# Shearlet-based texture analysis and deep learning for osteoporosis classification in lumbar vertebrae

# Poorvitha Hullukere Ramakrishna<sup>1</sup>, Chandrakala Beturpalya Muddaraju<sup>1</sup>, Bhanushree Kothathi Jayaramu<sup>2</sup>, Shobha Narasimhamurthy<sup>3</sup>

# **Article Info**

# Article history:

Received Sep 2, 2024 Revised Apr 9, 2025 Accepted May 24, 2025

# Keywords:

Classification Lumbar vertebrae Neural network Osteoporosis Shearlet transform

#### **ABSTRACT**

Osteoporosis is a bone disorder characterized by reduced bone density and increased fracture risk. It challenges society's health, remarkably among the elderly population. This research proposed an innovative method by combining Shearlet-transform (ST) spectral analysis with a deep learning neural network (DLNN) and a convolutional neural network (CNN), for osteoporosis classification in lumbar vertebrae (LV) L1-L4 of spine X-ray images. The ST enables precise extraction of texture features from images by capturing significant information regarding trabecular bone microarchitecture and bone mineral density (BMD) variations revealing in osteoporosis regions. These extracted features serve as input to a DLNN for automated classification of osteoporotic and non-osteoporotic vertebrae. Similarly, without extracting any features from ST image is directly used as an input to the CNN to classify the images. The experimental results highlight the framework's effectiveness, achieving 96% accuracy in osteoporosis image classification using CNN. Early and precise detection of osteoporosis, particularly in the lumbar vertebrae, is vital for effective treatment and fracture prevention. This study particularly emphasizes the potential and effectiveness of integrating image spectral analysis technique with NN, to improving diagnostic accuracy and clinical decision-making in osteoporosis management.

This is an open access article under the **CC BY-SA** license.



# Corresponding Author:

Chandrakala Beturpalya Muddaraju

Department of Information Science and Engineering, Dayananda Sagar College of Engineering Dayananda Sagar College of Engineering, Bangalore-560078, India

Email: chandrakalabm@yahoo.com

# 1. INTRODUCTION

Osteoporosis is characterized by reduced bone density and increased vulnerability to fractures affecting millions worldwide, particularly the elderly population. Especially in critical areas like the lumbar vertebrae are vital, it needs effective management of bone disorder condition because of its importance in the body. While traditional diagnostic methods like dual-energy X-ray absorptiometry (DEXA) scan provides valuable information of bone mineral density (BMD), though often lack of capturing detailed structural variations that affect fracture risk. Over the past few years, advancements in medical imaging and machine learning (ML) techniques are promising in improving osteoporosis diagnosis by considering bone microarchitecture in the image. This paper highlights the research gap and presents novel contributions such as:

- a. Osteoporosis, a disease-causing reduced bone density and structural deterioration, poses a significant risk to lumbar vertebrae.
- b. Current diagnostic methods are costly, invasive and reliant on subjective interpretation. Traditional texture analysis not able to capture multi-scale and directional features essential for accurate classification.
- c. The ST, with its capability to analyses anisotropic structures, combined with neural networks, offers a promising solution.
- d. This research paper aims to develop a Shearlet-based texture analysis with NN framework for reliable and accurate osteoporosis classification in lumbar vertebrae.

Dong et al. [1] aims to classify osteoporosis using deep learning on lumbar spine X-rays, achieving high accuracy through a convolutional neural network (CNN). It compares the model's performance with traditional methods. The key gap includes the need for more diverse datasets to improve generalization. Interpretability of the model decisions remains a challenge for clinical adoption. Additional investigation is required on practical integration and comparison with newer techniques. Zhang et al. [2] developed a deep learning neural network (DLNN) model to classify osteopenia and osteoporosis using lumbar spine X-ray images. The model demonstrated promising performance, achieving an area under the curve (AUC) of 0.767 for diagnosing osteoporosis and 0.787 for osteopenia in the test dataset. Diagnostic networks based on deep learning could be effective for screening osteoporosis and osteopenia based on lumbar spine radiographs. However, further research is necessary to verify and enhance diagnostic accuracy of DLNN models. Xue et al. [3] presents a multi-scale weighted fusion contextual transformer network (FCoTNet) for osteoporosis prediction in lumbar spine X-ray images. The model achieved high accuracy, sensitivity, and specificity, outperforming clinician performance in a 5-fold cross-validation. FCoTNet demonstrates potential for improving automated osteoporosis detection. However, the dataset size and number of clinicians involved were not specified, limiting the generalizability assessment. The study suggests that transformer-based models can enhance diagnostic tools in healthcare. Fan et al. [4] explores osteoporosis pre-screening with panoramic radiographs using a DLNN with an attention mechanism. The model aims to improve the identification of osteoporosis by focusing on relevant features in panoramic X-ray images. The attention mechanism improves model interpretability and accuracy. The method demonstrated promising results in automated screening. The study highlights the potential of integrating attention mechanisms to enhance DLNN for osteoporosis detection. Wang et al. [5] develops a method for estimating lumbar BMD from X-ray images through an anatomy-aware attentive multi-region of interest (ROI) model. The approach incorporates attention mechanisms to highlight pertinent region of interest (ROI) in X-ray images, improving BMD estimation accuracy. The method outperforms conventional approaches in terms of accuracy and robustness. However, the study primarily focuses on BMD estimation, and its performance on osteoporosis detection specifically remains unclear. Additional studies are required to evaluate the model clinical applicability in broader osteoporosis diagnosis.

Nguven et al. [6] presents a method for estimating lumbar BMD extracted from X-ray images through an anatomy-aware attentive ROI model. The model employs attention mechanisms to concentrate on the relevant aspects of ROI in the images, enhancing the accuracy of BMD estimation. The approach stands out conventional methods in accuracy and robustness. However, the model is focused on BMD estimation rather than direct osteoporosis diagnosis, which limits its applicability to broader clinical settings. Further work is required to evaluate its potential in osteoporosis detection. Soegijoko et al. [7] focuses on trabecular patterns in proximal femur radiographs for osteoporosis detection. The approach showed potential for automatic screening and diagnosis. However, the model does not specify the performance metrics, such as accuracy or sensitivity, which are vital for assessing model effectiveness. Ongoing research is essential to validate the method robustness and clinical applicability. Jeong et al. [8] investigates bone patterns in osteoporosis through texture parameters to characterize bone architecture. The method quantifies bone texture features in radiographic images to assist in the diagnosis of osteoporosis. The model displays the potential for identifying osteoporotic changes in bone structure. However, the model does not report specific performance metrics, which are essential for evaluating the model clinical utility. Additional validation and comparison with other diagnostic techniques are necessary to determine its effectiveness. Hong et al. [9] presents a deep-learning-based approach for detecting vertebral fractures and osteoporosis using lateral spine X-ray radiographs. The model demonstrated high accuracy in detecting both conditions, outperforming traditional methods. However, the model generalization to diverse populations and clinical settings was not fully explored and broader validation and comparison with other diagnostic techniques are essential. Mu et al. [10] explores the application of deep learning-based medical imaging in the treatment of lumbar degenerative diseases. The model holds promises for accurate diagnosis and treatment planning; however, it lacks detailed metrics for evaluating its performance in clinical settings. The method also does not address the dataset size, which could impact the model generalization. Additional research and validation on larger datasets are necessary to evaluate its practical clinical applicability.

