International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. #, No.# (Mon-20..)

Detection of Osteoporosis from Calcaneal Radiograph Images using Wavelet Texture Analysis and Machine Learning

Sai Kiran K S[1,2], Anu Shaju Areeckal[3]*

[1] *Department of Bioinformatics, Manipal School of Life Sciences, Manipal Academy of Higher Education, Manipal 576104, India*
[2] *Novo Nordisk GBS, Cignus, Whitefield, Bengaluru, Karnataka 560066, India*
[3] *Department of Electronics and Communication Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal 576104, India*
**Corresponding author*

E-mail address: saikiranks14@gmail.com, anu.areeckal@manipal.edu

Received ## Mon. 20##, Revised ## Mon. 20##, Accepted ## Mon. 20##, Published ## Mon. 20##

Abstract: Osteoporosis, a disorder defined by decreased bone mineral content and changes in bone microarchitecture, poses a challenge for accurate classification using X-ray images. This work aims to extract texture features from calcaneal radiographs and select the best texture features which can be used to train the machine learning classifier models for detection of osteoporosis. This work is based on multiresolution analysis and microstructural analysis to characterize trabecular bone microarchitecture from calcaneal radiograph. The image is transformed to extract the feature details using a two-level wavelet decomposition. Structural texture methods such as Local Binary Pattern, fractal dimension and Gabor filter are applied to the wavelet decomposed images. The most discriminating texture features are selected using independent sample t-test and feature selection methods. Machine learning models are constructed by training the classifiers using the best texture features to classify healthy images from osteoporotic images. The effectiveness of the proposed approach is evaluated using a public challenge dataset comprising calcaneal radiographic images. Notably, the best classification is obtained with k-Nearest Neighbour trained with the features selected using forward feature selection, with an accuracy rate of 78.24%. The results indicate the potential of the proposed approach as a possible alternative tool for screening osteoporosis.

Keywords: Osteoporosis, Health care, Machine learning, Texture analysis, Classification

1. INTRODUCTION

Osteoporosis is a prevalent metabolic disease characterized by reduced bone strength, primarily defined by bone mineral density (BMD). Osteoporosis is operationally recognized by the World Health Organization as having a BMD that is 2.5 standard deviations or more below average (T-score less than -2.5) [1]. The global impact of osteoporosis is significant, with an estimated 200 million individuals affected by the condition. The International Osteoporosis Foundation's statistics show the prevalence of osteoporotic fractures are all around the world. According to statistics, one in three women aged 50 or older and one in five males will have an osteoporotic fracture over their lifetime [2]. In India, the prevalence of osteoporosis among adults is estimated to be 22.9 percent. The prevalence is higher in females (26.3 percent) compared to males (10.9 percent), surpassing the global prevalence of 18.3 percent [1].

Osteoporosis poses a significant medical and socio-economic threat as it leads to a systemic deterioration of bone mass and microarchitecture, increasing the susceptibility to fragility fractures. In individuals with osteoporosis, the production of new bone fails to keep pace with the loss of old bone, resulting in decreased BMD. This reduced BMD makes even minor stresses potentially fracturing, and the associated chronic pain severely limits daily activities [3]. Common sites for osteoporosis-related fractures include the hip, wrist, and spine [4].

Various techniques, such as High-Resolution Peripheral QCT (HR-pQCT), Magnetic resonance imaging (MRI), Quantitative Computed Tomography

E-mail: anu.areeckal@manipal.edu

(QCT), Dual X-ray Absorptiometry (DXA), digital X-ray radiogrammetry (DXR) and quantitative ultrasound (QUS) are employed for diagnosing osteoporosis [5]. DXA is regarded as the gold standard technique; however, its availability is limited in low-income economies, and the cost of scans is high [6]. Cortical radiogrammetry, a cost-effective technique using radiographs, has proven useful in detecting bone loss. For the purpose of identifying bone disorders including osteoporosis and rheumatoid arthritis, the advanced cortical radiogrammetric method known as Digital X-ray Radiogrammetry (DXR) has been widely explored [7]. Cortical radiogrammetry does not examine the texture of trabecular bone, it only evaluates cortical bone characteristics. Before an apparent decrease in cortical bone is seen, osteoporosis first affects the trabecular bone structure. Additionally, measuring BMD alone is not a reliable way to predict fracture risk. Non-BMD variables like microarchitecture can also be used to describe the quality and strength of bone. Given the expense and limitations of existing diagnostic techniques, there is a pressing need to develop a low-cost device that utilizes bone microarchitecture for early detection and diagnosis of osteoporosis.

According to several studies, imaging is a potential way to support the diagnosis and aid the practitioner in making decisions, according to a number of studies. Texture analysis and classification enables an easy and less intrusive technique to describe the micro-architecture of the bone on X-ray images, which should help assess the risk of bone fracture and aid in the early detection of osteoporosis. Studies have demonstrated the value of calcaneal trabecular bone texture analysis in the diagnosis and early detection of fractures associated with osteoporosis and have reported the ability to distinguish between osteoporotic and healthy participants using trabecular texture characteristics from calcaneal radiographs.

Reference [8] used calcaneal radiographs of 174 women which included 87 images in both healthy and osteoporotic category. An effective fractal dimension estimator known as the anisotropic piecewise Whittle estimator was combined with an anisotropic fractional Brownian motion model in their oriented analysis technique and obtained an accuracy of 71.8% with 72% sensitivity, 71% specificity and AUC of 78%. Reference [9] used calcaneal TCB dataset which included 58 images in both healthy and osteoporotic categories. They made use of the gabor filter bank, which is made up of filters with 4 different scales and 4 orientations. The 1D-LBP histograms are then acquired. LBPs are calculated by taking into consideration 8 neighbors with a 1 pixel separation and obtained an accuracy of 72.71%.

Reference [10] used calcaneal TCB dataset which included 58 images in both healthy and osteoporotic categories. Multifractal analysis has been utilized to describe the bone texture. Here, local texture differences are effectively described by a set of local fractal dimensions rather than a single global fractal dimension. The effectiveness of the box-counting approach and the regularization dimension method are evaluated. The performance is then enhanced by combining both the characteristics and obtained accuracy of 55% with 59% sensitivity and 52% specificity.

Reference [11] used the ROIs that are taken from the trabecular region of individuals with varied ages and osteoporotic diseases using X-rays. They developed a fractal model that employs the differential box-counting method to calculate the fractal dimension (FD), which is done post image preprocessing step that ensures a reliable estimation approach and obtained a better result with p value of 4.8368e−4.

Reference [12] presented sparse representation-based technique for distinguishing healthy from diseased states using medical imaging patterns. Two classifiers based on log likelihood function and maximum a posteriori probability were developed. In order to handle the approximation problem's irregularities and construct a classifier ensemble that would produce more precise numerical answers than traditional sparse assessments of the entire spatial domain of the pictures, they suggest a spatial block decomposition approach from which they obtained accuracy of 67.8% with specificity 65.5% sensitivity 70.1% and AUC of 65%. Reference [13] used calcaneal TCB dataset which contains 58 images in both healthy and osteoporotic categories. Multifractals are used in their research to describe the trabecular bone microstructure.

