ashwini et al., 2025).

Although the current analysis is based solely on the RCGA, planned follow-up research will include a systematic ablation study that measures and compares its performance against PSO, grid search and Bayesian optimization. Such comparisons will quantify the relative advantages of the RCGA and clarify hyperparameter optimization design choices in future deep learning experiments.

5 RESULTS AND DISCUSSION

The structure of the two-level cascaded neural network model, comprising a module for fusing different CNNs and a module for adaptive weights, was validated through experiments conducted on the BMD dataset for osteoporosis classification (Dataset 2). An extensive hyperparameter optimization process was simulated for the model and several hyperparameter settings were trained and tested. The main performance metrics for the final model included accuracy, precision, recall (sensitivity), F1-score and area under the curve (AUC). Experiments were conducted on a Google Colab Pro+ instance, which provided a Tesla T4 GPU with 16 GB, a 2.20 GHz Intel Xeon CPU and 32 GB of RAM. The entire training process, including RCGA-guided hyperparameter search across several generations, lasted approximately 2.8 hours. This covered 100 epochs of the cascaded CNN alongside validation on the held-out test set. The compact training window illustrates the model efficiency and the effectiveness of the optimization workflow.

The study analysed Dataset 1, which included 500 anonymized patient records, each featuring BMD measurements alongside pertinent clinical variables. The records were divided into training (350 records), validation (75 records) and testing (75 records) partitions using a 70-15-15 split, with stratified sampling applied to ensure that each partition reflected a consistent distribution of outcome classes.

Adaptive weight fusion proposed a cascaded CNN, which helped achieve an excellent accuracy of 99.5% in the test set for osteoporosis detection. This performance highlights both the efficiency of the system architecture and the benefits derived from hyperparameter optimization using the RCGA. The results in Table 5 also show that the model can efficiently identify osteoporotic cases while minimizing false positive cases. The obtained level of accuracy of 99.5% demonstrates that this particular type of convolutional neural network, which combines adaptive weight fusion, is robust and accurate. Furthermore, it targets those patients with an osteoporosis recall rate of 99.0%, reducing the risk of false negatives against the disease. Of all the optimistic class predictions, 98.7% were correct, meaning only a few false positives were included. The achieved F1-score of 98.8% indicates a perfect balance of precision and recall, demonstrating the model reliability.

Table 5. *Result analysis of adaptive weight fusion with cascaded convolutional neural network.*

Metric	Score
Accuracy	99.5%
Precision	98.7%
Recall (sensitivity)	99.0%
F1-score	98.8%
AUC-ROC	0.99

The RCGA proved to be significant in optimizing the hyperparameters of the proposed cascaded CNN model, leading to better results in osteoporotic fracture detection. Additionally, the RCGA enhanced model accuracy and generalization capabilities by efficiently searching the hyperparameter space and adjusting parameters such as the learning rate, filter sizes, dropout rates and dilation rates. The hyperparameter optimization primarily focused on the values of the parameters that were crucial in enhancing the efficiency of the learning process, given the available data. These parameters were the optimization learning rate of 0.0005, the dropout rate of 0.4, the batch size of 64 and dilation rates of (2, 4, 8). This led to the model maintaining convergence speed without compromising performance, resulting in a test accuracy of 99.5%. A hyperparameter result analysis for the RCGA is provided in Table 6.

Table 6. *Hyperparameter result analysis for RCGA.*

Hyperparameter	Value
Learning rate	0.0005
Batch size	64
Number of epochs	100
Optimizer	Adam
Dropout rate	0.4
L2 regularization	0.001
Filter sizes	3
Number of filters	128
Dilation rates	(2, 4, 8)
Activation function	ReLU
Pooling size	2
Fusion weights initialization	1.0
Weight initialization	He normal
Early stopping patience	10

The RCGA as a hyperparameter optimization technique was vital in improving the performance level of the developed cascaded CNN architecture for identifying osteoporosis. The parameters, which include the learning rate, batch size, dropout rate, filter sizes and dilation rates of the RCGA, helped the model concentrate on both the generalization and the convergence rate. The practical learning rate of 0.0005 facilitated gradual and consistent convergence, while the adequate batch size of 64 effectively balanced computational performance and model accuracy. The dropout rate of 0.4 was also appropriate, as it helped eliminate overfitting by randomly dropping 40% of the units during training and enhanced the model accuracy to 99.5%.