Kim et al. [11] proposes a prediction model for spinal osteoporosis using lumbar spine X-ray images, employing transfer learning with deep CNNs. The model leverages pre-trained models to enhance accuracy in osteoporosis detection. It demonstrated promising results, but dataset size and diversity were not fully disclosed, raising concerns about its applicability to different populations. Further validation across larger, more diverse datasets is recommended to assess the model robustness. Kong et al. [12] develops a deep learning-based model for predicting spinal fractures using spine X-ray images. The model utilizes CNNs to identify fracture risk, showing high predictive accuracy, however, making it challenging to fully evaluate its clinical utility. Additionally, the dataset used was not large or diverse enough to generalize the results across different populations and more research with larger datasets and clinical validation is necessary. Kim et al. [13] compares deep learning models using anteroposterior and lateral X-rays for detecting osteoporotic vertebral compression fractures. Both models demonstrated promising performance by utilizing CNNs for fracture detection. However, the model lacks comparative advantage over traditional methods was not fully explored. Additional validation with varied datasets is required to verify its generalization and clinical applicability. Mebarkia et al. [14] reviews the progression of X-ray image analysis techniques for osteoporosis diagnosis, from shallow to deep learning methods. It explores various models, including traditional image processing and modern deep learning techniques for accurate detection. It highlights the advantages of deep learning in improving diagnostic accuracy. However, it does not provide specific performance metrics or a direct comparison of methods. More research is necessary to assess these techniques in actual clinical environments.

To address the gap in existing research, this paper presents Shearlet-transform (ST) basis spectral analysis technique, which has gained attention for its ability to extract and analyze texture features in X-ray spine images (LV from L1-L4) with high precision. This multiscale, multi-directional transform is designed to capture directional features and edges in images, preserving important directional information. Unlike traditional methods such as Fourier or wavelet transforms, the ST technique is particularly effective for analyzing complex textures in lumbar vertebrae. The study applies ST spectral analysis for osteoporosis classification by extracting features that reflect bone microarchitecture and density variations. With DLNN and CNN models, this approach demonstrates improved accuracy and robustness in osteoporosis detection.

# 2. PROPOSED MODEL

This paper introduces a technique for classifying lumbar spine (L1-L4) X-ray images to distinguish between healthy and osteoporotic conditions, as illustrated in Figure 1. The process involves three steps.

- a. Preprocessing: Enhancements like sharpening make fine textures in the X-ray image more visible.
- b. Transformation: Application of ST to the enhanced input images.
- c. Classification: Transformed images are classified directly with CNN. In addition, a DLNN is applied to features derived from the ST images.

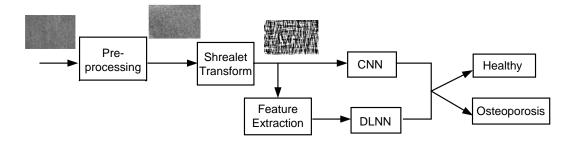


Figure 1. Outlined model

# 2.1. Pre-processing

This step enhances X-ray images (L1-L4) for better visibility and texture quality often involves advanced preprocessing techniques, like contrast limited adaptive histogram equalization (CLAHE) [15] combined with edge-preserving filter [16] for better enhancement. CLAHE improves contrast adaptively by limiting the amplification of noise. To further refine the results, edge-preserving filters like the bilateral filter are used to smooth textures while preserving edges.

#### 2.2. Shearlet-transform

The spectral analysis allows to analyze frequency information of the Shearlet-transformed coefficients across different scales and orientations [17], [18]. This provides disparities of intensity distribution of image features and textures in the spectral domain, which is vital to image aspects, including feature extraction, denoising and classification. The ST spectral analysis provides a robust framework for image analysis, particularly in tasks where directional and scale-dependent features are crucial. By decomposing images into Shearlet coefficients and analyzing their spectral properties to obtain underlying texture structure of the images. It covers the potential of traditional transforms techniques like Fourier or wavelet transforms by offering higher directional selectivity and sparsity properties. The ST decomposes the image into a group of localized and oriented basis functions, known as Shearlets, which are sensitive to edges and distinctive pixel intensity variations in the image. Shearlets are defined by scale, orientation and location, enabling them to accurately represent image features across various scales and directions.

#### 2.2.1. Shearlet transforms spectral analysis for an image

Osteoporosis classification in lumbar vertebrae (L1-L4), the ST spectral analysis involves applying the ST to radiographic images [19], [20] of lumbar vertebrae to extract texture features indicative of bone microarchitecture and density variations associated with osteoporosis. The spectral analysis aspect involves analyzing the frequency information of the Shearlet coefficients obtained from transformation. This spectral information provides insights into the fine texture information of image features across different scales and orientations, allowing for a comprehensive characterization of the textural characteristics of the vertebrae. Mathematically, in spectral analysis, the Fourier transform of the Shearlet coefficients is computed at each scale and orientation. The ST is a potent of multiscale and multidirectional transform that is particularly effective for representing geometry of multidimensional information. Hence well suited for texture analysis in enhanced images (L1-L4), where the texture patterns can indicate various conditions or abnormalities. The ST expressed as in (1).

$$SH_{W}(f)(a,s,t) = \langle f, \psi_{a,s,t} \rangle$$
 (1)

whereas (f) represents input image,  $(\psi_{a,s,t})$  represents Shearlet function with scale(a), shear(s) and translation (t). Hence Shearlet function  $(\psi_{a,s,t})$  is expressed as in (2).

$$\psi_{a,s,t}(x) = a^{-3/2} \psi(S_s A_a^{-1}(x-t))$$
 (2)

whereas  $(S_a)$  and  $(A_a)$  are the shear and scaling matrices, respectively.