In order to assess the overall regularity of the pixels, Hausdorff dimensions are first calculated for each Holder exponent. Finally, the Hausdorff dimensions are used to calculate lacunarity and obtained accuracy of 59% with 59% sensitivity and 59% specificity. Reference [14] used calcaneal X-ray image dataset which contains 87 images in both normal and osteoporosis categories. The proposed method follows different steps. The radiographic images are binarized after going through a median filtering stage of preprocessing. After computing the multi-fractal spectrum, numerous features are obtained to describe the microarchitecture of the trabecular bone. Finally, the two groups of osteoporotic as well as healthy radiographs are classified using the obtained features and obtained accuracy of 98.01%, with 97.26% Sensitivity, 98.78% Specificity and AUC of 98.37 %.

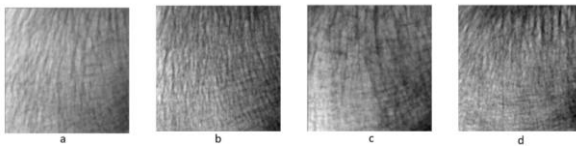

Figure 1. Calcaneal radiographs obtained from the dataset: (a) and (b) are calcaneal radiographs of class 0 category, and (c) and (d) are calcaneal radiographs of class 1.

Deep learning methods were used to predict osteoporosis. References [15]-[17] have used deep learning approach for classification of X-ray images as osteoporotic or healthy. Reference [15] obtained 93.9% sensitivity, 94.5% specificity and AUC of 97.5%. Reference [16] attained testing accuracy of 82% and validation accuracy of 84% on untrained test radiographic images. Reference [17] obtained an AUC of 76.80% and accuracy of 72.67%.

The proposed work aims to enhance the accuracy and precision of osteoporosis diagnostic tools by leveraging texture features extracted from calcaneal radiographs. By identifying significant texture features, we intend to train machine learning models and develop a classifier that can effectively differentiate between healthy and osteoporotic images. This approach holds promise in improving the diagnostic capabilities and overall effectiveness of osteoporosis detection.

2. MATERIALS AND METHODS

A. *Dataset*

The images used in this paper are obtained from a challenge dataset, released by International Society for Biomedical Imaging (ISBI) in 2014. This challenge involves applying texture analysis to classify osteoporotic patients from normal individuals based on X-ray images. The dataset consists of radiographs of calcaneal bone of two different population. Patients with osteoporosis and healthy individuals constitute the reference population. The images are of 400x400 pixels in 16-bit format. There are 58 photos in class 0 and 58 images in class 1 in the dataset. Fig. 1 shows some sample images.

B. *Methodology*

The proposed approach operates using different stages, as shown in Fig. 2. Prior to using a median filter to reduce noise content, the images are first preprocessed by increasing contrast using Contrast Limited Adaptive Histogram Equalization (CLAHE). Then, using various feature extraction approaches, texture features are obtained to characterize the trabecular bone microarchitecture. Feature selection strategies are used to choose the relevant features. Finally, osteoporotic and healthy radiographs are classified by training machine learning classifiers using the retrieved features.

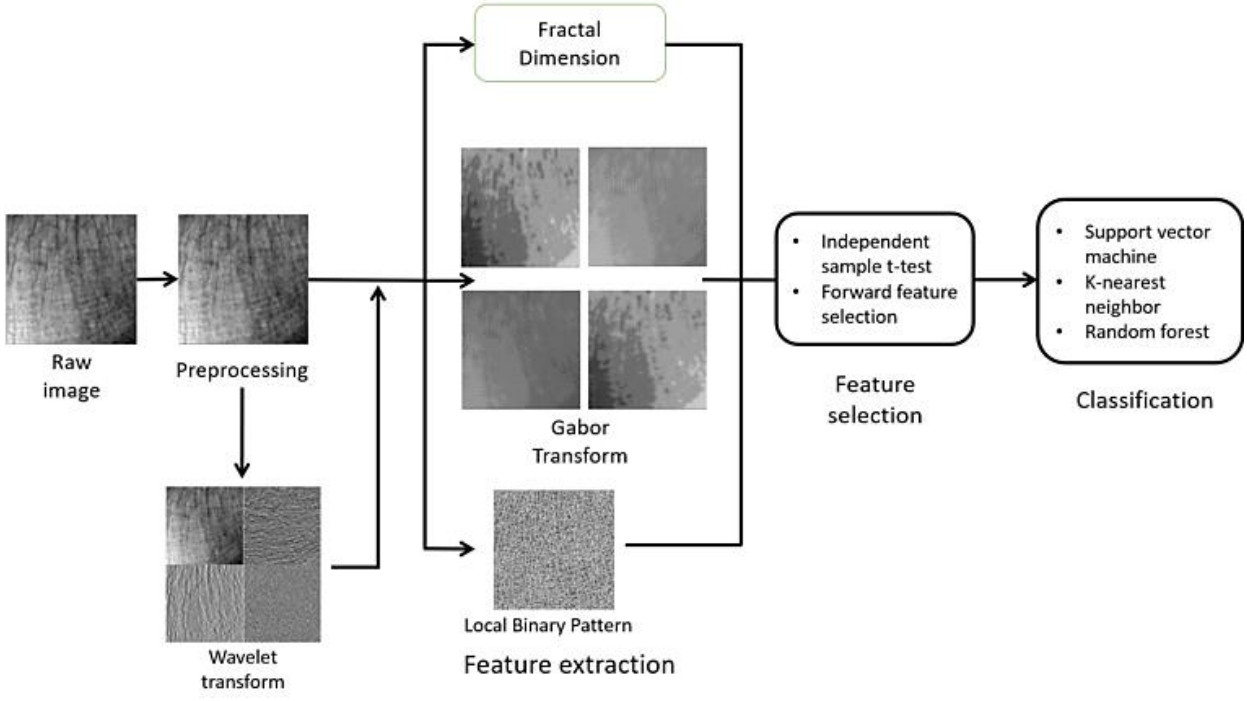

Figure 2. Methodology of the proposed work: The raw image is preprocessed and features are extracted using different texture analysis techniques, followed by feature selection. Classifiers are trained with the best texture features to classify healthy and osteoporotic images.

C. Preprocessing

Bone X-ray scans reveal striking similarities between osteoporotic patients and healthy individuals. Each picture is initially preprocessed in order to enhance image quality and improve the proposed technique's capacity to distinguish between two separate classes. A nonlinear median filter of size 3x3 is used on each grayscale picture to eliminate the impulse noise without changing the frequencies which are important for classifying osteoporosis. The impulsive noise created during acquisition is primarily eliminated by the median filter. The image intensity levels are then normalized in order to improve the contrast.

D. Feature extraction

Gravity (tension) as well as walking force (compression), which are applied to the heel, cause anisotropic qualities in the bone structure. The bone structure has anisotropic characteristics as a result of these forces. Additionally, normal (dense) and osteoporotic bones exhibit different levels of granularity, emphasizing the importance of multiresolution analysis that considers orientation and scale variations, as well as microstructural analysis [9]. To characterize the trabecular bone structure, several analysis methods based on fractal, structural, and texture transforms are employed.

