

Article

Qualitative Classification of Proximal Femoral Bone Using Geometric Features and Texture Analysis in Collected MRI Images for Bone Density Evaluation

Mojtaba Najafi ¹, Tohid Yousefi Rezaii ¹, Sebelan Danishvar ^{2,*}  and Seyed Naser Razavi ¹

¹ Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz 51666-16471, Iran; mojtaba.najafi@tabrizu.ac.ir (M.N.); yousefit@tabrizu.ac.ir (T.Y.R.); razaviii@tabrizu.ac.ir (S.N.R.)

² College of Engineering, Design and Physical Sciences, Brunel University London, Uxbridge UB8 3PH, UK

* Correspondence: sebelan.danishvar@brunel.ac.uk

Abstract: The aim of this study was to use geometric features and texture analysis to discriminate between healthy and unhealthy femurs and to identify the most influential features. We scanned proximal femoral bone (PFB) of 284 Iranian cases (21 to 83 years old) using different dual-energy X-ray absorptiometry (DEXA) scanners and magnetic resonance imaging (MRI) machines. Subjects were labeled as “healthy” (T-score > −0.9) and “unhealthy” based on the results of DEXA scans. Based on the geometry and texture of the PFB in MRI, 204 features were retrieved. We used support vector machine (SVM) with different kernels, decision tree, and logistic regression algorithms as classifiers and the Genetic algorithm (GA) to select the best set of features and to maximize accuracy. There were 185 participants classified as healthy and 99 as unhealthy. The SVM with radial basis function kernels had the best performance (89.08%) and the most influential features were geometrical ones. Even though our findings show the high performance of this model, further investigation with more subjects is suggested. To our knowledge, this is the first study that investigates qualitative classification of PFBs based on MRI with reference to DEXA scans using machine learning methods and the GA.



Citation: Najafi, M.; Yousefi Rezaii, T.; Danishvar, S.; Razavi, S.N. Qualitative Classification of Proximal Femoral Bone Using Geometric Features and Texture Analysis in Collected MRI Images for Bone Density Evaluation. *Sensors* **2023**, *23*, 7612. <https://doi.org/10.3390/s23177612>

Academic Editor: Massimo Mischi

Received: 2 July 2023

Revised: 15 August 2023

Accepted: 29 August 2023

Published: 2 September 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: osteoporosis; magnetic resonance imaging; dual energy X-ray absorptiometry; machine learning

1. Introduction

Osteoporosis is a systemic skeletal disease in which bone fragility and risk of fracture increase with loss of bone density and microarchitectural deterioration of bone tissue. Many parameters such as age, race, gender, hormonal disorders, heredity, and poor dietary habits correlate with the progression of the disease [1]. Osteoporosis can be considered to be a result of a lack of calcium as well as other minerals in the bones. This disease can be classified into the following three categories: primary osteoporosis including idiopathic osteoporosis, type 1 osteoporosis due to the estrogen deficiency common in postmenopausal women, and type 2 osteoporosis due to an estrogen deficiency with a loss of bone density. Vertebral fractures can be considered to be one of the common consequences of osteoporosis, which leads to constant pain, skeletal–muscular abnormalities, respiratory abnormalities, and deterioration of physical performance and quality of life. Also, there will be a chance of sustaining a fracture after an initial fracture. Physiotherapy can be considered a special non-pharmacological method that can be useful and effective for patients who have osteoporosis, through manual methods, resistance exercises, etc. [2,3].

Osteoporosis is common, preventable, and treatable [4,5]. However, it is a silent disease as it may be asymptomatic for many years until the fracture occurs, which can be physically debilitating and consequently cause a financial burden on the patient [6,7]. Without prevention and screening, the osteoporosis-related morbidity and mortality will

greatly burden healthcare systems throughout the world [8,9]. Thus, it is key to diagnose and treat osteoporosis before a fracture occurs.

Bone mineral density (BMD) is currently the main factor for determining bone quality and risk of osteoporosis [10]. Dual-energy X-ray absorptiometry (DEXA) is considered the clinical gold standard for BMD evaluation. According to the guidelines of the National Osteoporosis Foundation, the indications for DEXA examination are: (a) women 65 years old and men 70 years and older, (b) postmenopausal women and men over 50 who have suffered bone fractures, and (c) adults who suffer from rheumatoid arthritis or use medications such as glucocorticoids and have low bone mass [2,3,11]. However, DEXA scans do not give any information regarding the bone microarchitecture and mechanical properties, which are also crucial factors in bone resistance to fracture [12,13]. To obtain these factors, other non-invasive methods such as conventional radiology, computed tomography, magnetic resonance imaging (MRI), and ultrasound have been investigated [14,15].

MRI is a common imaging technique used in imaging centers. The feasibility of structural analysis of proximal femoral bone (PFB), one of the most clinically important sites of fracture, using MRI has been established [16,17]. Use of quantitative MRI techniques in assessing osteoporosis and fracture risk in the PFB has been documented [18,19]. Image segmentation and quantification to analyze osteoporosis and bone fracture risk using these methods are time-consuming and impractical [20]. The use of machine learning techniques (deep learning) such as convolutional networks [21] and deep convolutional neural networks [22] has shown promising results in solving segmentation problems.