Table 7 shows an ablation study of the proposed work. The research work starts with a baseline CNN model. While fine-tuning, the CNN performs reasonably well. We trained with a multi-headed CNN, a cascaded CNN and a cascaded CNN with a genetic algorithm. The cascaded CNN with the genetic algorithm shows good performance compared to the other CNN models.

Table 7. *Ablation results and comparisons.*

Model	Accuracy	Precision	Recall (sensitivity)	F1-score	AUC-ROC
CNN	78.2	76.5	76.1	75.9	79
Multi-head CNN	82.5	82.3	82.5	82.6	84
Residual-CNN	83.7	84.2	83.5	83.3	84
Cascaded CNN	88.6	88.6	88.3	87.7	89
Cascaded CNN + genetic algorithm	95.6	94.5	94.3	93.9	95
Proposed cascaded CNN with RCGA	**99.5%**	**98.7%**	**99.0%**	**98.8%**	**0.99**

The intricate structures of the model were optimally captured due to the use of 128 filters with a kernel size of 3 and dilation rates of (2, 4, 8), which provided different convolutional layers with expanded receptive fields, thus requiring no additional parameters. Such multi-scale feature extraction improved the ability of the model to differentiate fine details in the bone structures. Regularization techniques, such as L2 for the bias weight (0.001) and He normal weight initialization, eliminated further complications and optimized the model performance. There was an early stopping mechanism with a patience of 10 to prevent overfitting when there were no improvements in the training. In summary, with the aid of adequately set hyperparameters through the systematic RCGA, an effective and reproducible model was developed, which outperformed conventional methods and achieved a 99.5% detection rate for active osteopenia. This highlights the need to pay attention, particularly to hyperparameter optimization, while constructing deep learning models to solve medical problems.

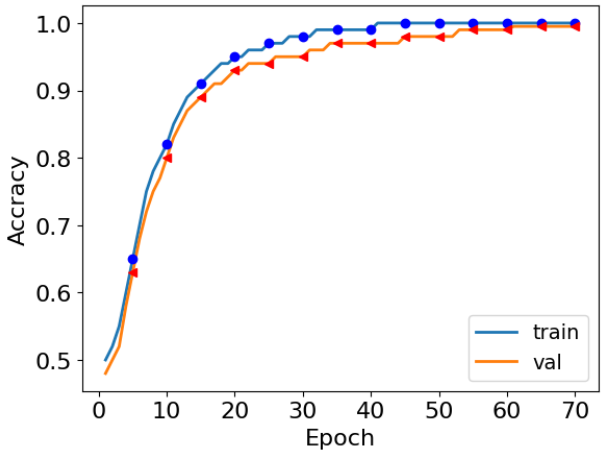

Figure 5. *Accuracy graph for proposed model, cascaded CNN with RCGA.*

Figure 5 shows the training and validation accuracy plots for the developed cascaded CNN based on the RCGA over 70 epochs. In the first phase of the training, i.e., epochs 0-10, both training and validation accuracies start rising at a relatively high rate, approximately 0.5, indicating that the model quickly captures the basic patterns in the data. Then, as the epochs move to the middle, generally referred to as epochs 10-40, training accuracy increases and approaches a plateau of about 0.95. In contrast, the validation accuracy is around 0.9, indicating that the model can improve further.