In this proposed method, we utilize various features to analyze the trabecular features extracted from the preprocessed calcaneal X-ray images. Specifically, we employ structural features such as local binary pattern (LBP), fractal features such as fractal dimension (FD), and transform-based features including Gabor transform (GT) and discrete wavelet transform (DWT). These features enable a comprehensive analysis of the trabecular bone structure, considering its unique properties and variations.

The Local Binary Pattern (LBP) approach utilize the statistical distribution of local patterns to characterize textures. It assigns labels to pixels by comparing their intensity values with the surrounding pixels and it computed a rotation-invariant metric known as the uniformity measure U. Patterns are given the LBP code if their U value is less than two, indicating that the center pixel is labeled as uniform [18]. Three different neighbor configurations are considered, namely 8, 16, and 24, with respective pixel spacing of 1, 2, and 3. Energy and entropy are then extracted from the resulting LBP images, generating a feature vector with 6 features for each preprocessed image.

Gabor transform is employed to represent and differentiate textures by examining the presence of specific frequency content in the localized area surrounding the point or region of interest [9]. Gabor filters are applied, resulting in Gabor images of different scales and orientations. Mean and standard deviation features are extracted from the set of Gabor filters with various orientations and frequencies, resulting in a feature vector with 16 features for each preprocessed image.

Fractal Dimension (FD) analysis, suitable for evaluating bone microstructure on radiographs, utilize the box-counting algorithm to characterize trabecular bone patterns [19]. Hurst coefficients are extracted from each image, generating a feature vector containing 4 coefficients.

Wavelet transform (WT) decomposes the images into four sub-images by convolving them with specific filters. These sub-images capture frequency information across various frequency ranges [20]. In this work, two-level decomposition is applied to each preprocessed image, resulting in a set of 7 images as shown in Fig. 3.

LBP, Gabor transform, and FD analysis are then applied to all 7 images, generating a feature vector with 96 features. This feature extraction process is repeated for all preprocessed images, resulting in a comprehensive set of features that capture different aspects of the trabecular bone structure and texture.

E. Feature selection

The extracted feature vectors may contain features that are not significant. It is important to remove any extraneous features since they might lower classification accuracy and make the classifier more complicated. It is possible to choose the most important and dominating features using a variety of feature selection techniques. Here, an independent sample t-test and forward feature selection method are used for feature selection.

In this work, the p-value of each feature obtained is calculated using an independent sample t-test. According to the test, features with a p-value of 0.05 or less are deemed to be significant [11]. Hence the feature with significance value i.e. p-value of 0.05 or less are considered to form feature vectors, which are used for training classifiers.

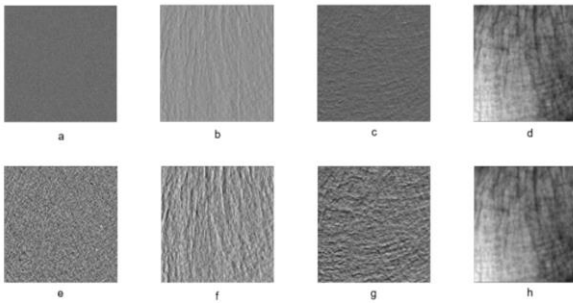

Figure 3. Two-level Wavelet Decomposition: (a) High High (HH), (b) High Low (HL), (c) Low High (LH), and (d) Low Low (LL) sub-band images of the first level of Haar decomposition. (e)-(h) are the HH, HL, LH and LL sub-band images of the second level, respectively.

Forward feature selection method is an iterative process where features are progressively added to a subset based on their impact on the performance of a machine learning model. It starts with an empty set and successively adds features that improve model performance until a stopping criterion is met. The process evaluates different feature subsets and selects the one that yields the best model performance [21]. We use this method to obtain a feature vector containing significant features which showed best accuracy amongst all others.

F. Classification

The different classifiers used are support vector machine (SVM), K-nearest neighbor (KNN) classifier and random forest (RF) classifier. For KNN, 6 nearest neighbours are considered. For the RF classifier, the number of estimators varies depending on the number of features. If the features are less than 15, 10 estimators are used, while if the features are more than 25, 100 estimators are used. This choice helps optimize the performance of the classifier based on the complexity of the feature set.

Due to the limited amount of training data, a 5-fold stratified cross-validation (CV) is performed. Utilizing performance measures obtained from the confusion matrix, the trained classifiers' effectiveness is assessed. The output findings are divided into four categories using a confusion matrix: true positive (TP), false positive (FP), true negative (TN), and false negative (FN). Sensitivity (Sn), specificity (Sp), and accuracy (Acc) are the performance measures employed, as given in Equations 1-3. By analyzing these performance measures, the classifiers' effectiveness in distinguishing between osteoporotic and healthy individuals based on the extracted features is evaluated.

$$\text{Sensitivity} = TP/TP + FN \quad (1)$$

$$\text{Specificity} = TN \,/\, (TN + FP) \quad (2)$$

$$\text{Accuracy} = (TP + TN) \,/\, (TP + TN + FP + FN) \quad (3)$$

3. RESULTS

Significant features are derived from different texture feature analysis methods through various feature selection techniques. Then the obtained feature vectors are used to train the classifier.

A. Feature selection

An independent sample t-test is conducted to determine the p-value for each feature, and all significant features, with p-value less than 0.05, is grouped as Feature Set-1 (FS1). As observed in Table 1, all features obtained from the LBP analysis are significant. One feature is significant in the GT analysis. In the FD analysis, two features are found to be significant.

TABLE I. INDEPENDENT SAMPLE T-TEST RESULTS OF FEATURES OBTAINED FROM VARIOUS TEXTURE METHODS

Feature method	Feature vector	p-value
Local binary pattern (LBP)	$LBP_{1,8}$_energy	0.0081*
	$LBP_{1,8}$_entropy	0.0146*
	$LBP_{2,16}$_energy	0.0006*
	$LBP_{2,16}$_entropy	0.0006*
	$LBP_{3,24}$_energy	0.0027*
	$LBP_{3,24}$_entropy	0.0024*
Gabor transform	$GT_{0,0.05}$_mean	0.6568
	$GT_{0,0.05}$_std	0.1446
	$GT_{0,0.4}$_mean	0.9213
	$GT_{0,0.4}$_std	0.4907
	$GT_{1,0.05}$_mean	0.9695
	$GT_{1,0.05}$_std	0.0078*
	$GT_{1,0.4}$_mean	0.7947
	$GT_{1,0.4}$_std	0.8602
	$GT_{2,0.05}$_mean	0.9343
	$GT_{2,0.05}$_std	0.7976
	$GT_{2,0.4}$_mean	0.8091
	$GT_{2,0.4}$_std	0.7128
	$GT_{3,0.05}$_mean	0.9359
	$GT_{3,0.05}$_std	0.2197
	$GT_{3,0.4}$_mean	0.7947
	$GT_{3,0.4}$_std	0.642
Fractal dimension	FDTA_HurstCoeff_1	0.0003*
	FDTA_HurstCoeff_2	0.0041*
	FDTA_HurstCoeff_3	0.2419
	FDTA_HurstCoeff_4	0.2526

*p-value < 0.05,
LBPn,m_feature- denotes the feature measured from LBP image obtained with radius of n and neighbourhood of m pixels,
GTn,m_mean- denotes the mean calculated from the Gabor image with threshold n and frequency m,
GTn,m_std- denotes the standard deviation calculated from the Gabor image with threshold n and frequency m,
FDTA_HurstCoeff- denotes Hurst coefficient of fractal dimension.