If we can determine the bone quality from MRI, when patients are referred to imaging centers to check the health of body organs such as the bladder, intestine, liver, kidneys, ovaries, and breasts in women and the prostate in men, we can detect many patients with osteoporosis or who are on the verge of developing osteoporosis earlier and use appropriate treatments to reduce the number of patients in society. However, different acquisition protocols influence the results of segmentation [23,24] and texture analysis [25].

The aim of our study is to investigate whether the discrimination of unhealthy femurs using images with different scan parameters (and from different hospitals) is practical or not. Another goal in our study is to identify important features in the classification of healthy and unhealthy femurs.

To our knowledge, this is the first study that uses machine learning and a genetic selection algorithm to evaluate proximal femoral bone quality based on MRI with reference to the BMDs from DEXA scans. We used machine learning methods to qualitatively classify proximal femoral bones into two classes (healthy and unhealthy bones (bones with osteopenia or osteoporosis)) using geometric features and texture analysis in MRI. Finally, we verified our classification results against a classification based on the results of DEXA scans as our reference.

The main contributions of this article are as follows:

- The compilation of a comprehensive MRI database available for the diagnosis of osteoporosis;
- The provision of a smart algorithm based on machine learning techniques to evaluate bone density;
- Achievement of the highest detection speed for bone density assessment;
- A comparison of the proposed method with other deep learning methods.

2. Methods

In this section, the collection of MRI images, pre-processing operations, data segmentation, classifications, and features used are fully described.

2.1. Data Collection

Our collected data included 284 images taken from 21 to 83-year-old Iranians with no history of bone fracture (135 men, average age of 56, standard deviation of 13, and with an average BMI of 27.8 and 149 women, average age of 49.5, standard deviation of

16, and with an average BMI of 24.3). The data were collected from the Ronash imaging center in Dezfoul city, Allameh Bohlul hospital in Gonabad city, and Noh-e-Dey hospital in Torbat-e-Heydarieh city in Iran. We obtained the informed consent of all subjects in writing prior to the tests and the tests were approved by the Ethics Committee of the hospital (with license number IRB.1398.9.26). In general, the following patients were excluded from the testing process: patients with a lack of personal consent for testing, people who used a pacemaker, and people who had an implant or a metal foreign body in their body (such as having a piece of metal in the brain or eye or platinum in bone).

Subjects were referred by a physician to imaging centers to be scanned from the pelvic region (and in most cases other regions as well). Scan parameters were set by specialists according to their experience. All three centers used 1.5T MRI machines (two had GE Optima 450 MRI machines and the other had a Siemens MAGENTOM Avanto MRI machine). These scan parameters are presented below:

- T1/TR = 415.00 ms/TE = 19.00 ms/Flip-Angle = 150 (44 referred, 33 healthy, 11 unhealthy)
- T1/TR = 536.00 ms/TE = 11.00 ms/Flip-Angle = 180 (43 referred, 31 healthy, 12 unhealthy)
- T1/TR = 4070.00 ms/TE = 33.00 ms/Flip-Angle = 180 (42 referred, 29 healthy, 13 unhealthy)
- T1/TR = 420.00 ms/TE = 22.00 ms/Flip-Angle = 180 (45 referred, 36 healthy, 9 unhealthy)
- T2/TR = 3600.00 ms/TE = 80.00 ms/Flip-Angle = 150 (45 referred, 19 healthy, 26 unhealthy)
- T2/TR = 7840.00 ms/TE = 109.00 ms/Flip-Angle = 150 (41 referred, 28 healthy, 13 unhealthy)
- T1/TR = 389.00 ms/TE = 13.42 ms/Flip-Angle = 110 (24 referred, 9 healthy, 15 unhealthy)

The BMDs of the subjects were obtained using three DEXA scanners (NORLAND bone densitometer). The scan parameters were set by the scanner for the femur automatically and not by the operator.

The person who recorded the MRI images had an MRI fellowship (specialized in radiology) and the person who recorded the DEXA scans was a radiologist. The time it took to obtain MRI images of the femur was approximately 15 min, and the time it took to obtain the BMD of the femur using the DEXA technique was approximately 2 min.

2.2. Programming Environment and Settings

Reading of the MRI images (saved in DICOM format), their conversion into usable numeric arrays, image segmentation, and feature extraction were implemented in Python and in the Anaconda-Spider environment. We used MATLAB R2021b to implement classification and feature selection processes.

We have made our program available to the public on <https://github.com> (file name: Mojtab2023/Classification-of-Proximal-Femoral-Bone-Using-Geometric-Features-and-Texture-Analysis-in-MR-Images-f (1 January 2022)).

2.3. Overall Procedure

Figure 1 shows the flowchart of the procedure we used to achieve classification of the PFB. We used a 10-fold cross-validation technique to carry out our experiment. To this end, the dataset was divided into 10 non-overlapping sections. The remaining section was used for testing after the nine sections were used for training. The process was repeated 10 times until all sections were taken into account in the testing procedure.

The PFBs were divided into the two groups of unhealthy (with the risk of osteopenia or osteoporosis; T score ≤ -0.9) and healthy (T score > -0.9) based on the result of the DEXA scans and labeled 1 and 0, respectively.