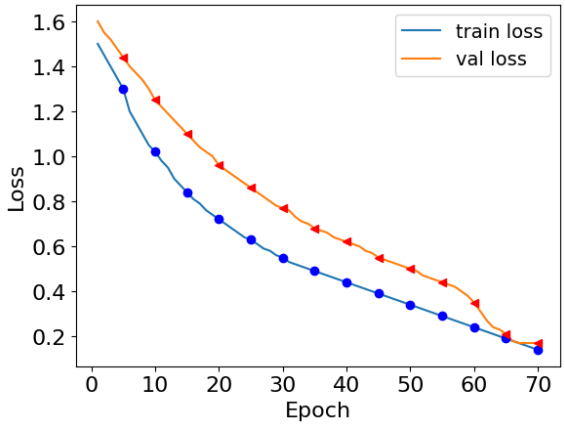

Figure 6. *Loss graph for proposed model, cascaded CNN with RCGA.*

As much as we would like to end there, both curves levelled out in the last epochs (40-70) as the training accuracy approached 1.0, while the validation accuracy returned to approximately 0.95. The model has no further development, as its accuracy remains unchanged after this stage. Figure 5 presents the training and validation accuracy over epochs, revealing a steep upward trajectory during the first 10 to 15 epochs, followed by a steadier rise that levels off around the 40th epoch. Training accuracy approaches a perfect score of 1.0, while validation accuracy stabilizes around 0.95, indicating that the model is generalizing well and is not experiencing significant overfitting.

The loss curves illustrated in Figure 6 show a pronounced drop in both training and validation loss during the initial epochs, with the training loss levelling off at approximately 0.2 and the validation loss at approximately 0.3 by epoch 70. The training and validation curves remain close to one another, suggesting that the learning process is solid and that dropout and L2 regularization are effectively mitigating overfitting. The fact that the losses do not begin to separate supports the conclusion that the model generalizes well to unseen data.

Concerning Figure 6, the training of the cascaded CNN model with RCGA optimization can be further assessed using the losses for both the model and the validation set over the 70 epochs. The two curves have high initial loss values, mostly around 1.5, which indicates that the model initially predicts a fair number of inaccuracies. As the training continues, the training and validation loss values decrease, which can be attributed to a reduction in error due to improvement and minimization of errors using the backpropagation method. In the early epochs(the initial 0 to 10 epochs), both sets of data rapidly decline in loss values. In fact, this rapid reduction is especially true concerning the loss values relative to the training data rather than to the validation data; in other words, the training loss is much lower than the validation loss. This means that the model can accurately capture the basic structures in the data. During the middle phase of the training process, which spans epochs 10 to 40, the decrease in loss begins to slow down as more refined adjustments are made to the model parameters and more sophisticated features are learned from the data. At epoch 40, both curves exhibit a general decrease, which, in all likelihood, indicates that the model is making progress. However, a wobble has been reached whereby progression is made less rapidly. During the training period, which spans epochs 40 to 70, both curves converge; the training loss tends towards 0.2, while the validation loss is approximately 0.3.

Table 8. *Comparison with baseline models.*

Model	Accuracy	Precision	Recall (sensitivity)	F1-score	AUC-ROC
Proposed cascaded CNN with RCGA	**99.5%**	**98.7%**	**99.0%**	**98.8%**	**0.99**
Support vector machine (SVM)	88.5%	87.2%	86.0%	86.6%	0.88
Random forest (RF)	85.8%	85.0%	84.5%	84.7%	0.87
K-nearest neighbours (KNN)	83.6%	82.5%	81.0%	81.7%	0.84
Logistic regression (LR)	80.2%	78.8%	79.5%	79.1%	0.82
Naive Bayes (NB)	78.0%	76.5%	77.2%	76.8%	0.80

Evaluating the cascaded CNN with the RCGA and five other baseline models, including SVM, RF, KNN, LR and NB, against the proposed model indicates that the proposed model is the best in all the evaluation metrics, as shown in Table 8. The deep learning model, CNN with RCGA, achieves an astounding 99.5% accuracy compared to 88.5% for the best-performing baseline model (SVM).

The outputs from the additional evaluation on Dataset 2 are presented in Table 9. Figure 7 explains the confusion matrix, illustrating the performance of the proposed model on the dataset during testing.