A 2-level wavelet decomposition is done on the original images and the LBP, Gabor transform and fractal dimension features are extracted from each sub-band image. The sub-images are Low-Low, Low-High, High-Low and High-High sub-bands of level-2 DWT decomposition, denoted as LL2, LH2, HL2 and HH2 images, respectively. The respective sub-images of level-2 are LL1, LH1, HL1 and HH1, respectively. The DWT analysis produced a total of 96 features, out of which 27 features are significant, as listed in Table 2 and 3.

TABLE II. INDEPENDENT SAMPLE T-TEST RESULTS OF FEATURES EXTRACTED FROM LEVEL-1 WAVELET TRANSFORM

Feature vector	Feature vector
LH1_$GT_{0,0.5}$_mean	HL1_$LBP_{1,8}$_energy*
LH1_$GT_{0,0.5}$_std	HL1_$LBP_{1,8}$_entropy*
LH1_$GT_{1,0.5}$_mean	HL1_$LBP_{2,16}$_energy*
LH1_$GT_{1,0.5}$_std	HL1_$LBP_{2,16}$_entropy*
LH1_$GT_{2,0.5}$_mean	HL1_FDTA_HurstCoeff_1*
LH1_$GT_{2,0.5}$_std	HL1_FDTA_HurstCoeff_2*
LH1_$GT_{3,0.5}$_mean	HH1_$GT_{0,0.5}$_mean
LH1_$GT_{3,0.5}$_std	HH1_$GT_{0,0.5}$_std
LH1_$LBP_{1,8}$_energy*	HH1_$GT_{1,0.5}$_mean
LH1_$LBP_{1,8}$_entropy*	HH1_$GT_{1,0.5}$_std
LH1_$LBP_{2,16}$_energy*	HH1_$GT_{2,0.5}$_mean
LH1_$LBP_{2,16}$_entropy*	HH1_$GT_{2,0.5}$_std
LH1_FDTA_HurstCoeff_1*	HH1_$GT_{3,0.5}$_mean
LH1_FDTA_HurstCoeff_2*	HH1_$GT_{3,0.5}$_std
HL1_$GT_{0,0.5}$_mean	HH1_$LBP_{1,8}$_energy*
HL1_$GT_{0,0.5}$_std	HH1_$LBP_{1,8}$_entropy*
HL1_$GT_{1,0.5}$_mean	HH1_$LBP_{2,16}$_energy
HL1_$GT_{1,0.5}$_std	HH1_$LBP_{2,16}$_entropy
HL1_$GT_{2,0.5}$_mean	HH1_FDTA_HurstCoeff_1*
HL1_$GT_{2,0.5}$_std	HH1_FDTA_HurstCoeff_2
HL1_$GT_{3,0.5}$_mean	
HL1_$GT_{3,0.5}$_std	

*p-value < 0.05,
XXi_GTn,m_feature- denotes the feature measured from Gabor transform (with threshold n and frequency m) of i^{th} level XX wavelet sub-image,
XXi_LBP n,m_feature- denotes the feature measured from LBP image (with radius n and neighbourhood m pixels) of i^{th} level XX wavelet sub-image,
XXi_FDTA_HurstCoeff_n- denotes the Hurst coefficient of fractal dimension of i^{th} level XX wavelet sub-image.

Next, a second feature extraction method is implemented using forward feature selection method. Through the forward feature selection method, we obtain a feature vector for each classifier that consists of significant features associated with the highest accuracy (FS2). The objective of this approach is to determine the most relevant features for our analysis. Initially, we start with a set of features and employ a stepwise procedure that iteratively adds the best-performing feature, based on the highest increase in accuracy. The process continues until no further improvement in performance is observed.

TABLE III. INDEPENDENT SAMPLE T-TEST RESULTS OF FEATURES EXTRACTED FROM LEVEL-2 WAVELET TRANSFORM

Feature vector	Feature vector
LL2_$GT_{0,0.5}$_mean	HL2_$GT_{0,0.5}$_mean
LL2_$GT_{0,0.5}$_std	HL2_$GT_{0,0.5}$_std
LL2_$GT_{1,0.5}$_mean	HL2_$GT_{1,0.5}$_mean
LL2_$GT_{1,0.5}$_std	HL2_$GT_{1,0.5}$_std
LL2_$GT_{2,0.5}$_mean	HL2_$GT_{2,0.5}$_mean
LL2_$GT_{2,0.5}$_std	HL2_$GT_{2,0.5}$_std
LL2_$GT_{3,0.5}$_mean	HL2_$GT_{3,0.5}$_mean
LL2_$GT_{3,0.5}$_std	HL2_$GT_{3,0.5}$_std
LL2_$LBP_{1,8}$_energy	HL2_$LBP_{1,8}$_energy*
LL2_$LBP_{1,8}$_entropy	HL2_$LBP_{1,8}$_entropy*
LL2_$LBP_{2,16}$_energy	HL2_$LBP_{2,16}$_energy
LL2_$LBP_{2,16}$_entropy	HL2_$LBP_{2,16}$_entropy
LL2_FDTA_HurstCoeff_1	HL2_FDTA_HurstCoeff_1*
LL2_FDTA_HurstCoeff_2	HL2_FDTA_HurstCoeff_2
LH2_$GT_{0,0.5}$_mean	HH2_$GT_{0,0.5}$_mean
LH2_$GT_{0,0.5}$_std	HH2_$GT_{0,0.5}$_std
LH2_$GT_{1,0.5}$_mean	HH2_$GT_{1,0.5}$_mean
LH2_$GT_{1,0.5}$_std	HH2_$GT_{1,0.5}$_std
LH2_$GT_{2,0.5}$_mean	HH2_$GT_{2,0.5}$_mean
LH2_$GT_{2,0.5}$_std	HH2_$GT_{2,0.5}$_std
LH2_$GT_{3,0.5}$_mean	HH2_$GT_{3,0.5}$_mean*
LH2_$GT_{3,0.5}$_std	HH2_$GT_{3,0.5}$_std
LH2_$LBP_{1,8}$_energy*	HH2_$LBP_{1,8}$_energy
LH2_$LBP_{1,8}$_entropy*	HH2_$LBP_{1,8}$_entropy
LH2_$LBP_{2,16}$_energy*	HH2_$LBP_{2,16}$_energy*
LH2_$LBP_{2,16}$_entropy*	HH2_$LBP_{2,16}$_entropy*
LH2_FDTA_HurstCoeff_1*	HH2_FDTA_HurstCoeff_1*
LH2_FDTA_HurstCoeff_2	HH2_FDTA_HurstCoeff_2

*p-value < 0.05,
XXi_GTn,m_feature- denotes the feature measured from Gabor transform (with threshold n and frequency m) of i^{th} level XX wavelet sub-image,
XXi_LBP n,m_feature- denotes the feature measured from LBP image (with radius n and neighbourhood m pixels) of i^{th} level XX wavelet sub-image,
XXi_FDTA_HurstCoeff_n- denotes the Hurst coefficient of fractal dimension of i^{th} level XX wavelet sub-image.