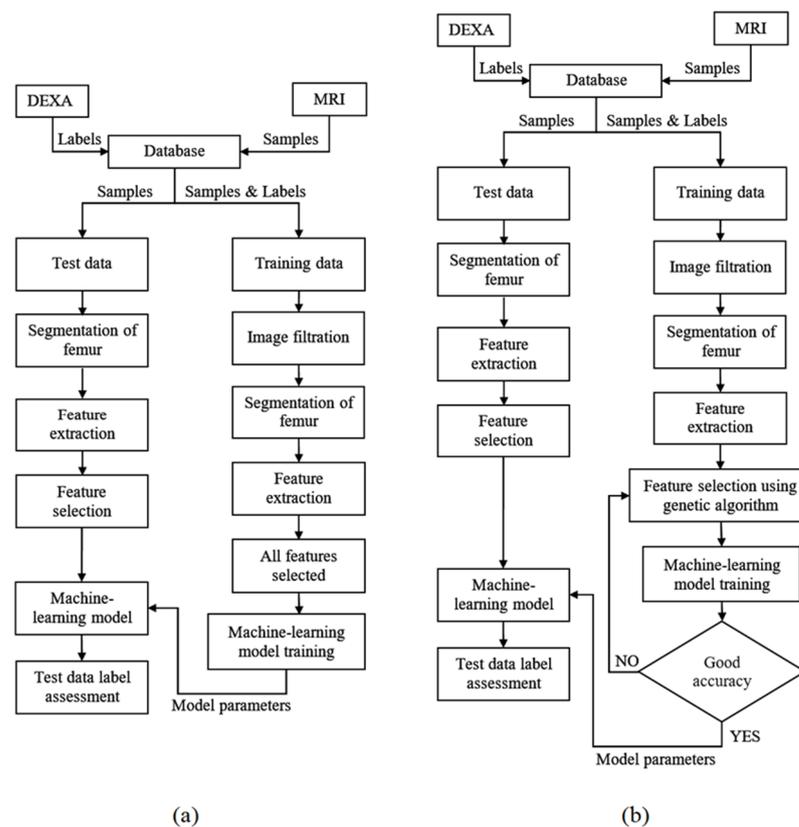


Figure 1. Overall procedure flowcharts. The proposed method for the diagnosis of patients at risk of osteoporosis using MRI and DEXA scans of the femur using different classifiers (a) and feature selection for classifiers using the Genetic algorithm (b).

Because the number of images associated with the healthy label exceeded the number of images associated with the unhealthy label, which may cause the majority class to be considered in the classification, the number of minority classes was increased. We artificially increased the data of the minority class using a data augmentation technique. In this way, the rotation range of the images was considered to be 40 degrees. Also, the horizontal shift of the images and the vertical shift of the images was considered to be 0.2%.

2.4. Segmentation

We used a 5×5 pixel median filter to reduce salt and pepper noise in the MRI. Then, segmentation was performed using the region growing method or the active contour method.

In the region growing method, we used the Canny edge detection technique. The starting point for the edge detection algorithm is important and affects the result. In order to find the best starting point, we used 2 different methods. One method was using the geometric features of PFB and the Hough transform to obtain the center of the Hough circle as the starting point and the other was choosing a pixel inside the bone with the highest intensity (where the image's histogram in the horizontal and vertical directions was the maximum). We evaluated the two methods used based on two indices (the Dice Similarity Coefficient (DSC) [26] and the Jacard Index (also called Intersection over Union, IoU)) [27] to determine the efficiency of each method. These two methods can be defined as Equations (1) and (2), where S and T , correspondingly, represent the segmentation result and the ground truth. Figure 2 shows a sample of the segmented results by both techniques. The best result from these two methods was chosen for feature extraction. In thirteen cases, segmentation failed by both methods and the boundary of the femur was determined

manually. Four out of these thirteen images had scan parameters of (T1/TR = 389.00 ms/TE = 13.42 ms/Flip-Angle = 110).

$$DSC = \frac{2|S \cap T|}{|S| + |T|} \quad (1)$$

$$IoU = \frac{|S \cap T|}{|S \cup T|} \quad (2)$$

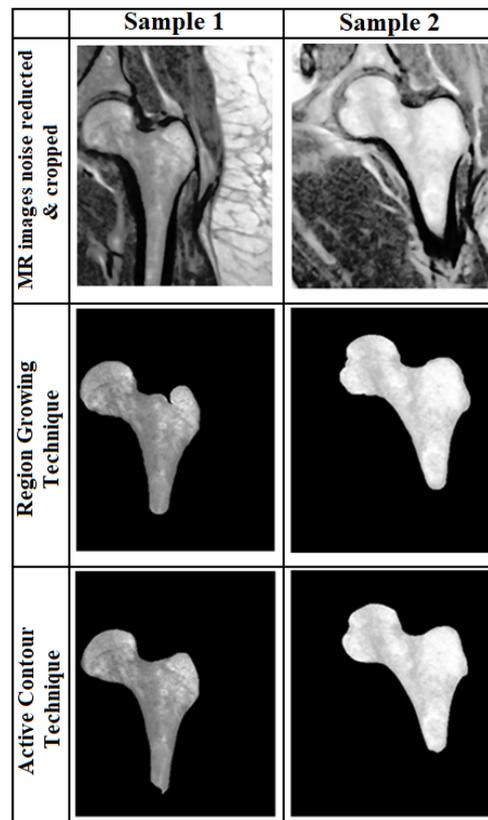


Figure 2. Segmentation results. Two segmentation results for two samples based on the active contour (Bottom row) and region growing (middle row) techniques.

