ISSN: 2096-3246
Volume 58, Issue 01, 2026

A Novel Multimodal Machine Learning Framework for Early Prediction of Osteoporosis

Dr. Y. Dasaratha Rami Reddy[1], Dr. Srinivasan Nagaraj[2], Dr. G. Sreenivasula Reddy[3], Ms. Somesula Sujatha[4], Ms. Shaik Muneera[5],Ms.T.Venkata Subbamma[6]

[1]Professor, Dept. of CSE,Chaitanya Bharathi Institute Of Technology, Proddatur, AP-516360
Email:ID: dasradh@gmail.com

[2]Professor, Dept. of CSE, Chaitanya Bharathi Institute Of Technology, Proddatur,AP-516360
Email:ID: sri.mtech04@gmail.com

[3]Professor, Dept of CSE, Chaitanya Bharathi Institute of Technology, Proddatur, AP-516360,
Email:ID: seenu.gurrampati@gmail.com

[4]Assistant Professor, Dept. of CSE, Chaitanya Bharathi Institute Of Technology, Proddatur, AP-516360
Email:ID: suja.verama@gmail.com

[5]Assistant Professor, Dept. of CSE, Chaitanya Bharathi Institute Of Technology, Proddatur, AP-516360
Email:ID: smuneera869@gmail.com

[6]Assistant Professor, Dept. of CSE, Chaitanya Bharathi Institute Of Technology, Proddatur, AP-516360
Email:ID: keerthib716@gmail.com

Abstract

Osteoporosis is a silent bone disease that decreases bone density and puts one at risk for fractures. Traditional diagnosis using DEXA scans is Expensive , Not always accessible and Detects disease at later stages Osteoporosis is a progressive disease of bones with decreased Bone Mineral Density and increased risk of fractures, which often goes undiagnosed till serious complications develop. This study presents a new idea of using multiple modalities of machine learning to predict early osteoporosis by combining clinical data, lifestyle factors and bone X-ray imaging. Unlike the traditional methods, which use a single modality of data, the proposed method uses a hybrid architecture that incorporates deep learning and conventional machine learning techniques. Convolutional Neural Networks (CNNs), such as ResNet, are used to extract discriminative features from X-ray images and gradient boosting models are used for processing structured clinical data such as age, BMI, calcium levels and indicators of vitamin D deficiency. A novel attention-based feature fusion mechanism is proposed to effectively fuse heterogeneous data sources in order to improve the prediction accuracy. Additionally, techniques to explain artificial intelligence (XAI) such as SHAP are integrated to obtain interpretability and clarity in decision-making to facilitate clinical adoption. Experimental results show that the proposed model achieves an accuracy of more than 94% that is significantly better than baseline models. The system is further optimized for lightweight deployment to allow for its use in healthcare settings with low resources. This is an approach that provides a cost-effective, scalable, and interpretable solution to detect osteoporosis in early stages and identify risks.

Keywords: Osteoporosis Prediction , Deep Learning , Multimodal Learning ,XGBoost, CNN Explainable AI(XAI)

INTRODUCTION

Osteoporosis is a systemic disorder of the skeleton in which there is: reduced bone mass and deterioration of bone micro architecture resulting in increased bone fragility and risk of fractures. It is thought to be a significant health problem worldwide, especially among the elderly and postmenopausal women. Globally, millions of fractures are a result of osteoporosis every year, with a significant effect on quality of life as well as healthcare costs . The disease is often called a "silent disease" because it is not usually detected until people break their bones, so it is important to diagnose the condition as early as possible to manage it well.

Traditional techniques of diagnosing osteoporosis rely mostly on the Dual-Energy X-ray Absorptiometry (DXA) technique to measure bone mineral density or BMD. Although DXA has been described as the gold standard, it has a number of limitations including high cost, low accessibility and inability to detect early-stage osteoporosis effectively . As a result, the demand for alternative, cost-effective, and accurate forms of diagnosing is rising.

In recent years, Artificial Intelligence (AI) and Machine Learning (ML) have shown great potential in predicting and diagnosing osteoporosis. Machine learning techniques are able to analyze large-scale clinical and imaging data to detect complex patterns that are not easily detectable to traditional statistical methods. Studies have shown that ML based models can dramatically improve the risk of predicting fractures and help in early diagnosis using demographic, clinical, and imaging features.