Table 9. *Test performance results of femoral dataset in proposed model.*

Model	Accuracy	Precision	Recall (sensitivity)	F1-score	AUC-ROC
Proposed cascaded CNN with RCGA	**97%**	**97%**	**96%**	**96.6%**	**0.97**

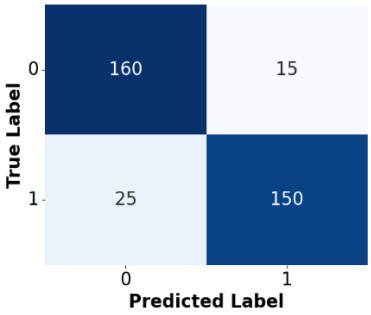

Figure 7. *Confusion matrix of Dataset 2 in testing.*

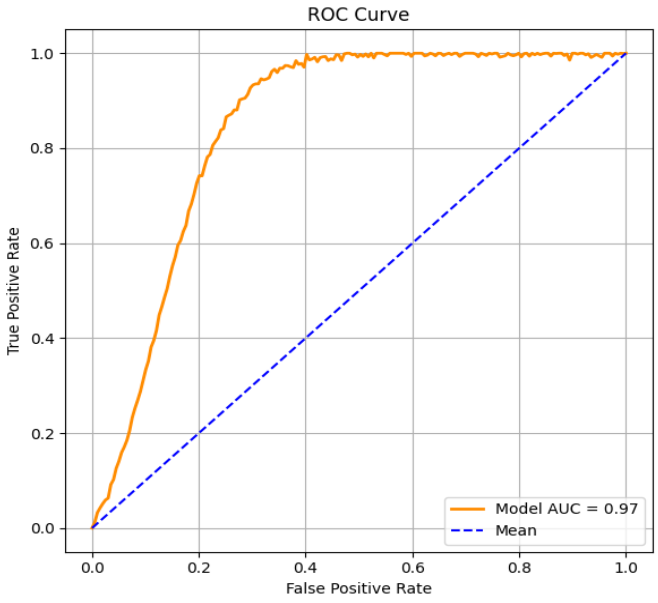

Figure 8. *RoC curves of proposed model in Dataset 2 testing performance.*

Table 10. *Result comparison with literature survey.*

Model	Input type	Accuracy
Proposed cascaded CNN with RCGA	Tabular (BMD, demographics)	**99.5%**
Boruta-XGBoost (Zhang et al., 2025)	Tabular BMD-NHANES	90.03%
LASSO-logistic regression (LR) (Zhang et al., 2025)	Tabular BMD-NHANES	74.8%
Logistic regression (Al-Husaini et al., 2025)	Femoral neck BMD	85%

The various models for classifying osteoporosis and their respective accuracies are benchmarked in Table 10. Performance results are captured for each classification approach. The proposed cascaded CNN with the RCGA achieved a maximum accuracy of 99.5%; the combination of deep learning with hyperparameter tuning effectively learns complex associations within the relational tabular BMD and demographic data. The accuracy of Boruta-XGBoost (90.03%) demonstrates the successful application of feature selection to optimize prediction precision; however, it still falls short of that achieved by the deep learning model. The accuracy of the LASSO-logistic regression model of 74.8% highlights a significant limitation of linear models and their application in the complexities of osteoporosis detection. A moderate accuracy of 85% was attained with logistic regression on femoral neck BMD data, thereby providing a more practical yet less comprehensive approach. Inevitably, it was overshadowed by ensemble and deep learning techniques. In addition to the contexts in which simpler models, such as logistic regression, possess utility and efficiency, the cascaded CNN with the RCGA remains the most accurate, demonstrating the sophistication of contemporary deep learning models for complex medical classification.

Among other baselines, the proposed model performs exceptionally well in essential measures such as precision and recall, surpassing all the baselines at 98.7% and 99.0%, respectively. This means that when positive cases are reported, very few false positives are generated. An F1-score of 98.8% is a good score, indicating that the model has performed well across all stages and can even identify actual cases while remaining free of false-positives. Additionally, the presented work demonstrated an AUC-ROC of the model design at 0.99, indicating the qualitative ability of the cascaded CNN with the RCGA in accurately differentiating between osteoporotic and non-osteoporotic patients, contrasted with the baseline models and suggesting a diminishing return as the baseline models with the best AUC-ROC for SVM at 0.88 and the lowest naive Bayes at 0.80 are introduced.