After segmentation, features were extracted. Figure 3 shows the basic geometrical features that were used in this study. The parameters $W1$, $W2$, etc. are defined as geometric features in Figure 3. The purpose of this definition of these geometric parameters that are marked on the figure was to find out whether there is a relationship between the defined geometric parameters with osteoporosis. To put it more simply, parameters $W1$, $W2$, etc. were defined by ourselves as geometric features and are not shorthand for a specific phrase or word.

We created an averaged image from unhealthy (based on DEXA results) PFBs and another averaged image from healthy PFBs in our training dataset after flipping (for right femurs), cropping, and framing PFB images at similar positions. For images with the same scan parameters, the difference between these two images indicated the important regions in which these two classes differ the most, i.e., the boundaries of the PFB. This means that characteristics of the boundaries may represent the difference between the two PFB classes. When we used all images with different scan parameters, the inside region also became more apparent in the subtracted image (see Figure 4).

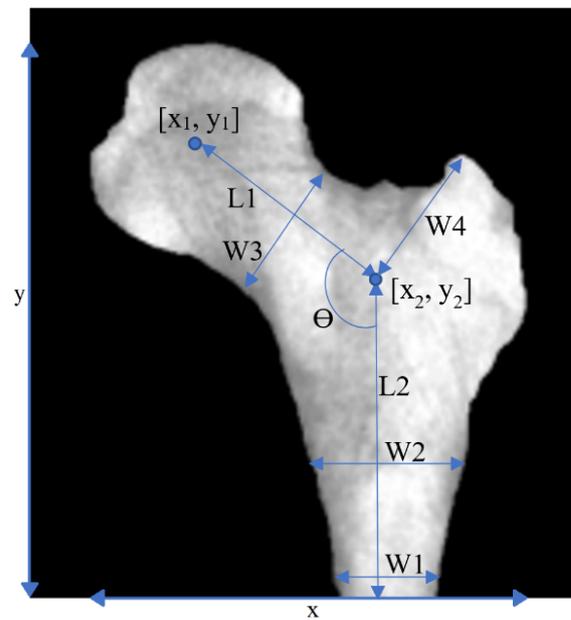


Figure 3. Geometrical features. Geometrical features extracted from a PFB image.

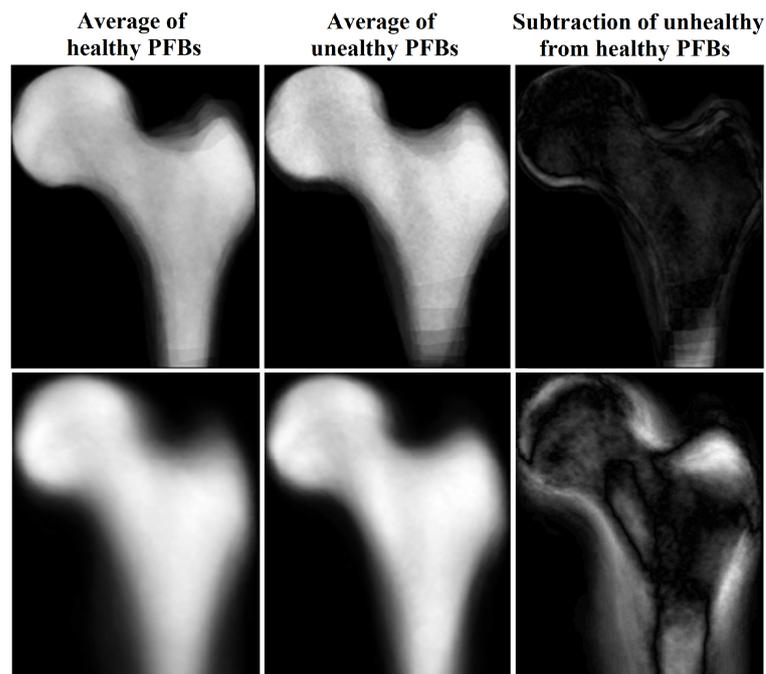


Figure 4. Average and subtraction of PFB images.

Five descriptive curves were also extracted from the cropped image of the bone. These curves were named C1-0, C1-1, C1-2, C2-0, and C4-0 (see Figure 5).

Each curve was obtained by measuring the distance of the bone's boundaries from the edge of the image (in the direction indicated by the arrow) along the marked regions. These curves describe the characteristics of the bone boundaries. Figure 6 shows two of the rustled curves for a PFB as an example.

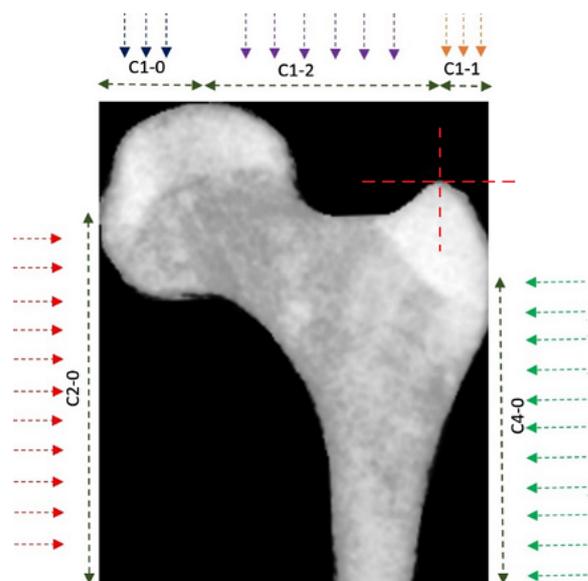


Figure 5. Feature curves for boundaries. Five feature curves were used to define the characteristics of the PFB’s boundaries.