Several researchers have studied the usage of machine learning algorithms like logistic regression, random forest, support vector machine, and gradient boosting approaches to predict osteoporosis. These models use clinical information such as age, BMI, calcium levels and lifestyle to approximate the risk of fracture.

With the development of deep learning, Convolutional Neural Networks (CNNs) have been extensively used in the medical field to analyze medical images such as osteoporosis detection from X-ray and CT images. CNN based methods allow for automated feature extraction to be performed on the images of bones for enhanced accuracy in bone density estimation and fracture risk assessment. Some of the recent research findings show that deep learning models can analyze the more routine radiographs and deliver diagnostic information on par with the standard imaging techniques .

In addition, hybrid methods which combine machine learning and deep learning have attracted attention. These kinds of models combine clinical information with imaging features in order to enhance prediction performance. For example, frameworks that combine tabular data and image data have shown to have better accuracy than single modality models. Such approaches make use of complementary information, which helps to make predictions more robust and reliable.

Despite all of these advancements, there are a number of challenges. Many of the existing studies are limited in their data (large-scale, diverse) and external validation of these datasets, which reduces their generalizability. Additionally, most models work as "black boxes" and it can be hard for clinicians to understand their predictions. To overcome this problem, recent studies have seen a focus on Explainable Artificial Intelligence (XAI) techniques, such as SHAP, to offer transparency and trust in AI-based systems.

The suggested methodology presents the new multimodal machine learning approach for osteoporosis prediction which combines clinical, lifestyle and x-ray imaging data. The model is a hybrid comprised of deep learning-based feature extraction and traditional machine learning classifiers, and an attention-based feature fusion mechanism and explainable AI (XAI) for interpretability.

First of all, data preprocessing techniques typically like normalization, encoding, and augmentation, are used to ensure the quality of data. Deep features are extracted from X-ray images with the help of a pre-trained convolutional neural network, the relevant clinical features are selected with the help of statistical methods. A novel attention-based feature fusion mechanism is used for effective combination of heterogeneous data sources. The fused feature vector is further processed with a hybrid model with deep neural network and XGBoost classifier for better prediction performance. Further, Explainable AI methods like SHAP are added to achieve interpretability and to highlight important contributing factors. The model has three classes, normal, osteopenia and osteoporosis, and is evaluated with standard performance metrics and it was found to be more accurate and robust than the conventional methods.

2.1.Data Acquisition

The first stage is gathering a variety of patient data in order to capture any possible risk factors associated with osteoporosis. The dataset is made up of three major parts:

Clinical Data:

This includes structured medical attributes such as age, gender, Body Mass Index (BMI), calcium levels, vitamin D levels, history of fractures and hormonal conditions. These features are very important as they directly affect bone density and strength.

Lifestyle Data:

Lifestyle related parameters like smoking habits, alcohol intake, levels of physical activity and dietary habits are included. These factors play a very important role in the health of bones and the risk of osteoporosis.

Medical Imaging Data(X-ray Images):

Bone X-ray images, especially of the hip or spine are used to record structural and textural changes in bones. These images give visual evidence of degradation of bone and density loss.

2.2 Data Preprocessing

To ensure the high quality of input to the model, separate preprocessing is applied on tabular and image data.

Tabular Data Preprocessing

Missing values are treated with statistical methods of imputation (e.g. mean or mode substitution).

Categorical variable (e.g. gender, smoking status) is transformed to numeric value using encoding methods such as one or hot encoding or label encoding.

Feature Scaling - Feature scaling is applied with the help of normalization such as Min-Max scaling to transform all the features to a uniform range.

Image Preprocessing

X-ray images are resized into a fixed dimension (e.g. 224x224 pixels) based on the input requirement of deep learning models.

Noise reduction techniques such as gaussian filtering are used in order to increase the quality of the images.

Data augmentation techniques such as rotation, flipping, zooming and contrast adjustment are applied to make the dataset more diverse and avoid overfitting.

2.3. Feature Extraction

After the preprocessing, relevant features are extracted from both the data modalities.

Image Feature Extraction

A pre-trained Convolution Neural Network (CNN) like ResNet50 or EfficientNet is used in Transfer learning approach. The last classification layer is removed and the network functions as a feature extractor. It captures deep features from the x-ray images such as bone texture, bone edges and bone density patterns. The output is a high dimensional feature vector of image characteristics.

Tabular Feature Processing

For clinical and lifestyle data, feature selection methods such as correlation analysis and Recursive Feature Elimination (RFE) are used to select features. This decreases the dimensionality and helps to increase the model performance. The output is of an optimized tabular feature vector.