The stated accuracy of 99.5% results from a combination of architectural choices and training techniques, all designed to minimize overfitting. Datasets were divided via stratified random sampling, ensuring clean splits with no overlap between training, validation and test sets. This strict separation effectively eliminated data leakage. To curb overfitting, we employed dropout at a rate of 40%, an L2 penalty of 0.001 and early stopping after 10 epochs with no validation improvement. Loss curves for training and validation, shown in Figure 8, tracked closely, reinforcing the notion of consistent generalization.

Additionally, adaptive weight fusion directed attention to the most informative features, while a genetic algorithm for configuration of hyperparameters (RCGA) guided the search for a sweet spot between complexity and predictive power. We recognize that variability in medical imaging is often mentioned. However, employing curated structured clinical data instead of raw imaging may have bolstered the model performance. We plan to test the model on external and cross-institutional datasets to assess its generalizability and robustness further.

6 CONCLUSION AND FUTURE WORK

The proposed approach of constructing a cascaded CNN using adaptive weight fusion and the RCGA for hyperparameters has proven to be successful and robust enough for osteoporosis detection. The model achieved an accuracy rate of 99.5%, precision of 98.7% and recall of 99.0%, with an AUC-ROC of 0.99, surpassing the performance of other ML models. While the cascaded CNN configuration utilizes both low-detail and high-detail components of the medical photograph by adopting the appropriate weight fusion technique, this technique also enhances the relevance of the features used by the model derived from heterogeneous network layers. Additionally, it is worth noting that the RCGA supports hyperparameter optimization, which enables the achievement of good results on the training dataset and generalizes the model to new datasets. Given the structured nature of the input data, traditional visual explanation tools such as Grad-CAM or saliency maps are not directly applicable. However, the model provides interpretability through its adaptive weight fusion mechanism, which assigns importance to different feature hierarchies during learning.

Furthermore, we analysed feature-level contributions (e.g., BMD, age or fracture history), which can provide insights into the drivers of prediction. For improved transparency, future work will incorporate SHAP or LIME-based analysis to give the clinicians individual-level explanations for each diagnostic prediction. These additions will support better trust and adoption in real-world healthcare settings. While the proposed work shows promising results, we have a few limitations on the BMD dataset. In this research, we utilized two datasets which are not widely used, and previously there were no benchmark studies on them. In future research, it is essential to validate the

proposed model using other datasets such as NHANES, enable comparison with existing studies and improve the clinical reliability.

The findings demonstrate that the proposed model is precise and reliable, making it suitable for clinical practice in osteoporosis diagnosis. With its capacity to discern hidden complexities within medical images and its robust ability to generalize, it can enhance the correct identification of such diseases by clinical doctors in time to avert the risk of developing osteoporosis-related fractures. Based on the model performance, it can support automated assessment of medical images. Although the suggested model yields current and relevant results, some gaps remain for further development and investigation. Future research directions may include extending the capabilities of the model to perform multi-class tasks, such as identifying cancerous bones beyond osteoporosis. It is also believed that the current work can be enhanced by incorporating additional patient data, such as clinical and demographic information, in conjunction with imaging data.

ADDITIONAL INFORMATION AND DECLARATIONS

Conflict of Interests: The authors declare no conflict of interest.

Author Contributions: H.B.: Conceptualization, Methodology, Software, Data curation, Writing – Original draft preparation. M.L.P.: Visualization, Investigation. V.S.: Supervision. V.K.: Software, Validation, Writing – Reviewing and Editing.

Statement on the Use of Artificial Intelligence Tools: The authors declare that they didn't use artificial intelligence tools for text or other media generation in this article.

Data Availability: The data that support the findings of this study are available from the corresponding author. The used datasets can be obtained here: https://www.kaggle.com/datasets/amarsharma768/bmd-data and https://doi.org/10.17632/kys6x6wykj.1.

Data Availability: The data that support the findings of this study are available from the corresponding author.