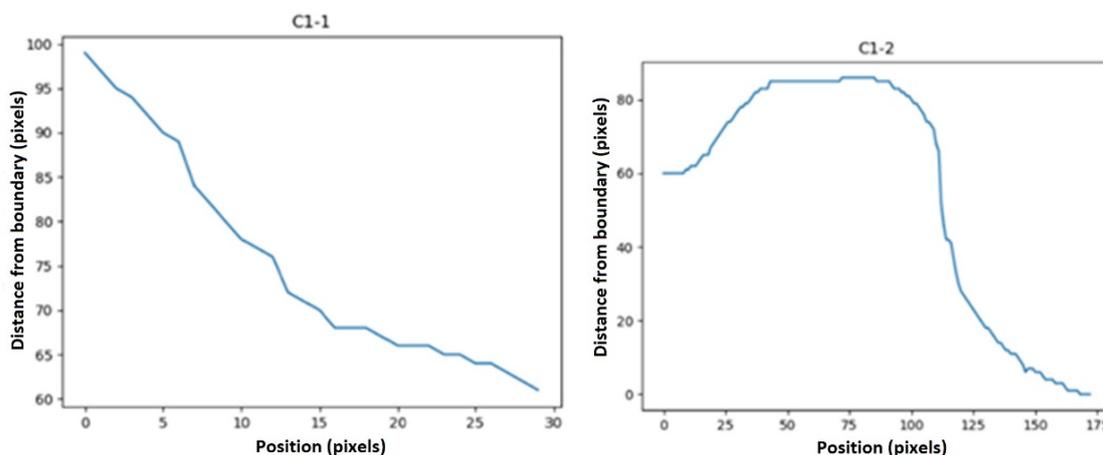


Figure 6. Sample of the resulting feature curves in Figure 5. Two samples of the resulting feature curves from the PFB in Figure 5.

Then, statistical functions such as the minimum, maximum, mean, variance, elongation, skewness, and first and second moments of the signal (moment1 and moment2 in Table 1) of these descriptors were calculated. We defined some features as below:

$$\text{Average ramp} = \Delta y / \Delta x \tag{3}$$

$$\text{First to last point ramp} = \Delta y_1 / \Delta x_1 \tag{4}$$

Figure 7 shows how these deltas were defined. Furthermore, we defined another feature to remove the effect of slope from our curves (to have more normal-distribution-like characteristics) to make functions such as skewness and kurtosis meaningful. This feature was defined as below:

$$\text{Ramp removed sig} = \text{signal} - \text{averageramp} \tag{5}$$

Table 1. Primary features. Primary extracted features used in this study (the second and third power of each feature was used as a feature as well).

Features	Description	No. of Features
Geometrical features	As in Figure 2, i.e.,: x1, y1, L1, L2, θ , W0, W1, W2, W3, W4, x, y	12
C1-1	min, max, average ramp, kurtosis (ramp_removed_sig), skew (ramp_removed_sig), first to last point ramp, moment1 (ramped_remov_sig), moment 2 (ramp_removed_sig), variance, max. of second derivative	10
C1-2	min, max, mean, variance, kurtosis, skew, moment1, moment2	8
C1-0	min, max, average ramp, kurtosis (ramp_removed_sig), skew (ramp_removed_sig), first to last point ramp, moment1 (ramped_remov_sig), moment 2 (ramp_removed_sig), variance, max of second derivative	10
C2-0	min, max, average ramp, kurtosis (ramp_removed_sig), skew (ramp_removed_sig), first to last point ramp, moment 1 (ramped_remov_sig), moment2 (ramp_removed_sig), variance, max of second derivative	10
C4-0	min, max, average ramp, kurtosis (ramp_removed_sig), skew (ramp_removed_sig), first to last point ramp, moment1 (ramped_remov_sig), moment2 (ramp_removed_sig), variance, max of second derivative	10
Features based on texture	mean, variance, skew	3
Fractal dimensions	fractal dimensions	1
Features based on texture	gray-level co-occurrence matrix (GLCM)	4

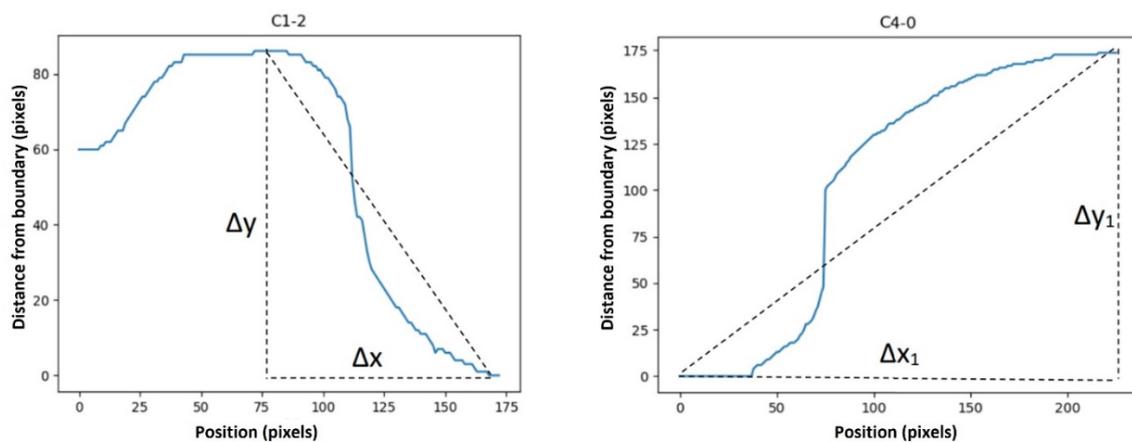


Figure 7. Δx , Δy , Δx_1 , and Δy_1 in Equations (3) and (4). Δx , Δy , Δx_1 , and Δy_1 were defined as shown in the images above to calculate the “average ramp” and “first to last point ramp”.