For image texture extraction, the ST is performed to decompose the image into different scales and orientations, capturing the intrinsic texture features at multiple resolutions. Apply the ST to decompose the enhanced image into Shearlet coefficients. Analyze the obtained features to identify and classify osteoporosis. The equations of ST as in (3),

$$SH_{\psi}(f)(a,s,t) = \langle f, \psi_{a,s,t} \rangle$$
 (3)

The Shearlet function is expressed as in (4),

$$\psi_{a,s,t}(x) = a^{-3/2} \psi(S_s A_a^{-1}(x-t)) \tag{4}$$

The shear matrix  $S_s$  and its scaling matrix  $A_a$  as in (5).

$$S_{s} = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} \quad A_{a} = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix} \tag{5}$$

Together, the shear and scaling matrices provide a framework for decomposing an image into a set of directional components at multiple scales. This combination gives the ST its unique ability to efficiently handle geometrical features such as edges, curves and textures, making it particularly powerful for image classification. The ST provides a robust method for image texture features in enhanced images (L1-L4), it enables detailed texture information. The features extracted using ST are highly versatile and applicable to a broad spectrum in image analysis and boost the potency of classification algorithms with increase diagnostic accuracy.

# 2.3. Osteoporosis classification techniques

DLNN [21], [22] and CNN [23], [24] are revealed remarkable capabilities in image classification tasks because of their capacity to learn hierarchical representations of data on images. The DLNN model is trained on ST image features dataset with the corresponding label and every image is related with a binary label indicating the existence or non-existence of osteoporosis. Similarly, the CNN model is trained directly on Shearlet transformed images. The network learns to map the Shearlet transformed spectral features to the corresponding class labels through an iterative optimization process. By leveraging the directional and spectral information derived from the ST with the learning capabilities of networks can achieve accurate and robust classification of osteoporotic and non-osteoporotic vertebrae, which helps in the prompt identification of osteoporosis.

# 2.3.1. Shearlet with DLNN

The DLNN architecture presented here is specifically designed for binary classification tasks, such as distinguishing between osteoporosis and healthy cases. The architecture is carefully structured to extract meaningful patterns from feature vectors derived from Shearlet coefficients while mitigating overfitting and optimizing performance. The DLNN is organized sequentially input and output sizes for batch sizes of 10 with input shape (60) as follows.

- a. Dense: Input Shape (10, 60), Output Shape (10, 128)
- b. Dropout Layer: Input Shape (10, 128), Output Shape (10, 128). Dropout does not change shape, only randomly drops some connections during training
- c. Second Dense Layer: Input Shape: (10, 128), Output Shape: (10, 64)
- d. Second Dropout Layer: Input Shape (10, 64), Output Shape (10, 64). Dropout does not change shape
- e. Third Dense Layer: Input Shape (10, 64), Output Shape (10, 32)
- f. Output Layer: Input Shape (10, 32), Output Shape (10, 1)
  - (i) *Input Layer:* This layer serves as the entry point for the feature vector derived from Shearlet coefficients. The length of the feature vector as 60 and each feature acts as a neuron in this layer.
  - (ii) *Hidden Layers:* The hidden layers are composed of fully connected (dense) layers and dropout layers for enhanced feature extraction and regularization.
  - *Dense Layers:* Each dense layer has fully connected neurons that process and transform the input features using learned weights and biases. The first dense layer has 128 neurons, activation function (AF) is ReLU (Rectified Linear Unit), which introduces non-linearity by allowing only positive values to pass through. Second dense layer has 64 neurons, AF is ReLU. Third dense layer has 32 neurons, AF is ReLU.
  - Dropout Layers: Dropout layers are introduced after each dense layer to prevent overfitting by randomly setting a fraction of input units to zero during training. The first dropout layer has a dropout rate of 0.5 (50% of the neurons are randomly deactivated during training). The second dropout layer has a dropout rate of 0.5.
  - (iii) *Output Layer:* The output layer purpose is to provide the final prediction for binary classification. (single neuron, representing the binary outcome). AF as Sigmoid, which maps the output to a probability value between 0 and 1. A threshold (e.g., 0.5) is applied to decide between the two classes: osteoporosis or healthy.

The advantages of architecture are the features of learning, the dense layers with ReLU activation allow the network to learn complex, non-linear representations of the input data. The regularization, dropout layers reduce overfitting by introducing randomness in the training process, ensuring the network generalizes well to unseen data. The sigmoid activation in the output layer is well-suited for binary classification, providing probabilistic outputs. The scalability can be adjusted easily to handle feature vectors of different lengths by modifying the input layer size. This architecture provides a robust framework for the classification task, combining effective feature representation with regularization to achieve reliable predictions.

# 2.3.2. Shearlet with CNN

The CNN architecture is organized sequentially as follows, input size of (62, 62, 1) and a batch size of 32.

- a. Input Tensor: (32, 62, 62, 1)
- b. Conv1 (32 filters, 3x3): Output shape (32, 62, 62, 32)
- c. MaxPool1 (2x2): Output shape (32, 31, 31, 32)
- d. Conv2 (64 filters, 3x3): Output shape (32, 29, 29, 64)
- e. MaxPool2 (2x2): Output shape (32, 14, 14, 64)
- f. Conv3 (128 filters, 3x3): Output shape (32, 12, 12, 128)
- g. MaxPool3 (2x2): Output shape (32, 6, 6, 128)

- h. Flatten Layer: Output shape (32, 6\*6\*128 = 4608)
- i. Dense Layer (128 units): Output shape (32, 128)
- j. Dropout: No change in shape (32, 128)
- k. Dense Layer (64 units): Output shape (32, 64)
- 1. Dropout: No change in shape (32, 64)
- m. Output Layer (1 unit): Output shape (32, 1)

Input layer takes a 2D matrix as input, which represents the Shearlet transformed, the input shape is defined based on the image's dimensions, such as (height=64, width=64). Convolutional layers, perform convolution operations to extract spatial features from the input image. The layers used 32 filters and each size of the filter kernel 3×3. Activation function is rectified linear unit (ReLU) to introduce non-linearity. Pooling layers are used to reduce the spatial dimensions of the feature maps to prevent overfitting and reduce computational complexity. Pooling type is max pooling (retains the maximum value in each pooling region). The size of the pooling window (e.g., 2×2). Flatten layer, converts the 2D output of the convolutional layers into a 1D feature vector. Fully connected layers, dense layers like those in the DLNN structure. Includes: Dense Layer 1, 128 units with ReLU activation, Dense Layer 2, 64 units with ReLU activation. Regularization with dropout layers (rate = 0.5). Output layer, a single neuron with a sigmoid activation function for binary classification (e.g., 0 = healthy, 1 = osteoporosis).