REFERENCES

Al-Husaini, N., Razali, R., Al-Haidose, A., Al-Hamdani, M., & Abdallah, A. M. (2025). Characterizing low femoral neck BMD in Qatar Biobank participants using machine learning models. *BMC Musculoskeletal Disorders*, 26(1), 492. https://doi.org/10.1186/s12891-025-08726-5

Alden, Z. N. a. M. S., & Ata, O. (2024). A comprehensive analysis and performance evaluation for osteoporosis prediction models. *PeerJ Computer Science*, *10*, e2338. https://doi.org/10.7717/peerj-cs.2338

Amani, F., Amanzadeh, M., Hamedan, M., & Amani, P. (2024). Diagnostic accuracy of deep learning in prediction of osteoporosis: a systematic review and meta-analysis. *BMC Musculoskeletal Disorders*, 25(1), 991. https://doi.org/10.1186/s12891-024-08120-7

Ashwini, A., Chirchi, V., Balasubramaniam, S., & Shah, M. A. (2025). Bio inspired optimization techniques for disease detection in deep learning systems. *Scientific Reports*, 15(1), 18202. https://doi.org/10.1038/s41598-025-02846-7

Belali, M. H., Khan, F. R., Rajin, N. M., Sarkar, H., & Uddin, K. M. M. (2025). Bone Osteoporosis Fractures Detection with Deep Learning: An X-ray Image Analysis Approach. *Biomedical Materials & Devices*, 4(1), 533–551. https://doi.org/10.1007/s44174-025-00285-6

Chan, W. P., Liu, J., & Chi, W. (2004). Evaluation of bone mineral density of the lumbar spine and proximal femur in population-based routine health examinations of healthy asians. *Acta Radiologica*, 45(1), 59–64. https://doi.org/10.1080/02841850310003221

Chen, Y., Chan, W. P., Zhang, H., Tsai, Z., Peng, H., Huang, S., Jang, Y., & Kuo, Y. (2024). Automated osteoporosis classification and T-score prediction using hip radiographs via deep learning algorithm. *Therapeutic Advances in Musculoskeletal Disease*, 16, 1759720X241237872. https://doi.org/10.1177/1759720x241237872

Cooper, C., O'Neill, T., & Silman, A. (1993). The epidemiology of vertebral fractures. *Bone*, 14, 89–97. https://doi.org/10.1016/8756-3282(93)90358-h

Dhanagopal, R., Menaka, R., Kumar, R. S., Raj, P. T. V., Debrah, E. L., & Pradeep, K. (2024). Channel-Boosted and Transfer Learning Convolutional Neural Network-Based Osteoporosis Detection from CT Scan, Dual X-Ray, and X-Ray Images. *Journal of Healthcare Engineering*, 2024, 1–10. https://doi.org/10.1155/2024/3733705

D'Souza, M., Nimma, D., Pokkuluri, K. S., Ramesh, J. V. N., Kondaveeti, S. B., & Kongala, L. (2024). Multiclass Osteoporosis Detection: Enhancing Accuracy with Woodpecker-Optimized CNN-XGBoost. *International Journal of Advanced Computer Science and Applications*, 15(8), 0150889. https://doi.org/10.14569/ijacsa.2024.0150889

Carberry, G. A., Pooler, B. D., Binkley, N., Lauder, T. B., Bruce, R. J., & Pickhardt, P. J. (2013). Unreported vertebral body compression fractures at abdominal multidetector CT. *Radiology*, 268(1), 120–126. https://doi.org/10.1148/radiol.13121632

Ha, T.J., Kim, H.S., Kang, S.U., Lee, D., Kim, W.J., Moon, K.W., Choi, H.S., Kim, J.H., Kim, Y., Bak, S.H., & Park, S.W. (2024). Multi-classification of osteoporosis grading stages using abdominal computed tomography with clinical variables: Application of deep learning

with a convolutional neural network. *Journal of the Korean Society of Radiology*, 18(3), 187–201. https://doi.org/10.7742/jksr.2024.18.3.187