Features such as the pixel gray-scale average (as an indicator of bone marrow fat and the BMD) [28], the fractal dimension (using the Boxing count method), and the gray-level co-occurrence matrix were also derived from texture analysis (for further information on the image segmentation and texture analysis refer to the work of Larroza et al. [29]). A total of 68 primary features were extracted. We created a feature vector (FV) with these features as well as their 2nd and 3rd powers, i.e., a total of 204 features. Next, we used decision tree, logistic regression, and support vector machine (SVM) algorithms to classify the data.

2.5. Classification Algorithms

In this section, the classifications used and compared in this study are reviewed.

2.5.1. Classification Using Decision Tree

Decision tree (DT) is a versatile nonlinear machine learning model composed of hierarchical decision rules established during the training procedure to perform a classification task by recursively splitting independent factors into different groups. In [30–32], DT was used to perform disease diagnoses. Some studies also have employed and compared DT with other machine learning algorithms [33]. Here, we utilized DT as well as other classifiers to evaluate the extracted features.

2.5.2. Classification Using Logistic Regression

Logistic regression is a simple linear supervised binary classifier. Some studies have used logistic regression for medical purposes [34–36]. This classifier was employed in our experiments as a benchmark because, in recent years, it has been widely used in studies related to the classification of different types of osteoporosis.

2.5.3. Classification Based on the SVM Algorithm

SVM is one of the most-used algorithms for texture analyses of MRI [37]. It is computationally cheaper than a neural network and can produce similar or even better results [38].

Polynomial, linear, radial basis function (RBF), and Gaussian kernels were used in the implementation of the SVM algorithm in our study.

2.6. Feature Selection Using the Genetic Algorithm

Feature selection is necessary to find the most relevant features to improve the classification performance due to the variable discrimination power of the extracted features. The classification performance is hampered by redundant or irrelevant features; when more features are added to a model, the classification performance suffers [39].

The Genetic algorithm (GA) has been successfully used for feature selection in MRI segmentation and classification and also in combination with the SVM [40,41].

In the GA, the total population was set to 100, with a maximum of 500 iterations. The flowchart of this procedure is shown in Figure 1b. In our study, we used a cost function that would result in the best accuracy with the least number of features (NOF) as defined below:

$$\text{Cost (FV)} = (\text{Accuracy} + 1/\text{NOF})^{-1} \quad (6)$$

3. Results

Based on the results of their DEXA scans and T scores ≥ -0.9 , 185 subjects were labeled as “healthy”, while 99 subjects were labeled as “unhealthy”.

As the first part of our study, we employed SVM with different kernels (SVM-linear, SVM-Polynomial, SVM-RBF, and SVM-Gaussian), decision tree, and logistic regression classifiers to classify healthy and unhealthy cases using all 204 features. All the experiments were performed using the 10-fold cross-validation technique. Given the inherently random characteristic of the classifier training procedure and the fact that data partitioning impacts the final result, we repeated the experiments ten times. The results are shown in Figure 8.

As we can see, with all 204 features, the highest accuracies were obtained from the SVM with RBF and Gaussian kernels (more than 80%) and the lowest accuracy was from the logistic regression (above 60%).

Figure 9 shows how the classification performance of different algorithms improved in 500 iterations of the GA. When we selected features based on the GA, the accuracy of the SVM-RBF increased from about 83% to 89% using 56 features, while the accuracy of SVM-Gaussian, using only 7 features, increased from about 81% to 86%. Reducing the number of features improved the classification performance and increased the computational speed as was expected based on the literature [40,41].

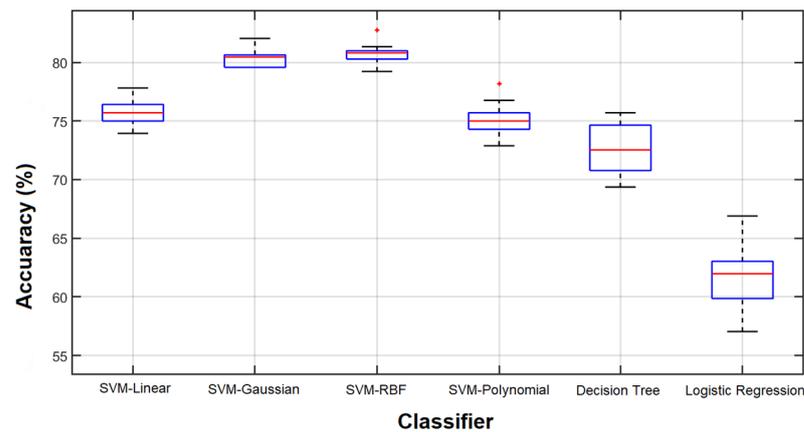


Figure 8. Classification results of 10 repetitions. Results obtained from 10 classification repetitions using various classifiers and all 204 features.