For the three classifiers, three distinct feature vectors are obtained. In the SVM classifier, our feature vector comprised 5 features, out of which 4 features exhibited a p-value below 0.05, indicating their statistical significance. For the KNN classifier, the feature vector consisted of 15 features, with 5 features displaying a p-value below 0.05. In the RF classifier, the feature vector encompassed 12 features, among which 7 features possessed a p-value less than 0.05, as shown in Table 4. Table 4 provide insights into the selection of significant features for each classifier and demonstrate the importance of feature selection in improving the accuracy and interpretability of the classification models.

TABLE IV. FIVE-FOLD CROSS VALIDATION (CV) RESULTS OF FEATURES FROM FORWARD FEATURE SELECTION (FS2)

Classifier	Features	CV fold	CV score	Average CV score
KNN	LH1_$GT_{0,0.5}$_std HH2_$GT_{0,0.5}$_std LH1_$LBP_{1,8}$_energy* HH2_$GT_{2,0.5}$_std LH1_$LBP_{2,16}$_entropy* HH2_$LBP_{2,16}$_energy* HL1_$GT_{2,0.5}$_std HL1_$LBP_{2,16}$_energy* HL1_FDTA_HurstCoeff_2* HH1_$GT_{0,0.5}$_std HH1_$GT_{1,0.5}$_std HH1_$GT_{2,0.5}$_std HH1_$LBP_{1,8}$_energy* HH1_$LBP_{2,16}$_energy HH1_$LBP_{2,16}$_entropy	1	0.68	0.78
		2	0.89	
		3	0.66	
		4	0.88	
		5	0.77	
SVM	HH1_FDTA_HurstCoeff_1 LH1_FDTA_HurstCoeff_1* HL1_$LBP_{2,16}$_entropy* LH2_$LBP_{2,16}$_entropy* HH1_$LBP_{2,16}$_entropy	1	0.68	0.72
		2	0.73	
		3	0.77	
		4	0.77	
		5	0.61	
RF	LH1_$GT_{0,0.5}$_mean HH1_FDTA_HurstCoeff_1* LH1_FDTA_HurstCoeff_2* HH2_$LBP_{1,8}$_entropy HL1_$GT_{2,0.5}$_mean HL1_$LBP_{1,8}$_energy* HL1_$LBP_{1,8}$_entropy* LH2_$LBP_{2,16}$_entropy* HH1_$GT_{1,0.5}$_std HL2_$GT_{1,0.5}$_std HH1_$LBP_{1,8}$_energy* HH1_$LBP_{1,8}$_entropy*	1	0.78	0.78
		2	0.78	
		3	0.88	
		4	0.77	
		5	0.66	

*p-value < 0.05,
XXi_GTn,m_feature- denotes the feature measured from Gabor transform (with threshold n and frequency m) of ith level XX wavelet sub-image.
XXi_LBPn,m_feature- denotes the feature measured from LBP image (with radius n and neighbourhood m pixels) of ith level XX wavelet sub-image.

B. Classification

In this work, we employ feature extraction methods and feature selection methods to construct feature vectors. These feature vectors are then used to train three different classifiers. The feature extraction methods allowed us to extract significant features from the input data, capturing important patterns and characteristics. Subsequently, the feature selection methods further refine the feature vectors by selecting the most relevant features, enhancing the efficiency and interpretability of the classifiers. We train three classifiers, namely SVM, KNN and RF, using the constructed feature vectors. These classifiers are evaluated based on their accuracy, which measures the overall correctness of the predictions. Additionally, sensitivity and specificity, derived from the confusion matrix, provided insights into the classifiers' ability to correctly identify positive and negative instances. Fig. 4 shows the cross validation accuracy obtained by SVM, KNN and random forest classifiers on all feature sets. Similarly, Fig. 5 and 6 show the sensitivity and specificity of the classifiers on feature sets.

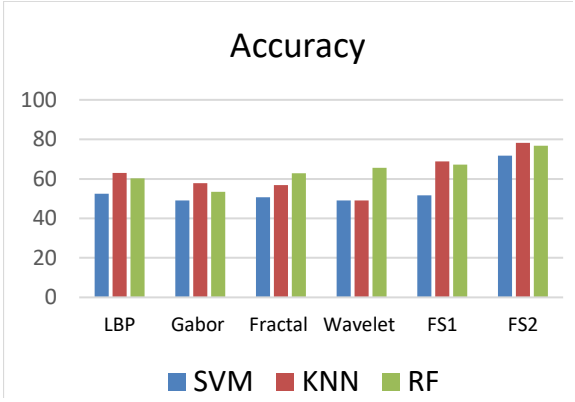

Figure 4. Five-fold cross validation accuracy of SVM, KNN and Random Forest classifiers trained using extracted texture feature sets.

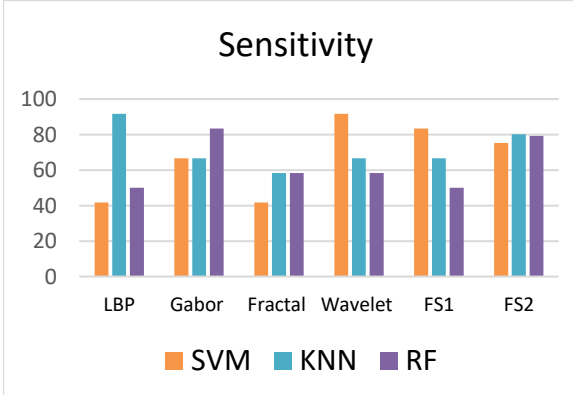

Figure 5. Sensitivity values of SVM, KNN and Random Forest classifiers trained using extracted texture feature sets.

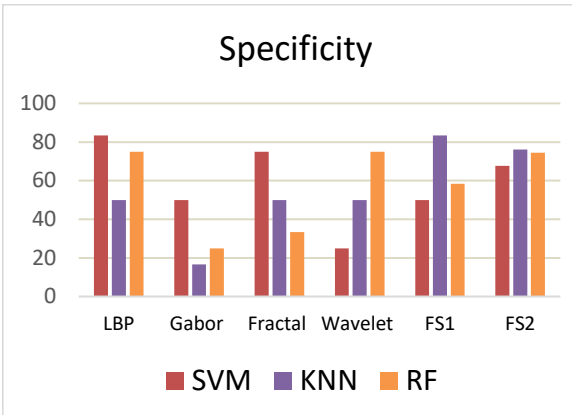

Figure 6. Specificity values of SVM, KNN and Random Forest classifiers trained using extracted texture feature sets.

Table 5 tabulates the accuracy, sensitivity and specificity values of KNN, SVM and random forest classifiers trained using different sets of trabecular texture features are tabulated. It is observed that in general, the best results are obtained when the classifiers are trained with texture features selected by the feature selection methods, such as independent sample t-test and forward feature selection method. The feature set FS2 outperforms other feature sets, indicating that forward feature selection is a better feature selection technique for identifying significant features.

Among the three classifiers, weighted KNN trained using FS2 features demonstrates the best performance. It achieves a 5-fold stratified cross-validation accuracy of 78.24%, with a sensitivity of 80.14% and specificity of 76.14%. This indicates that the classifier was able to accurately classify osteoporotic patients and healthy individuals, with a relatively high overall accuracy and balanced sensitivity and specificity.