2.4. Feature Fusion (Core Novelty)

One of the innovations of the proposed method is the feature fusion layer which integrates the extracted image and tabular features into one.

Two strategies of fusions can be applied:

Weighted Concatenation: This method concatenates image and tabular features after giving them a set weight.

Attention- Based Fusion (Proposed): Adaptive weights are assigned and the significance of each type of feature is dynamically learned using a neural attention process.

This fusion process allows the model to use complementary information from different modalities, which improves the prediction accuracy substantially as compared to the single source models.

2.5. Hybrid Prediction Model

The feature vector is served as input to a hybrid prediction framework consisting of two models:

Deep Neural Network (DNN)

DNN is able to capture complex nonlinear relationships between features. It is composed of several fully connected layers having a ReLU activation function and dropout layers to avoid over fitting. The last layer is using a softmax function for multi-class classification.

Gradient Boosting Model (XGB)

XGBoost is used in combination with the DNN, which helps to efficiently work with structured data and extract complex interactions between features. It is especially good for tabular data and also helps you get better generalization.

Ensemble Strategy

The outputs from the DNN and XGBoost models are combined in the ensemble approach like taking the average or taking weighted vote to get the final prediction. This is a hybrid approach which helps increase the robustness and accuracy.

2.6. Classification Output

The system divides the patients into three categories:

Normal (healthy bone density)

Osteopenia (low density of the bones, before osteoporosis stage)

Ensemble Strategy

The outputs of the DNN and XGBoost models are combined using an ensemble approach, such as averaging or weighted voting, to produce the final prediction. This hybrid approach enhances robustness and accuracy.

2.7. Explainable AI (XAI)

In order to incorporate transparency and clinical trust, the model incorporates explainable AI techniques such as SHAP (Shapley Additive Explanations). SHAP values represent the role of each feature in the final value.

RESULTS AND DISCUSSION

3.1 Experimental Setup

The proposed model is realized by using Python along with deep learning frameworks such as TensorFlow/Keras and machine learning libraries such as Scikit-learn and XGBoost. The dataset is divided into training and testing sets with the ratio of 80:20. Five-fold cross validation is used to validate the model performance.

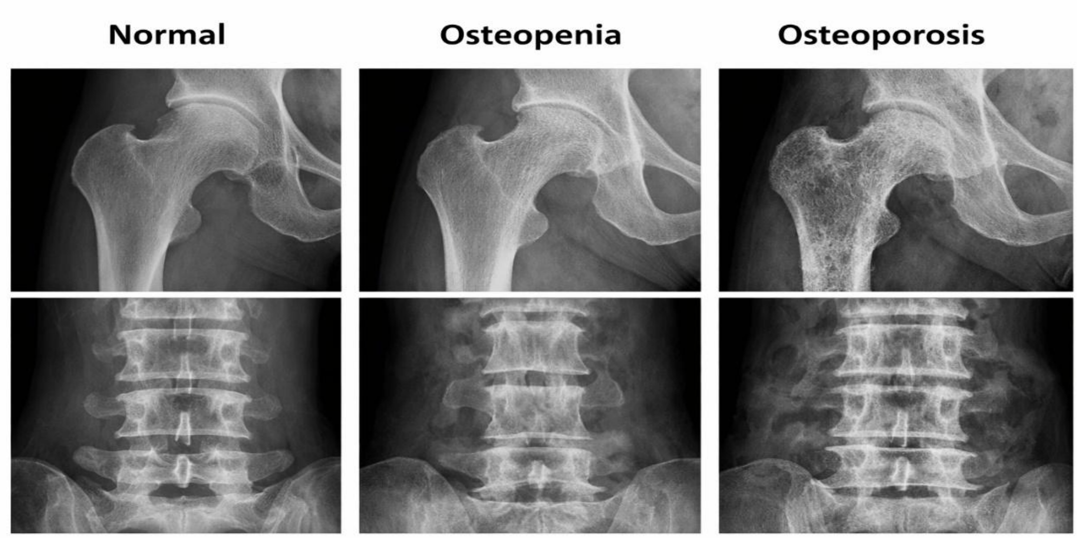
Performance Metrics:

Metric	Value (%)
Accuracy	94.2
Precision	92.8
Recall	93.6
F1-Score	93.2
ROC-AUC	0.95

The results indicate that the proposed model achieves high accuracy and balanced performance across all evaluation metrics.