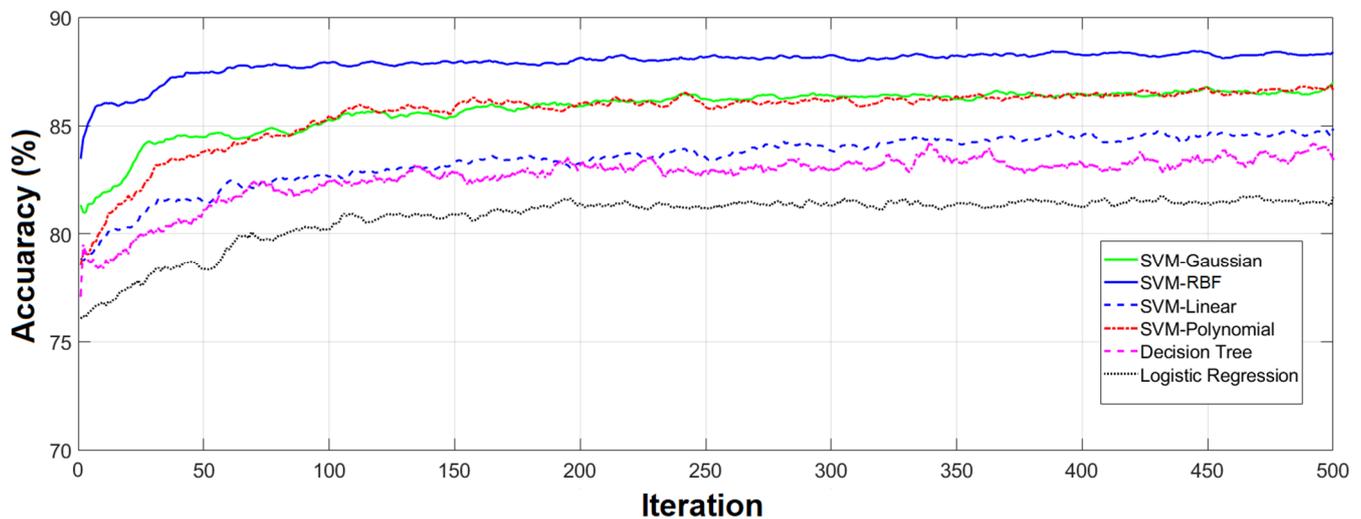


Figure 9. Classification results of 10 repetitions. Results obtained from repeating the classification 10 times using various classifiers and all 204 features.

Table 2 shows the best results (using a confusion matrix) achieved based on the features selected by the GA for each classifier and its corresponding number of selected features. According to Table 2, as can be seen, the GA has selected different features and the number of features for different categories. For example, 60 features, 56 features, and 7 features are selected for SVM-Linear, SVM-RBF, and SVM-Gaussian, respectively. The types of features selected by the GA are discussed below.

Table 2. Final classification results. Results based on each model using a confusion matrix (TH, True Healthy; FH, False Healthy; TU, True Unhealthy; FU, False Unhealthy; Acc., Accuracy (TH + TU)/(TH + TU + FH + FU)).

Features	Classifier	TH	FH	TU	FU	Acc. (%)	F1 Score (%)	No. of Selected Features
Using selected features by the Genetic algorithm	SVM-Linear	167	24	75	18	85.21	83.48	60
	SVM-RBF	172	18	81	13	89.08	87.84	56
	SVM-Gaussian	173	24	75	12	87.32	85.61	7
	SVM-Polynomial	161	11	88	24	87.68	86.80	54
	Decision Tree	168	24	75	17	85.56	83.83	62
	Logistic Regression	158	24	75	27	82.04	80.37	80

In SVM-Gaussian, the seven selected features were W0, C1-1 (max. of second derivative), C1-1 (variance) 2, C2-0 (first to last point ramp) 2, C4-0 (moment2 (ramp_removed_sig)) 2, and fractal dimensions. On the other hand, the main selected features for SVM-RBF were based on geometric features (15 features) and C2-0 (9 features). The most influential features in our study in descending order were geometrical features (in Table 1), C1-2, C2-0, fractal dimensions, C1-0, C1-1, and, finally, features based on the texture. The most-used features across all classifiers were y1, W2, and C1-1 (kurtosis (ramp_removed_sig)).

Although every primary feature was used either in its first, second, or third order, a total of 105 features were not used by any classifier, half of which belonged to the third power of the primary features. This was expected, as we designed the cost function of the GA not only to improve the classification performance but also to reduce the dimensionality of the feature vector.

Each run of the GA may result in a different set of features. Figure 10 shows the selection process of the GA for SVM-RBF, SVM-Polynomial, SVM-Gaussian, and decision tree during the 500 iterations. Each of these binary images has 500 columns (corresponding to 500 iterations) and 204 rows (associated with 204 features). Ones (white pixels) and zeros (black pixels) represent selected or missed corresponding features in the relevant iteration. The horizontal white line represents a feature that was steadily chosen by the GA. The most unstable accuracy based on feature selection by the GA was for the decision tree model with less than 2% fluctuations in accuracy.