Halesh, T.G., & Sathish, P. (2024). Femur bone volumetric estimation for osteoporosis classification based on deep learning with tuna jellyfish optimization using X-ray images. *Multiagent and Grid Systems*, *20*(1), 1–25. https://doi.org/10.3233/mgs-230123

Ho, C., Fan, T., Kuo, C., Yen, T., Chang, S., Pei, Y., & Chen, Y. (2025). HarDNet-based deep learning model for osteoporosis screening and bone mineral density inference from hand radiographs. *Bone*, *190*, 117317. https://doi.org/10.1016/j.bone.2024.117317

Huo, X., Sun, G., Tian, S., Wang, Y., Yu, L., Long, J., Zhang, W., & Li, A. (2024). HiFuse: Hierarchical multi-scale feature fusion network for medical image classification. *Biomedical Signal Processing and Control*, 87, 105534. https://doi.org/10.1016/j.bspc.2023.105534

Houssein, E. H., Saber, E., Ali, A. A., & Wazery, Y. M. (2024). Integrating metaheuristics and artificial intelligence for healthcare: basics, challenging and future directions. *Artificial Intelligence Review*, 57(8), 205. https://doi.org/10.1007/s10462-024-10822-2

Iliou, T., Anagnostopoulos, C., Stephanakis, I. M., & Anastassopoulos, G. (2017). A novel data preprocessing method for boosting neural network performance: A case study in osteoporosis prediction. *Information Sciences*, 380, 92–100. https://doi.org/10.1016/j.ins.2015.10.026

Kale, K. D., Ainapure, B., Nagulapati, S., Sankpal, L., & Satpute, B. S. (2024). Chronological-hybrid optimization enabled deep learning for boundary segmentation and osteoporosis classification using femur bone. *The Imaging Science Journal*, 72(2), 183–203. https://doi.org/10.1080/13682199.2023.2200513

Kale, K., Naik, A., & Naik, P. (2025). Femur Dual-Energy X-ray Absorptiometry (DEXA) Scan Dataset of the Indian Population. *Mendeley Data*, V1. https://doi.org/10.17632/kys6x6wykj.1

Kang, J., Park, C., Lee, D., Yoo, J., & Kim, M. (2023). Prediction of bone mineral density in CT using deep learning with explainability. *Frontiers in Physiology*, 13, 1061911. https://doi.org/10.3389/fphys.2022.1061911

Kunjali, S., & Malarvizhi, S. (2024). Optimized Support Vector Machine using Sparrow Search Optimization for Osteoporosis Disease Diagnosis. In *2024 2nd International Conference on Sustainable Computing and Smart Systems*, (pp. 1523–1528). IEEE. https://doi.org/10.1109/icscss60660.2024.10625185

Liu, L. (2024). Implemented classification techniques for osteoporosis using deep learning from the perspective of healthcare analytics. *Technology and Health Care*, 32(3), 1947–1965. https://doi.org/10.3233/thc-231517

Lee, S., Kim, J., Kang, H., Kang, D., & Park, J. (2021). Genetic algorithm based deep learning neural network structure and hyperparameter optimization. *Applied Sciences*, 11(2), 744. https://doi.org/10.3390/app11020744

Müller, D., Bauer, J. S., Zeile, M., Rummeny, E. J., & Link, T. M. (2008). Significance of sagittal reformations in routine thoracic and abdominal multislice CT studies for detecting osteoporotic fractures and other spine abnormalities. *European Radiology*, 18(8), 1696–1702. https://doi.org/10.1007/s00330-008-0920-2

Oh, S., Kang, W. Y., Park, H., Yang, Z., Lee, J., Kim, C., Woo, O. H., & Hong, S. (2024). Evaluation of deep learning-based quantitative computed tomography for opportunistic osteoporosis screening. *Scientific Reports*, 14(1), 363. https://doi.org/10.1038/s41598-023-45824-7

Johnell, O., & Kanis, J. A. (2006). An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. *Osteoporosis International*, 17(12), 1726–1733. https://doi.org/10.1007/s00198-006-0172-4