# 2.4. Implementation workflow

The proposed workflow is divided into five parts based on X-ray image texture analysis with deep learning techniques to classify osteoporosis or healthy image in lumbar spine L1-L4.

- a. Preprocessing: Sharpening enhancements help reveal finer textures in the X-ray image more clearly.
- b. Apply the shearlet transform to the enhanced image: Decompose the image into two scales
- c. Extract ST coefficients: Extract horizontal  $(0^{\circ})$ , vertical  $(90^{\circ})$  and diagonal  $(45^{\circ})$  coefficients at each scale and obtain the corresponding coefficient matrix.
- d. Calculate statistical Features: For each coefficient matrix, calculate [25], [26] mean coefficient (MC) is the mean of Shearlet coefficients at the given scale and orientation. Std deviation coefficient (SDC) of Shearlet coefficients. Energy (E) is the sum of squared coefficients, providing a measure of the image activity at the given scale and orientation. Entropy (Et) is a measure of randomness or complexity in the coefficient distribution. Skewness (Sk) is a measure of the asymmetry distribution of coefficients. Kurtosis (K) is a measure of the tailedness distribution of coefficients. Contrast (C) is a measure of the local variations in the coefficient values. Correlation (Cor) is a measure of how correlated the coefficient values are in different regions of the image. Homogeneity(H) is a measure of the resemblance of coefficient values in the image.
- e. Fusion: Used ST features to analyze the structural integrity of bone patterns for osteoporosis classification by concatenation of scale-1 and scale-2 features by making it as feature vector.

The above-revealed features provide rich statistical information about the image's texture, structure, and complexity. Together, they capture both global and local image characteristics that are essential for effective classification. They can be particularly useful in distinguishing different types of textures, identifying patterns in specific regions within images. These features are useful in DLNN model, to classify images accurately based on their intrinsic properties.

#### 3. DATA SET DESCRIPTION

The research study collected X-ray spine images with a focal distance of 0.812 meters and X-ray parameters set at 75-80 kV and 80 mAs for all patients. The images were provided by a prestigious hospital in Bangalore, Karnataka, India, and are 2D radiographic JPEGs. To test the system experimentally, datasets comprising 52 X-ray images from control subjects (CS/normal (healthy)) and 56 from pathological cases (osteoporotic or abnormal subjects (OS)), accompanied by DXA reports for the same individuals. These DXA reports detail the statistical status of the lumbar spine (L1-L4) as either normal or osteoporotic. Table 1 lists the database of region of interest (ROI) X-ray images of the lumbar spine (L1-L4) employed for experimental validation along DXA reports.

Table 1. Data description

Images	Numbers of X-ray images	ROI sub images (L1-L4)	Training	Testing
CS	52	$208 (52 \times 4)$	168	40
OS	56	$224 (56 \times 4)$	178	46

#### 4. OUTCOME AND EVALUATION

The result generated by this work is described in two stages, ST with feature extraction stage and classification stage. The ST is applied to enhanced image, which enhances specific structural features such as edges and textures. ST helps in capturing multi-scale and directional information for analyzing bone structures and then calculated features from ST image. In classification stage, calculated features are used in DLNN and ST images are used in CNN to distinguish between osteoporosis and healthy samples.

#### 4.1. Shearlet transform and feature extraction

Figure 2 illustrates the application of the ST on an enhanced image. Figure 2(a) represents a healthy bone (enhanced), where the internal texture details are not clearly visible. In contrast, Figure 2(b) shows the corresponding ST output, highlighting the internal structural features by emphasizing edges and patterns that reflect the trabecular network's integrity. Figure 2(c) depicts an osteoporotic bone (enhanced), exhibiting a less uniform texture compared to Figure 2(a). Similarly, Figure 2(d) presents the corresponding ST output, revealing fragmented structural patterns with more discontinuities and irregularities compared to Figure 2(b), this reflects the weakened and porous nature of osteoporotic bone. These results highlight the potential of ST in diagnosing osteoporosis by detecting and quantifying structural anomalies in bone images.

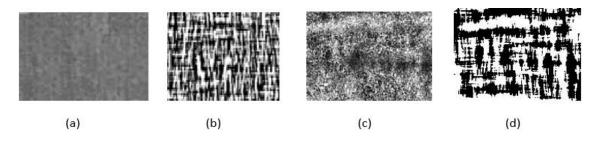


Figure 2. Input and output of the ST: (a) enhanced healthy bone image, (b) ST output for the healthy bone, (c) enhanced osteoporosis bone image and (d) ST output for the osteoporosis bone

The application of the ST provides several key metrics for performance evaluation like edge detection accuracy (EDA) assesses how effectively the transform identifies edges. Structural similarity index (SSI) quantifies the similarity between the transformed image and the enhanced. Feature extraction time (FET) measures the time required to extract features from the ST image. Compression ratio (CR) represents the ratio of the enhanced image dimensions to those of the transformed image. Noise reduction efficiency (NRE) evaluates the transformability to reduce noise while preserving essential features. Table 2 presents the results for these metrics across five sample ST images.

TD 11 0	D C	C C	1	•
Table 7	Pertormance	Of TIME	cample	1mages
Table 2.	Performance	OI II VC	Sampic	mages

Image ID	EDA	SSI	FET (Seconds)	CR	NRF (%)
Image - 01	0.92	0.88	1.5	2.8	85
Image - 02	0.89	0.86	1.7	2.8	82
Image - 03	0.94	0.89	1.6	3.0	88
Image - 04	0.90	0.87	1.4	2.7	83
Image - 05	0,91	0.88	1.5	2.9	84

The EDA values range from 0.89 to 0.94, demonstrating high edge detection accuracy across all images. Image-03 achieves the highest EDA (0.94), highlighting the technique of strong edge detection capability for that specific image. The SSI, which typically ranges from 0 to 1 (higher values indicate better similarity), varies between 0.86 and 0.89. This indicates consistent structural preservation. Image-03 also shows the highest SSI (0.89), suggesting greater retention of details. The FET ranges from 1.4 to 1.7 seconds, with Image-04 being the fastest (1.4 seconds), indicating computational efficiency with minimal variation. The CR values span from 2.7 to 3.0, with Image-03 achieving the highest value (3.0), demonstrating effective image compression while preserving essential features. Lastly, the NRE values range from 82% to 88%, with Image-03 exhibiting the highest noise reduction (88%), reflecting strong denoising capability.