Fig. 7 provides the Receiver Operating Characteristic (ROC) plots and the area under the ROC curve (AUC) of the 5-fold cross validation of KNN, RF and SVM classifiers trained on FS2 feature set. The ROC of every fold along with the mean curve is plotted. This further illustrates the superior performance of the weighted KNN classifier. Overall, this work showcases the effectiveness of feature extraction, feature selection, and classifier training in analyzing and classifying osteoporosis based on the constructed feature vectors.

TABLE V. RESULTS OBTAINED FROM THE CLASSIFIERS TRAINED WITH DIFFERENT SETS OF FEATURES

Features	Classifiers	Sensitivity (%)	Specificity (%)	5-fold CV accuracy (%)
Local Binary Pattern	SVM	41.67	83.34	52.42
	KNN	91.67	50.00	63.01
	RF	50.00	75.00	60.28
Gabor Transform	SVM	66.67	50.00	49.09
	KNN	66.67	16.64	57.82
	RF	83.34	25.00	53.40
Fractal Dimension	SVM	41.67	75.00	50.76
	KNN	58.36	50.00	56.88
	RF	58.36	33.34	62.86
Wavelet Transform (all)	SVM	91.67	25.00	49.09
	KNN	66.67	50.00	49.09
	RF	58.36	75.00	65.54
Wavelet Transform (FS1)	SVM	83.34	50.00	51.60
	KNN	66.67	83.34	68.91
	RF	50.00	58.36	67.24
Wavelet Transform (FS2)	SVM	75.24	67.70	71.75
	KNN	80.14	76.14	78.24
	RF	79.27	74.50	76.84

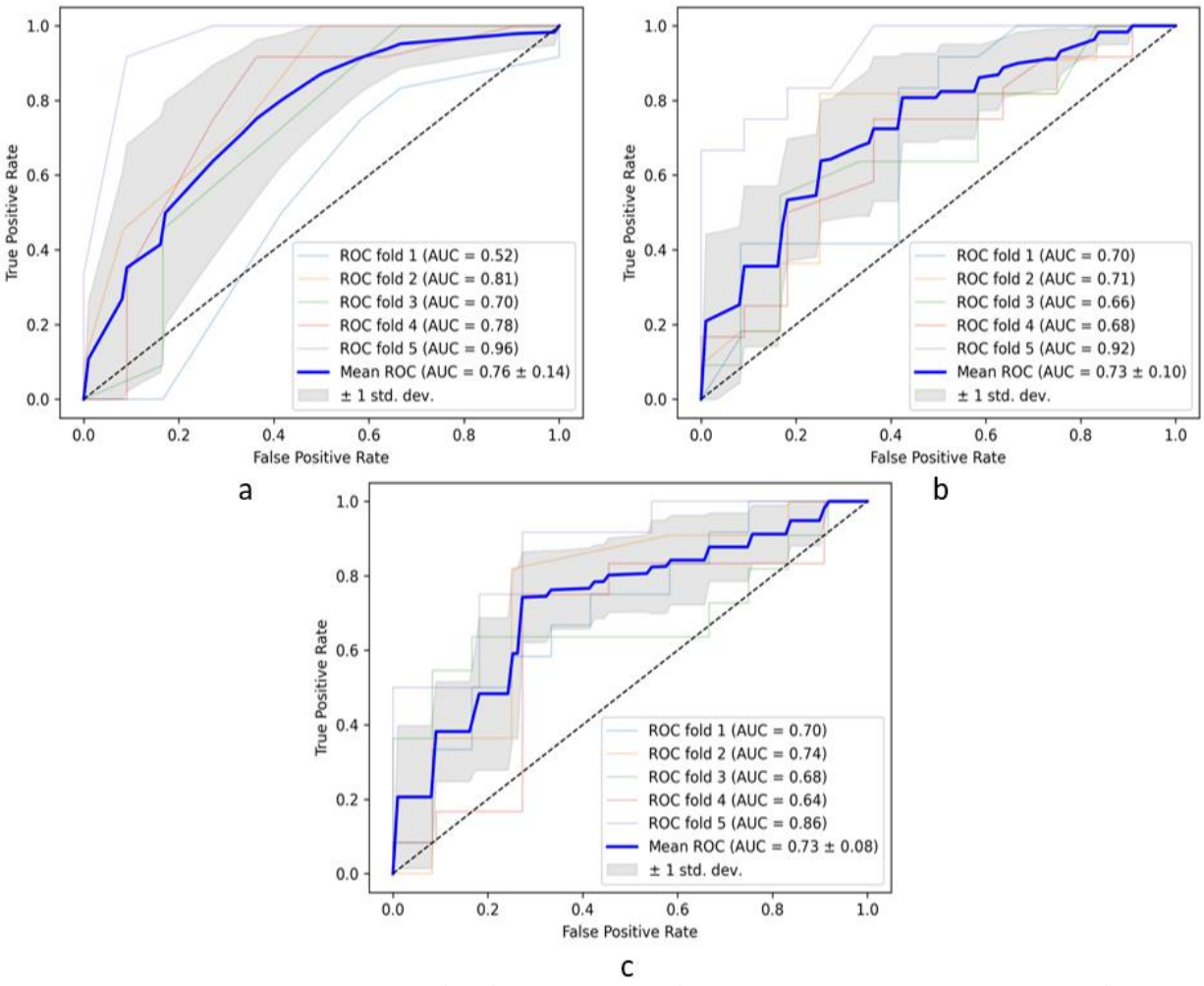

Figure 7. ROC plots and AUC values obtained for 5-fold cross validation of (a) KNN, (b) Random Forest, and (c) SVM classifiers

4. DISCUSSION

The outcomes of our proposed method were contrasted with those of existing studies that used the same dataset. Table 6 compares the results of the proposed method with related work using the same dataset. Reference [13] computed Holder exponents, then Hausdorff dimensions are determined. Finally, lacunarity is computed and obtained accuracy of 59% with 59% sensitivity and 59% specificity. Reference [9] used Gabor filter bank, created using filters with four scales and four orientations. Afterwards, 1D-LBP histograms are acquired from which an accuracy of 72.71% is obtained. Reference [10] utilized trabecular bone texture features which are extracted using the regularization and box-counting dimensions, and obtained accuracy of 55% with 59% sensitivity and 52% specificity.

Other studies, such as those conducted by [22] and [23] utilized the same dataset, consisting of 116 images. Their respective results showed an accuracy of 63.8% and 74.1%. Furthermore, References [8], [24] and [25] also utilized the same dataset in their studies. However, they had included an additional blind data of 58 images, bringing the total dataset size to 174 images. By including these blind data, they were able to expand the dataset and potentially obtain more comprehensive insights and improved results compared to our proposed work. The availability of a larger dataset have provided a broader perspective and potentially enhanced the accuracy of their findings.

In addition to the present work, several other research works, including those conducted by [26]-[29] have explored similar topics but utilized different datasets. Table 7 compares the proposed method with related work using different dataset. Notably, the studies that incorporated larger datasets have shown promising results. The availability of a larger data pool has likely provided these studies with a more comprehensive representation of the underlying patterns and improved the reliability and generalizability of their findings.