Panse, V., & Gupta, R. (2021). Medical image enhancement with brightness preserving based on local contrast stretching and global dynamic histogram equalization. In *2021 10th IEEE International Conference on Communication Systems and Network Technologies*, (pp. 164–170). IEEE. https://doi.org/10.1109/CSNT51715.2021.9509670

Sarhan, A. M., Gobara, M., Yasser, S., Elsayed, Z., Sherif, G., Moataz, N., Yasir, Y., Moustafa, E., Ibrahim, S., & Ali, H. A. (2024). Knee Osteoporosis diagnosis based on Deep Learning. *International Journal of Computational Intelligence Systems*, 17(1), 241. https://doi.org/10.1007/s44196-024-00615-4

Sarmadi, A., Razavi, Z. S., Van Wijnen, A. J., & Soltani, M. (2024). Comparative analysis of vision transformers and convolutional neural networks in osteoporosis detection from X-ray images. *Scientific Reports*, 14(1), 18007. https://doi.org/10.1038/s41598-024-69119-7

Samala, S., Prasad, C.R., Kumar, P.R., Akash, K., Afreena, M.D., & Kumar, M.S. (2024). CNN-Based Osteoporosis Diagnosis: A Novel Approach to Bone Fragility Analysis. In *2024 International Conference on IoT Based Control Networks and Intelligent Systems*, (pp. 1111–1116). IEEE. https://doi.org/10.1109/ICICNIS64247.2024.10823107

Sampath, K., Rajagopal, S., & Chintanpalli, A. (2024). A comparative analysis of CNN-based deep learning architectures for early diagnosis of bone cancer using CT images. *Scientific Reports*, 14(1), 2144. https://doi.org/10.1038/s41598-024-52719-8

Sivasakthi, B., Preetha, K., & Selvanayagi, D. (2025). Osteoporosis Disease Detection using Optimized Elman Recurrent Neural Network based on Hybrid Bacterial Colony Optimization and Tabu Search Algorithm. *International Research Journal of Multidisciplinary Technovation*, 7(1), 1–16. https://doi.org/10.54392/irjmt2511

Sharma, A. (2021). BMD dataset. https://www.kaggle.com/datasets/amarsharma768/bmd-data

Tong, X., Wang, S., Zhang, J., Fan, Y., Liu, Y., & Wei, W. (2024). Automatic Osteoporosis Screening System Using Radiomics and Deep Learning from Low-Dose Chest CT Images. *Bioengineering*, 11(1), 50. https://doi.org/10.3390/bioengineering11010050

Woud, W., Hamad, M. M., & Alrawi, A. T. H. (2025). Enhancing Osteoporosis Detection with Fuzzy Logic Preprocing and Pre-Trained Deep Convolutional Neural Networks. *Journal of Intelligent Systems and Internet of Things*, 16(2), 214–235. https://doi.org/10.54216/jisiot.160216

Wright, N. C., Looker, A. C., Saag, K. G., Curtis, J. R., Delzell, E. S., Randall, S., & Dawson-Hughes, B. (2014). The recent prevalence of osteoporosis and low bone mass in the United States based on bone mineral density at the femoral neck or lumbar spine. *Journal of Bone and Mineral Research*, 29(11), 2520–2526. https://doi.org/10.1002/jbmr.2269

Zhang, B., Chen, Z., Yan, R., Lai, B., Wu, G., You, J., Wu, X., Duan, J., & Zhang, S. (2024). Development and validation of a Feature-Based Broad-Learning system for opportunistic osteoporosis screening using lumbar spine radiographs. *Academic Radiology*, 31(1), 84–92. https://doi.org/10.1016/j.acra.2023.07.002

Zhang, D., Yang, X., Wang, F., Qiu, C., Chai, Y., & Fang, D. (2025). Comparative analysis of feature extraction methods and machine learning models for predicting osteoporosis prevalence. *Journal of Medical Systems*, 49(1), 72. https://doi.org/10.1007/s10916-025-02203-1

Acta Informatica Pragensia is published by the Prague University of Economics and Business, Czech Republic | eISSN: 1805-4951