Image-03 consistently outperforms other images across most metrics, achieving the highest EDA, SSI, CR, and NRE, making it the most effectively processed image in the dataset. All images exhibit high

EDA and SSI values, confirming the technique's strong performance in edge detection and structural preservation. Computational efficiency (FET) remains stable and reasonable, ranging from 1.4-1.7 seconds. Additionally, the high NRE ensures clear and noise-free outputs. These results demonstrate the technique's effectiveness and reliability across multiple performance metrics, highlighting its robustness and versatility for image analysis tasks.

The overall performance proposed work by providing mean values and standard deviations of the metrics as demonstrated in Table 3. The low standard deviations for EDA, FET, CR and NRF highlight the method robustness and reliability, performing consistently across various images. High accuracy specifies that the high mean EDA (0.912) and SSI (0.876) values indicate that the technique effectively identifies edges and preserves structural details. The low mean FET (1.54 seconds) suggests the method is computationally efficient and suitable for large-scale processing tasks. Quality assurance indicates that the high NRF mean (84.4%) ensures the outputs are clean and of high quality, while the CR (2.78) shows an excellent balance between compression and feature preservation, making the technique useful for storage. This highlights the technique as a promising choice for image texture analysis.

Table 3. Overall outcomes of ST

rusie s. o verum outcomes of s.										
Mean	Standard deviation									
0.912	0.018									
0.876	0.11									
1.54	0.11									
2.78	0.17									
84.4	2.05									
	Mean 0.912 0.876 1.54 2.78									

A Table 4 that displays an example of four Shearlet coefficients (ST-Cs) at two different scales and three different orientations for one image. The ST decomposes an image into multiple scales and orientations, to extract coefficients, used the Shearlet filter oriented at 0° to extract coefficients capturing horizontal (H) structures in the image. Used the Shearlet filter oriented at 90° to capture vertical (V) structures. Used Shearlet filters oriented at 45° to capture diagonal (D) features. The results show a clear dominance of horizontal structures, consistent vertical patterns and weaker diagonal features. This information is significant for understanding the nature of the images and tailoring techniques for task texture pattern recognition. Dominance of horizontal features are consistently higher across both scales, indicating the predominance of horizontal structures in the analyzed images. Diagonal feature strengths are moderate in strength, whereas diagonal features are the weakest, suggesting that the dataset or technique prioritizes horizontal and vertical edges over diagonal ones. Scale analysis, scale 1 (finer scale) generally has higher coefficients than scale 2 (coarser scale), which is expected since finer scales capture more detailed features, while coarser scales capture broader structures. The variation between scale 1 and scale 2 demonstrates the effectiveness of the ST in capturing both fine and coarse details, which is essential for texture analysis in feature extraction of medical image.

Table 4. ST coefficients of an image for two different scales

ST coefficients	Scale 1 - H	Scale 1 - V	Scale 1 - D	Scale 2 - H	Scale 2 - V	Scale 2 - D
Shearlet coefficient 1	0.12	0.08	0.05	0.11	0.07	0.04
Shearlet coefficient 2	0.10	0.09	0.06	0.10	0.08	0.05
Shearlet coefficient 3	0.11	0.07	0.05	0.09	0.06	0.03
Shearlet coefficient 4	0.13	0.09	0.06	0.12	0.08	0.05

Figure 3 illustrates a grouped bar chart that visualizes the ST-Cs across different scales and orientations. The x-axis represents the various pairings, combinations of scales and orientations. There are two scales (Scale 1 and Scale 2) and three orientations within each scale (H, V and D). The y-axis shows the values of the ST-Cs values. The plot is significant for several reasons, it allows us to visually compare how the ST-Cs vary across different scales and orientations and have significant implications in osteoporosis detection, provide a robust representation of image textures at various scales and directions. By leveraging multi-directional and multi-scale, this approach enhances the sensitivity and specificity of osteoporosis detection.

A Table 5 indicates features are calculated from the ST image sample image-01 at two scales and three different orientations, these features are useful for osteoporosis image classification. Variations in MC across different orientations and scales can reveal directional texture information. Lower SDC at certain scales could suggest smoother regions, while higher values could indicate detailed or noisy areas. Energy

4326 ISSN: 2088-8708

differences across scales could highlight regions with dominant features (e.g., edges or corners). Increasing entropy with certain orientations or scales may reveal regions with intricate details. High skewness could point to outlier features or abrupt changes in intensity. Higher kurtosis could imply regions with highly concentrated features or edges.

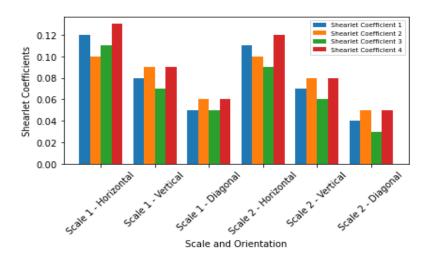


Figure 3. ST coefficients grouped bar chart

Table 5	Calculated	features	from	ST image
raine	Calculated	icatures.	пош	or image

Image - ID	S	Or	MC	SDC	Е	Et	Sk	K	С	Cor	Н
Image - 01	1	$0_{\rm o}$	0.045	0.012	0.034	1.54	0.23	3.21	0.45	0.76	0.89
Image - 01	1	90°	0.0325	0.011	0.036	1.56	0.25	3.19	0.46	0.73	0.86
Image - 01	1	45°	0.0548	0.010	0.039	1.58	0.27	3.15	0.49	0.72	0.85
Image - 01	2	$0_{\rm o}$	0.038	0.015	0.028	1.49	0.20	3.30	0.40	0.78	0.83
Image - 01	2	90°	0.040	0.013	0.031	1.51	0.22	3.28	0.42	0.73	0.86
Image - 01	2	45°	0.043	0.014	0.033	1.53	0.24	3.25	0.44	0.75	0.87

Regions with high contrast might correspond to areas of interest for feature detection. Correlation changes across scales could reveal patterns or periodic structures. Homogeneity might decrease in areas with sharp edges or complex textures. Directional dependence refers to variations across orientations (0°, 45°, 90°), which reveal anisotropic structures or dominant directions within the image. Scale sensitivity highlights differences across scales, showing how image features change in size or resolution. Combined metrics offer a comprehensive representation of regions with rich textures or prominent features, making them especially valuable for tasks like image classification.