TABLE VI. COMPARISON WITH RELATED WORK USING CALCANEAL IMAGE DATA

Reference	Features used	Classifier	Dataset	Sensitivity (%)	Specificity (%)	Accuracy (%)
[22]	Wavelet Marginals-Haar	SVM	58 cases &58 control	62.1	65.5	63.8
[23]	FD, wavelet transform, DFT, DCT, Gabor, LBP, Laws masks, edge	Random forest	58 cases &58 control	74.10	74.10	74.1
[13]	Multifractal-based lacunarity analysis	SVM	58 cases &58 control	59.00	59.00	59
[10]	Regularization dimension and box-counting dimension	SVM	58 cases &58 control	59.00	52.00	55
[9]	Gabor filters and 1D local binary pattern (1D-LBP)	KNN	58 cases &58 control	-	-	72.71
[24]	Histogram, GLCM, PCA	SVM	87 cases &87 controls	97.70	95.40	96.60
[25]	Anisotropic discrete dual-tree wavelet transform	SVM	87 cases &87 controls	-	93.10	91.90
[8]	Oriental fractal analysis	-	87 cases &87 controls	72.00	71.00	71.80
[14]	Multi-fractal spectrum	LR	87cases &87 controls	97.26	98.78	98.01
[17]	Sparse analysis method	Deep learning	87cases &87 controls	-	-	72.67
Proposed method	**Features from Wavelet images using LBP, GT and FD**	**KNN**	**58 cases &58 control**	**80.14**	**78.24**	**78.24**

TABLE VII. COMPARISON WITH RELATED WORK USING DIFFERENT DATASET

Reference	Features used	Classifier	Dataset	Sensitivity (%)	Specificity (%)	Accuracy (%)
[29]	1D-LBP	KNN	39 cases & 41 controls	_	43.99	71.30
[26]	Fractional Brownian model and Rao geodesic distance	KNN	348 cases & 348 controls	97.80	95.40	96.60
[28]	Cortial and hLLBP	Logistic Regression	Distal radius 60 cases & 60 controls	81.70	76.70	79
[16]	Deep neural network architecture	Deep learning	186 images in both case and control	_	_	84.06
Proposed method	**Features extracted from Wavelet transform images using LBP, GT and FD**	**KNN**	**58 cases &58 control**	**80.14**	**78.24**	**78.24**

In recent years, several research studies, including works by [15]-[17] have successfully employed deep learning techniques. These studies have leveraged the power of deep learning models to improve the accuracy and effectiveness of classification tasks. However, in our research, we did not utilize deep learning techniques due to the consideration that a substantial amount of data is typically required to train and effectively optimize deep learning models. Given the constraints of our dataset size, we opted for a machine learning technique that could yield meaningful results within the available data limitations. By acknowledging the previous works that utilized deep learning techniques, we highlight the potential of such methods in the classification of healthy and osteoporotic images. However, our work explored alternative approaches that can still provide valuable insights despite the challenges posed by limited data availability.

5. CONCLUSION

The proposed method uses the texture feature analysis method to extract features from calcaneal radiographs. Our method was used to categorize 116 X-rays of bone tissue, half of which were of patients with osteoporosis and the other half were of healthy individuals. There were four steps to the proposed method. The radiograph preprocessing in the first step was done to emphasize the trabecular bone network. The second step included feature extraction and texturing feature analysis. In the third stage, feature selection techniques were used to obtain significant features. In the last stage, three different classifiers were trained with the obtained significant features and the trained model was used classify osteoporotic patients and healthy subjects. Based on the results obtained, we can draw the following conclusions. The features extracted from wavelet transform images using LBP, GT, and FD proved to be valuable in the analysis of bone trabecular network. This suggests that both the orientation and fractal features play a significant role in this analysis. Furthermore, the feature vector obtained through forward feature extraction demonstrated better performance compared to the feature vector obtained through independent sample t-test. This highlights the effectiveness of the forward feature extraction method in capturing relevant and discriminative features for the classification. Thus, the developed model can be used as an effective screening tool for Osteoporosis in clinical settings.

COMPLIANCE WITH ETHICAL STANDARDS

All compliance and ethical standards have been followed for data collection and experimental analysis.

FUNDING DETAILS

No funding is received for this research work.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

[1] S. K. Sabat, S. Panda, B. S. Sahoo, and P. Barik, "Prevalence of osteoporosis in India", International Journal of Health Sciences, pp. 10154–10166, 2022. https://doi.org/10.53730/ijhs.v6nS4.11033.

[2] T. Sözen, L. Özışık, and N. Ç. Başaran, "An overview and management of osteoporosis", European Journal of Rheumatology, vol. 4. no. 1, p.46, 2017. https://doi.org/10.5152/EURJRHEUM.2016.048.

[3] T. D. Rachner, S. Khosla, and L. C. Hofbauer, "Osteoporosis: now and the future", The Lancet, vol. 377, no.9773, pp. 1276–1287, 2011. https://doi.org/10.1016/S0140-6736(10)62349-5.

[4] N. Harvey, E. Dennison, and C. Cooper, "Osteoporosis: impact on health and economics", Nature Reviews Rheumatology, vol. 6, no. 2, pp. 99–105, 2010. https://doi.org/10.1038/nrrheum.2009.260.

[5] A. S. Areeckal, M. Kocher, and S. S. David, "Current and Emerging Diagnostic Imaging-Based Techniques for Assessment of Osteoporosis and Fracture Risk", IEEE Reviews in Biomedical Engineering, vol. 12, pp. 254-268, 2019. https://doi.org/10.1109/RBME.2018.2852620.

[6] A. Mithal, B. Bansal, S. C. Kyer, and P. Ebeling, "The Asia-Pacific Regional Audit-Epidemiology, Costs, and Burden of Osteoporosis in India 2013: A report of International Osteoporosis Foundation", Indian Journal of Endocrinology and Metabolism, vol. 18, no.4, pp. 449–454, 2014. https://doi.org/10.4103/2230-8210.137485.

[7] A. Rosholm, L. Hyldstrup, L. Bæksgaard, M. Grunkin, and H. H. Thodberg, "Estimation of bone mineral density by digital x-ray radiogrammetry: Theoretical background and clinical testing", Osteoporosis International, vol. 12, no. 11, pp. 961–969, 2001. https://doi.org/10.1007/S001980170026/METRICS.

[8] K. Harrar, R. Jennane, K. Zaouchi, T. Janvier, H. Toumi, and E. Lespessailles, "Oriented fractal analysis for improved bone microarchitecture characterization", Biomedical Signal Processing and Control, vol. 39, pp. 474–485, 2018. https://doi.org/10.1016/J.BSPC.2017.08.020.

[9] F. Riaz, R. Nemati, H. Ajmal, A. Hassan, E. Edifor, and R. Nawaz, "Osteoporosis classification using texture features", Proceedings - IEEE Symposium on Computer-Based Medical Systems, pp. 575–579, June 2019. https://doi.org/10.1109/CBMS.2019.00119.

[10] D. A. Palanivel, S. Natarajan, S. Gopalakrishnan, and R. Jennane, "Trabecular Bone Texture Characterization Using Regularization Dimension and Box-counting Dimension", IEEE Region 10 Annual International Conference, Proceedings/TENCON, pp. 1047–1052, October 2019. https://doi.org/10.1109/TENCON.2019.8929524.