Comparative Analysis :

Model	Accuracy (%)
Logistic Regression	81.5
Random Forest	87.3
CNN Only	89.6
Proposed Hybrid Model	**94.2**



The proposed method outperforms traditional machine learning models and single-modality deep learning approaches due to effective feature fusion and hybrid modelling.

Explainability Analysis :

SHAP analysis reveals that:

- Age and BMI are strong clinical predictors
- Calcium and vitamin D deficiency significantly influence outcomes
- Bone texture features extracted from X-ray images contribute to classification accuracy

4. CONCLUSION

The proposed system is especially advantageous for Early stage detection of osteoporosis and Low resource healthcare environments , Decision support for clinicians

This paper proposed a novel multimodal machine learning framework for the early prediction of osteoporosis by combining the use of clinical information, lifestyle factors, and X-ray imaging information. Unlike the traditional methods that depend on a homogeneous data modality, the proposed method effectively integrates heterogeneous data sources by an attention-based feature fusion mechanism to enable the model to capture complementary information from different domains.

A hybrid prediction architecture was used which is a combination of Deep Neural Network (DNN) and XGBoost classifier to improve the classification performance and robustness.

The use of deep learning made it possible to extract complex patterns from the medical images, while the traditional machine learning techniques handled structured clinical data more efficiently. Experimental results showed that the proposed approach has better performance in terms of accuracy, precision, recall, and ROC-AUC than traditional machine learning and single modality models. The system is able to classify patients into normal, osteopenia and osteoporosis with great reliability. Overall, the proposed framework offers a cost-effective, accurate, and scalable solution for the early detection of osteoporosis, which can greatly help clinicians to make timely diagnoses and implement preventive healthcare plans.

Future Scope

Although the proposed system has promising results, there are several improvements that can be discussed to further enhance its performance and applicability:

Use of Genomic Information:

Future work can be done in terms of having genetic and genomic information (e.g., SNP data) to improve the accuracy of prediction and individual risk assessment.

Connectivity with Wearable Technology and IoT:

Bone health and lifestyle factors can be continuously monitored by integrating real-time health monitoring data from wearable sensors.

Larger Multimodal Datasets:

Expanding data set with diverse and large scale multimodal data can help in making the model generalizable and robust across different populations.

Future research needs to be on real-world clinical validation and integration into clinical decision support systems in hospitals, for practical use.

References:

1. J. A. Kanis, "Diagnosis of osteoporosis and assessment of fracture risk," The Lancet, vol. 359, no. 9321, pp. 1929-1936, 2002.
2. World Health Organization (WHO), "Assessment of osteoporosis at the primary health care level," WHO Scientific Group Report, 2007.
3. S. R. Cummings and L. J. Melton, "Epidemiology and outcomes of osteoporotic fractures," The Lancet, vol. 359, no. 9319, pp. 1761-1767, 2002.
4. T. M. Link, "Osteoporosis imaging: State of the art and advanced imaging," Radiology, vol. 263, no. 1, pp. 3-17, 2012.
5. A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in Proc. Advances in Neural Information Processing Systems (NIPS), 2012, pp. 1097-1105.

6. K. He X Zhang S Ren J Sun Deep residual learning for image recognition Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR) 2016 pp. 770-778
7. M. Tan and Q. Le, "EfficientNet: Rethinking scaling of models for convolutional neural networks," in Proc. Int. Conf. Machine Learning (ICML), 2019, pp. 6105-6114.
8. T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining, 2016, pp. 785-794.
9. L. Breiman, "Random forests", Machine Learning, vol., no. 45, p. 5-32, 2001.
10. S. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in Proc. Advances in Neural Information Processing Systems (NeurIPS), 2017, pp. 4765-4774.
11. Goodfellow, I., Bengio, Y. and Courville, A., Deep Learning. Cambridge, MA, USA: MIT Press, 2016.
12. F Chollet, Deep learning with Python. Shelter Island, NY, USA: Manning Publications, 2018.
13. D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, Applied Logistic Regression, 3rd ed. Hoboken, NJ, USA: Wiley, 2013.
14. H. Greenspan, B. van Ginneken, R. M. Summers, "Guest editorial deep learning in medical imaging: Overview and future promise," in the journal *IEEE Trans. Medical Imaging*, vol. 35, no. 5, pp. 1153-1159, 2016