# 4.2. Osteoporosis classification

In this work, a DLNN is trained for osteoporosis classification using the computed texture feature vector as input, with each ST image input vector having a total length of 60. Similarly, the Shearlet transformed images are fed into CNN. Both networks demonstrated promising performance in differentiating between osteoporotic and non-osteoporotic lumbar spine images (L1-L4). The classification task utilized a sigmoid activation function in the output layer for binary classification (osteoporosis vs. healthy). The dataset was divided into training and testing subsets, ensuring balanced classes to mitigate bias. Binary cross-entropy was employed as the loss function, with the Adam optimizer used for model training. The model was trained on the training set and hyperparameters were fine-tuned using a validation set.

Figure 4 graph illustrates CNN training and validation accuracy of a model over 20 epochs. Initially, both training and validation accuracy increase steadily, showing consistent improvement as the model learns. Around the middle epochs, the training accuracy continues to rise, while validation accuracy shows minor fluctuations, indicating slight overfitting or model instability. Toward the later epochs, validation accuracy aligns closely with training accuracy, suggesting good generalization and effective learning. Overall, the model achieves high performance, with accuracy more than 95% by the final epochs for both training and validation.

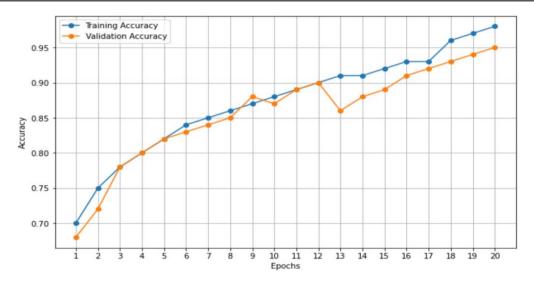


Figure 4. CNN training and validation over epochs

#### 4.3. Performance calculation

The metrics are calculated [19][20]as follows: Accuracy is (TP + TN) / (TP + TN + FP + FN). Precision is TP / (TP + FP). Recall (Sensitivity) is TP / (TP + FN). Specificity is TN / (TN + FP). F1-score is  $2 \times (Precision \times Recall) / (Precision + Recall)$ .

The baseline model uses traditional features without Shearlet transform, serving as a comparison point. Shearlet with DLNN demonstrates the improvement when incorporating Shearlet statistical values. Shearlet with CNN shows further enhancement when combining ST images with a CNN, which is adept at handling image data and enhances the classification model performance. The multi-scale and multi-directional features captured by the Shearlet transform enhance the ability to distinguish between osteoporosis and healthy individuals in enhanced images. Especially with CNN architectures, substantial gains in accuracy, precision, recall, specificity and F1-score as in Table 6. CNN generally performs better with ST image data due to its ability to capture spatial relationships. Based on the provided performance, the ST with CNN model is generally the best model for this task. It leverages the raw Shearlet coefficients more effectively through convolutional layers, which are adept at extracting spatial features from images. This leads to higher accuracy, precision, recall, specificity, and F1-score compared to the DLNN model.

Table 6. The results of the three approaches

			1.1		
Model	Accuracy	Precision	Recall	Specificity	F1-score
Baseline (No shearlet	91%	90%	89%	91%	89.5%
ST+ DLNN	92%	91%	90%	93%	90.5%
ST+ CNN	96%	95%	94%	96%	95.5%

Figure 5 a confusion matrix is a useful for evaluating the performance of a classification model, The confusion matrix shows the classification results of a binary model distinguishing between osteoporosis and healthy individuals. The model correctly classified 45 osteoporosis cases and 39 healthy cases, indicating high sensitivity and specificity. There is only 1 false positive (a healthy individual misclassified as osteoporosis) and 1 false negative (Osteoporosis case misclassified as healthy), demonstrating excellent performance. The low error rate suggests the model is well-trained and effective in distinguishing the two classes. Overall, the confusion matrix reflects a highly accurate classification system.

Figure 6 illustrates a graphical representation of different models' receiver operating characteristic (ROC) and area under curve (AUC) [27], needs the true positive rate (TPR) and the false positive rate (FPR) for each model. The TPR is equivalent to recall, and the FPR can be calculated as 1–specificity. Shearlet with CNN is the best ROC curve with the highest AUC, meaning it is the most effective at distinguishing between classes. Shearlet with DLNN performs better than the baseline but not as well as the CNN model. Baseline (No Shearlet) is good compared to the existing techniques, however in this case it is the least effective among the three in terms of the ROC and AUC. These metrics help in understanding the trade-offs each model makes in terms of true positives versus false positives, which is crucial in selecting the best model depending on the applications. Table 7 illustrates the proposed method (ST + CNN) outperforms the existing DLNN

4328 □ ISSN: 2088-8708

approaches in all three metrics accuracy, specificity, and sensitivity when analyzing X-ray lumbar images. This suggests that the new method is not only more accurate but also better at identifying both the presence and absence of osteoporotic conditions in these medical images.

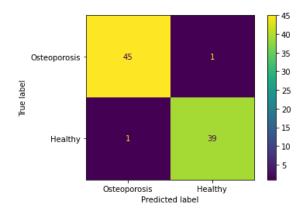


Figure 5. Performance of a CNN model

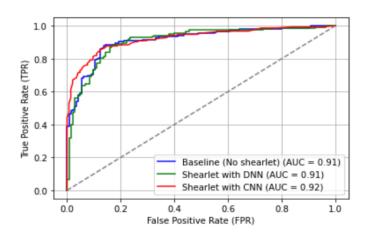


Figure 6. AUC-ROC for different models

Table 7. Evaluation of performance of different approaches

Ref. No.	Methods	Modality	Accuracy	Specificity	Sensitivity
[2]	DLNN	X-ray lumbar images	89%	90%	87%
[3]	FCoTNet	X-ray spine image	85%	90%	88%
[11]	DLNN	X-ray lumbar images	82%,	84%	86%,
[28]	DLNN	CT images of the spine	95%	94%	93%
[29]	DLNN	CT images of the spine	84%	87%	90.%
[30]	HT- CNN	CT images of the spine	95%	93%	91%
[31]	DLNN	X-ray spine image	89%	87%	85%
Proposed	ST+ CNN	X-ray lumbar image	96%	96%	94%

# 5. CONCLUSION

Osteoporosis, a condition characterized by reduced bone density and increased fracture risk, poses a significant health challenge, especially among the elderly. Accurate and early diagnosis is crucial for effective treatment and fracture prevention. This study addresses this challenge by proposing an innovative framework that integrates ST spectral analysis with DLNN and CNN for osteoporosis classification in lumbar vertebrae (L1-L4) X-ray images.