[11] S. Zehani, A. Ouahabi, M. Oussalah, M. Mimi, and A. Taleb-Ahmed, "Bone microarchitecture characterization based on fractal analysis in spatial frequency domain imaging", International Journal of Imaging Systems and Technology, vol. 31, no. 1, pp. 141–159, 2021. https://doi.org/10.1002/IMA.22512.

E-mail: anu.areeckal@manipal.edu

[12] K. Zheng, C. E. Harris, R. Jennane, and S. Makrogiannis, "Integrative blockwise sparse analysis for tissue characterization and classification", Artificial Intelligence in Medicine, vol. 107, p. 101885, 2020. https://doi.org/10.1016/J.ARTMED.2020.101885.

[13] D. A. Palanivel, S. Natarajan, S. Gopalakrishnan, and R. Jennane, "Multifractal-based lacunarity analysis of trabecular bone in radiography", Computers in Biology and Medicine, vol. 116, p. 103559, 2020. https://doi.org/10.1016/J.COMPBIOMED.2019.103559.

[14] O. Bouzeboudja, B. Haddad, A. Taleb-Ahmed, S. Ameur, M. El Hassouni, and R. Jennane, "Multifractal analysis for improved osteoporosis classification", Biomedical Signal Processing and Control, vol. 80, p. 104225, 2023. https://doi.org/10.1016/J.BSPC.2022.104225.

[15] T. Meena and S. Roy, "Bone Fracture Detection Using Deep Supervised Learning from Radiological Images: A Paradigm Shift", Diagnostics (Basel, Switzerland), vol. 12, no. 10, 2022. https://doi.org/10.3390/DIAGNOSTICS12102420.

[16] A. Kumar, R. C. Joshi, M. K. Dutta, R. Burget, and V. Myska, "Osteo-Net: A Robust Deep Learning-Based Diagnosis of Osteoporosis Using X-ray images", 2022 45th International Conference on Telecommunications and Signal Processing, TSP 2022, pp. 91–95, 2022. https://doi.org/10.1109/TSP55681.2022.9851342.

[17] C. E. Harris, and S. Makrogiannis, "Sparse Analysis of Block-Boosted Deep Features for Osteoporosis Classification", 2022 IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop, 2022. https://doi.org/10.1109/IVMSP54334.2022.9816199.

[18] A. Hadid, J. Ylioinas, M. Bengherabi, M. Ghahramani, and A. Taleb-Ahmed, "Gender and texture classification: A comparative analysis using 13 variants of local binary patterns", Pattern Recognition Letters, vol. 68, pp. 231–238, 2015. https://doi.org/10.1016/j.patrec.2015.04.017.

[19] L. Pothuaud, C. L. Benhamou, P. Porion, E. Lespessailles, R. Harba, and P. Levitz, "Fractal Dimension of Trabecular Bone Projection Texture Is Related to Three-Dimensional Microarchitecture", Journal of Bone and Mineral Research, vol. 15, no. 4, pp. 691–699, 2010. https://doi.org/10.1359/jbmr.2000.15.4.691.

[20] O. Regniers, L. Bombrun, V. Lafon, and C. Germain, "Supervised Classification of Very High Resolution Optical Images Using Wavelet-Based Textural Features", IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 6, pp. 3722–3735, 2016. https://doi.org/10.1109/TGRS.2016.2526078.

[21] A. Jović, K. Brkić, and N. Bogunović, "A review of feature selection methods with applications", 2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2015 - Proceedings, pp. 1200–1205, 2015. https://doi.org/10.1109/MIPRO.2015.7160458.

[22] F. Yger, "Challenge IEEE-ISBI/TCB : Application of Covariance matrices and wavelet marginals", 2014. https://arxiv.org/abs/1410.2663v1.

[23] K. Zheng and S. Makrogiannis, "Bone texture characterization for osteoporosis diagnosis using digital radiography", Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 1034–1037, 2016. https://doi.org/10.1109/EMBC.2016.7590879.

[24] A. Singh, K. Dutta, R. Jennane, E. Lespessailles, and M. Kishore Dutta, "Classification of the Trabecular Bone Structure of Osteoporotic Patients using Machine Vision", 2017. https://doi.org/10.1016/j.compbiomed.2017.10.011.

[25] H. Oulhaj, M. Rziza, A. Amine, H. Toumi, E. Lespessailles, M. El Hassouni, and R. Jennane, "Anisotropic Discrete Dual-Tree Wavelet Transform for Improved Classification of Trabecular Bone", IEEE Transactions on Medical Imaging, vol. 36, no. 10, pp. 2077–2086, 2017. https://doi.org/10.1109/TMI.2017.2708988.

[26] M. El Hassouni, A. Tafraouti, H. Toumi, E. Lespessailles, and R. Jennane, "Fractional Brownian Motion and Rao Geodesic Distance for Bone X-Ray Image Characterization", IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 5, pp. 1347–1359, 2017. https://doi.org/10.1109/JBHI.2016.2619420.

[27] A. S. Areeckal, N. Jayasheelan, J. Kamath, M. Kocher, and S. S. David, "Early diagnosis of osteoporosis using radiogrammetry and texture analysis from hand and wrist radiographs in Indian population", Osteoporosis International, vol. 29, pp. 665–673, 2018. https://doi.org/10.1007/s00198-017-4328-1.

[28] A. S. Areeckal, J. Kamath, S. Zawadynski, M. Kocher, and S. S. David, "Combined radiogrammetry and texture analysis for early diagnosis of osteoporosis using Indian and Swiss data", Computerized Medical Imaging and Graphics, vol. 68, pp. 25–39, 2018. https://doi.org/10.1016/j.compmedimag.2018.05.003.

[29] L. Houam, A. Hafiane, A. Boukrouche, E. Lespessailles, and R. Jennane, "One dimensional local binary pattern for bone texture characterization", Pattern Analysis and Applications, vol. 17, no.1, pp. 179–193, 2014. https://doi.org/10.1007/s10044-012-0288-4..

Sai Kiran K S received the Bachelor of Science in Biotechnology from Mysore University, India, in 2021, and the Master of Science in Bioinformatics from Manipal School of Life Sciences (MSLS), Manipal Academy of Higher Education, Manipal, India, in 2023. Since August 2023, he has been working as Associate clinical data programmer at Novo Nordisk, India.

Anu Shaju Areeckal received the B.Tech. degree in electronics and communication engineering from Mahatma Gandhi University, Kerala, India, in 2010, the M.Tech. degree in VLSI design from the Karunya Institute of Technology and Sciences, Tamil Nadu, India in 2012, and the Ph.D. degree from the National Institute of Technology Karnataka, Surathkal, India in 2019. Since 2019, she has been an Assistant Professor with the Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India. Her research interests include biomedical image analysis, computer vision, machine and deep learning. She received various awards for the thesis on ‘‘Early Diagnosis of Osteoporosis Using Metacarpal Radiogrammetry and Texture Analysis’’, which includes Gandhian Young Technological Innovation (GYTI) Award 2019, India, Fourth Best PhD Thesis Award in IEEE ComSoc Graduate Congress GRATE’7’ 2020, India, and BITES (Board for IT Education Standards) Best PhD thesis 2019- Special Mention of the Jury Award, Karnataka, India.