The proposed method effectively extracts trabecular bone micro-architecture information using ST, providing meaningful texture features for classification. The DLNN utilizes texture features for automated osteoporosis classification, while the CNN directly processes ST images without feature

extraction. The experimental results demonstrate the effectiveness of this approach, achieving a high classification accuracy of 96% using CNN. This study highlights the potential of combining spectral analysis techniques with deep learning to improve osteoporosis diagnosis. The framework enhances diagnostic accuracy and clinical decision making, offering a robust and reliable solution for early osteoporosis detection and management.

4329

To improve the accuracy and robustness of osteoporosis classification, future research can focus on the following areas. Expanding the dataset with a larger and more diverse population can enhance the generalizability of the model. Exploring more advanced neural network models, such as transformers or hybrid architecture, may further enhance classification performance.

# **FUNDING INFORMATION**

Authors state no funding involved.

# **AUTHOR CONTRIBUTIONS STATEMENT (mandatory) (10 PT)**

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	С	M	So	Va	Fo	I	R	D	0	E	Vi	Su	P	Fu
Poorvitha Hullukere	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			
Ramakrishna														
Chandrakala Beturpalya				$\checkmark$	$\checkmark$		✓			$\checkmark$		$\checkmark$	$\checkmark$	
Muddaraju														
Bhanushree Kothathi													$\checkmark$	
Jayaramu														
Shobha													$\checkmark$	
Narasimhamurthy														

Fo: Formal analysis E: Writing - Review & Editing

# CONFLICT OF INTEREST STATEMENT

None of the authors have any conflicts of interest, whether financial, personal, or professional.

#### DATA AVAILABILITY

- The data that support the findings of this study are available on request from the corresponding author BM. The data, which contains information that could compromise the privacy of research participants, are not publicly available due to certain restrictions.
- The data that support the findings of this study are completely provided by a prestigious private hospital in Bangalore, Karnataka, India, upon reasonable request from authors.

# REFERENCES

- [1] Q. Dong *et al.*, "Deep learning classification of spinal osteoporotic compression fractures on radiographs using an adaptation of the genant semiquantitative criteria," *Academic Radiology*, vol. 29, no. 12, pp. 1819–1832, Dec. 2022, doi: 10.1016/j.acra.2022.02.020.
- [2] B. Zhang et al., "Deep learning of lumbar spine X-ray for osteopenia and osteoporosis screening: A multicenter retrospective cohort study," Bone, vol. 140, p. 115561, Nov. 2020, doi: 10.1016/j.bone.2020.115561.
- [3] L. Xue et al., "Osteoporosis prediction in lumbar spine X-ray images using the multi-scale weighted fusion contextual transformer network," Artificial Intelligence in Medicine, vol. 143, p. 102639, Sep. 2023, doi: 10.1016/j.artmed.2023.102639.
- [4] H. Fan, J. Ren, J. Yang, Y.-X. Qin, and H. Ling, "Osteoporosis prescreening using panoramic radiographs through a deep convolutional neural network with attention mechanism," arXiv preprint arXiv:2110.09662, 2021.
- [5] F. Wang *et al.*, "Lumbar bone mineral density estimation from chest X-ray images: anatomy-aware attentive multi-ROI modeling," *IEEE Transactions on Medical Imaging*, vol. 42, no. 1, pp. 257–267, Jan. 2023, doi: 10.1109/TMI.2022.3209648.
- [6] H. T. Nguyen et al., "VinDr-SpineXR: A deep learning framework for spinal lesions detection and classification from radiographs," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes

- in Bioinformatics), vol. 12905 LNCS, pp. 291-301, 2021, doi: 10.1007/978-3-030-87240-3\_28.
- [7] S. Soegijoko, J. T. Pramudito, S. Soegijoko, T. R. Mengko, F. I. Muchtadi, and R. G. Wachjudi, "Trabecular pattern analysis of proximal femur radiographs for osteoporosis detection," *Journal of Biomedical & Pharmaceutical Engineering*, vol. 1, pp. 45–51, 2007, [Online]. Available: https://www.researchgate.net/publication/229035119.
- [8] H. Jeong, J. Kim, T. Ishida, M. Akiyama, and Y. Kim, "Computerised analysis of osteoporotic bone patterns using texture parameters characterising bone architecture," *The British Journal of Radiology*, vol. 86, no. 1021, pp. 20101115–20101115, Jan. 2013, doi: 10.1259/bjr.20101115.
- [9] N. Hong et al., "Deep-learning-based detection of vertebral fracture and osteoporosis using lateral spine X-ray radiography," *Journal of Bone and Mineral Research*, vol. 38, no. 6, pp. 887–895, 2023, doi: 10.1002/jbmr.4814.
- [10] S. Mu, J. Wang, and S. Gong, "Application of medical imaging based on deep learning in the treatment of lumbar degenerative diseases and osteoporosis with bone cement screws," *Computational and Mathematical Methods in Medicine*, vol. 2021, pp. 1– 10, Oct. 2021, doi: 10.1155/2021/2638495.
- [11] J. M. Kim, J. B. Woo, and H. Y. Kim, "Prediction model of spinal osteoporosis using lumbar spine X-Ray from Transfer learning deep convolutional neural networks," *The Nerve*, vol. 10, no. 2, pp. 98–106, Oct. 2024, doi: 10.21129/nerve.2024.00598.
- [12] S. H. Kong et al., "Development of a spine X-Ray-based fracture prediction model using a deep learning algorithm," Endocrinology and Metabolism, vol. 37, no. 4, pp. 674–683, Aug. 2022, doi: 10.3803/EnM.2022.1461.
- [13] C. Kim et al., "Comparative efficacy of anteroposterior and lateral X-ray based deep learning in the detection of osteoporotic vertebral compression fracture," Scientific Reports, vol. 14, no. 1, p. 28388, Nov. 2024, doi: 10.1038/s41598-024-79610-w.
- [14] M. Mebarkia, A. Meraoumia, L. Houam, and S. Khemaissia, "X-ray image analysis for osteoporosis diagnosis: From shallow to deep analysis," *Displays*, vol. 76, p. 102343, Jan. 2023, doi: 10.1016/j.displa.2022.102343.
- [15] Q. Yuan and S. Dai, "Adaptive histogram equalization with visual perception consistency," *Information Sciences*, vol. 668, p. 120525, May 2024, doi: 10.1016/j.ins.2024.120525.
- [16] Y. Yamaguchi et al., "Edge-preserving smoothing filter using fast M-estimation method with an automatic determination algorithm for basic width," Scientific Reports, vol. 13, no. 1, p. 5477, Apr. 2023, doi: 10.1038/s41598-023-32013-9.
- [17] A. Khare, M. Khare, and R. Srivastava, "Shearlet transform based technique for image fusion using median fusion rule," Multimedia Tools and Applications, vol. 80, no. 8, pp. 11491–11522, Mar. 2021, doi: 10.1007/s11042-020-10184-1.
- [18] V. Subbiah Parvathy, S. Pothiraj, and J. Sampson, "A novel approach in multimodality medical image fusion using optimal shearlet and deep learning," *International Journal of Imaging Systems and Technology*, vol. 30, no. 4, pp. 847–859, Dec. 2020, doi: 10.1002/ima.22436.
- [19] Y. A. Hamza, N. E. Tewfiq, and M. Q. Ahmed, "An enhanced approach of image steganographic using discrete Shearlet transform and secret sharing," *Baghdad Science Journal*, vol. 19, no. 1, p. 0197, Feb. 2022, doi: 10.21123/bsj.2022.19.1.0197.
- [20] K. Meshkini and H. Ghassemian, "Texture classification using Shearlet transform and GLCM," in 2017 Iranian Conference on Electrical Engineering (ICEE), May 2017, pp. 1845–1850, doi: 10.1109/IranianCEE.2017.7985354.
- [21] C. Affonso, A. L. D. Rossi, F. H. A. Vieira, and A. C. P. de L. F. de Carvalho, "Deep learning for biological image classification," Expert Systems with Applications, vol. 85, pp. 114–122, Nov. 2017, doi: 10.1016/j.eswa.2017.05.039.
- [22] S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi, and J. A. Benediktsson, "Deep learning for hyperspectral image classification: An overview," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 6690–6709, Sep. 2019, doi: 10.1109/TGRS.2019.2907932.
- [23] T. Guo, J. Dong, H. Li, and Y. Gao, "Simple convolutional neural network on image classification," in 2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)(, Mar. 2017, pp. 721–724, doi: 10.1109/ICBDA.2017.8078730.
- [24] M. Hussain, J. J. Bird, and D. R. Faria, "A study on CNN transfer learning for image classification," 2019, pp. 191–202.
- [25] G. Hatem, J. Zeidan, M. Goossens, and C. Moreira, "Normality testing methods and the importance of skewness and kurtosis in statistical analysis," *BAU Journal - Science and Technology*, vol. 3, no. 2, Jun. 2022, doi: 10.54729/KTPE9512.
- [26] E. Lhermitte, M. Hilal, R. Furlong, V. O'Brien, and A. Humeau-Heurtier, "A deep learning and entropy-based texture features for color image classification," *Entropy*, vol. 24, no. 11, p. 1577, Oct. 2022, doi: 10.3390/e24111577.
- [27] J. Muschelli, "ROC and AUC with a binary predictor: a potentially misleading metric," *Journal of Classification*, vol. 37, no. 3, pp. 696–708, Oct. 2020, doi: 10.1007/s00357-019-09345-1.
- [28] R. Dzierżak and Z. Omiotek, "Application of deep convolutional neural networks in the diagnosis of osteoporosis," *Sensors*, vol. 22, no. 21, p. 8189, Oct. 2022, doi: 10.3390/s22218189.
- [29] L. Xue et al., "A dual-selective channel attention network for osteoporosis prediction in computed tomography images of lumbar spine," Acadlore Transactions on AI and Machine Learning, vol. 1, no. 1, pp. 30–39, Nov. 2022, doi: 10.56578/ataiml010105.
  [30] J. Liu et al., "Hybrid transformer convolutional neural network-based radiomics models for osteoporosis screening in routine
- [30] J. Liu et al., "Hybrid transformer convolutional neural network-based radiomics models for osteoporosis screening in routine CT," BMC Medical Imaging, vol. 24, no. 1, p. 62, Mar. 2024, doi: 10.1186/s12880-024-01240-5.
- [31] Y. Ono *et al.*, "A deep learning-based model for classifying osteoporotic lumbar vertebral fractures on radiographs: a retrospective model development and validation study," *Journal of Imaging*, vol. 9, no. 9, p. 187, Sep. 2023, doi: 10.3390/jimaging9090187.

# **BIOGRAPHIES OF AUTHORS**



Poorvitha Hullukere Ramakrishna © received the B.E. degree in computer science and engineering from Visvesvaraya Technological University, Belagavi and an MTech degree in computer science and engineering from Visvesvaraya Technological University, Belagavi. She is currently pursuing a Ph.D. from Dayananda Sagar College of Engineering under Visvesvaraya Technological University, Belagavi. Her area of interest includes image processing, machine learning, and deep learning. She can be contacted at email: poorvitharam26@gmail.com.



Chandrakala Beturpalya Muddaraju is is working as an associate professor at the Department of Information Science and Engineering, Dayananda Sagar College of Engineering, Bengaluru, Karnataka, India. She received her Ph.D. Degree in Computer and Information Science and Engineering, Visvesvaraya Technological University, Belagavi, Karnataka. She has been working as an associate professor for the past 20 years. Her research interests are cloud computing, network security, cybersecurity, artificial intelligence, and machine learning. An active member of IEEE, has published 48 papers in journals and conferences and published 7 patents. Guiding Research Scholars in various areas. She can be contacted at this email: chandrakalabm@yahoo.com.



Bhanushree Kothathi Jayaramu received the B.E., M Tech. and Ph.D. degrees from VTU University, Karnataka. She is currently working as associate professor, the Department of Computer Science and Engineering, Bangalore Institute of Technology, Bengaluru, Affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India. She has teaching experience of more than 15 years. She has published more than 16 research papers. She can be contacted at this email: kjbhanushree@bit-bangalore.edu.in.



Shobha Narasimhamurthy is working as an associate professor in the Department of Computer Science & Design, Dayananda Sagar College of Engineering, Bengaluru. She obtained her PhD in computer and information science from Visvesvaraya Technological University, Karnataka. She has published around 18 papers in international/national journals and conferences, three book chapters and published six patents. Her research interests include data mining, machine learning, and data science. She can be contacted at this email: shobha-csd@dayanandasagar.edu.