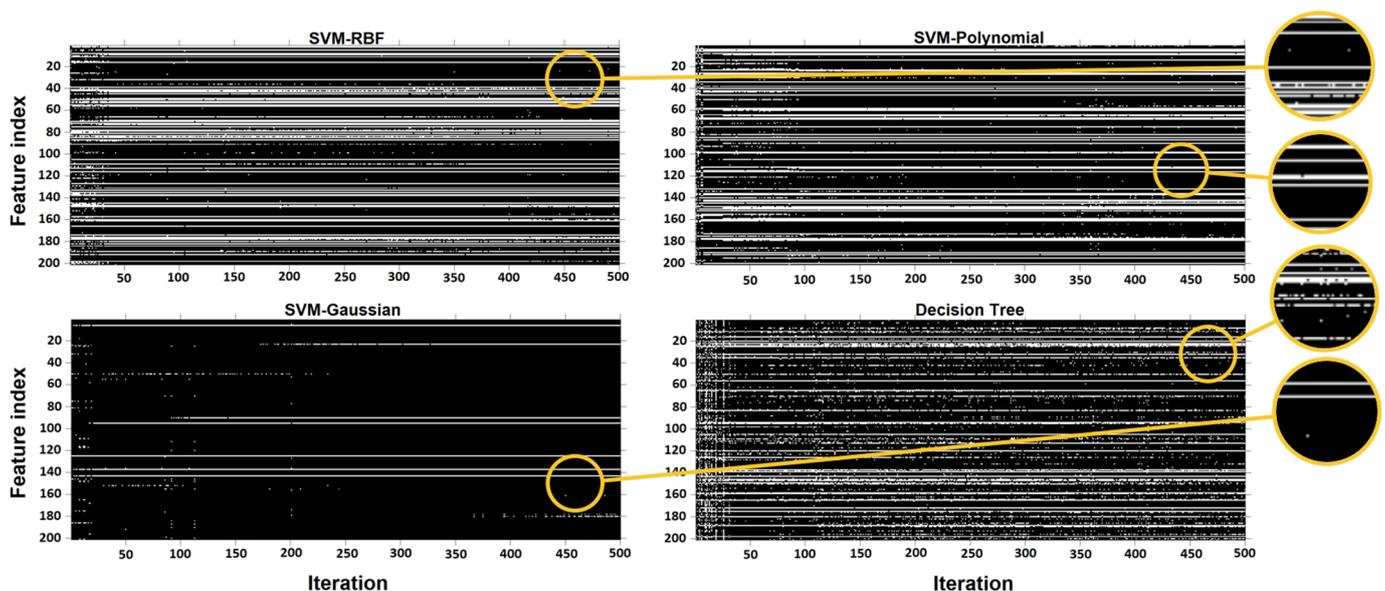


Figure 10. Representation of selected/missed features. Representation of selected/missed features during 500 iterations of the Genetic algorithm. White pixels represent a chosen feature. The feature index is based on the position of the features in Table 1, from left to right and up to down.

In another experiment, different classifiers were employed to evaluate the feature vector that resulted in the best classification performance by the SVM-RBF classifier. All experiments were performed 10 times and Figure 11 shows the results.

Comparing the results obtained by all features and the 56 features selected by the GA for the SVM-RBF, which are shown in Figures 8 and 11, respectively, reveals that the selected features outperform the result of using all features in terms of classification accuracy. However, the accuracy is still considerably lower than that of the GA-selected features.

In other words, the best performance of each classifier was achieved by the feature vector optimized by the GA for that classifier. Another notable finding from Figure 11 is the small variance in the SVM-RBF with the optimized feature vector. Since the feature vector is optimized for this specific classifier, the results of the 10 experiments are close. This may be considered to be a stability factor for the selected feature vector.

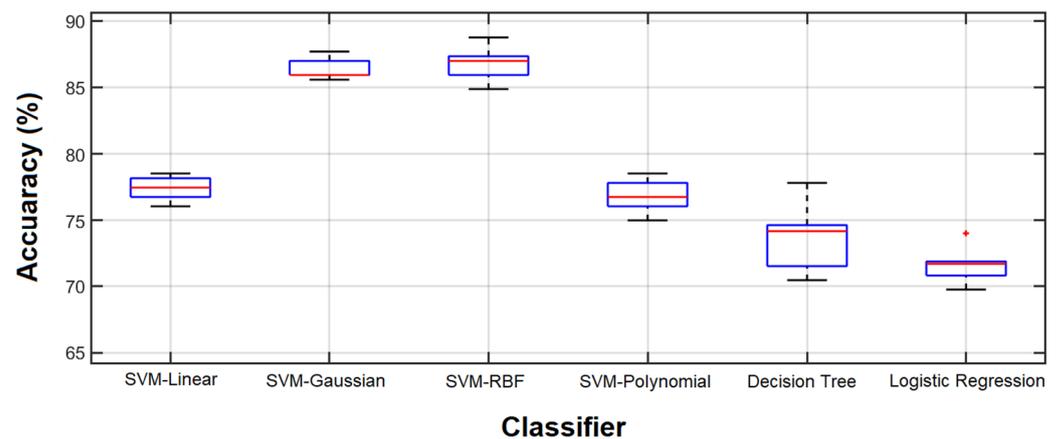


Figure 11. Classification accuracy. Classification accuracy using 56 features selected by the GA coupled with SVM-RBF.

Among the different scan parameters used in our study, images with scan parameters of T1/TR = 536.00 ms/TE = 11.00 ms/Flip-Angle = 180 produced the best results with an accuracy of 97.73% (only one false unhealthy detection).

The reason for our use of hand-crafted feature selection/extraction techniques and manual classification in this work is due to the low number of images obtained. Due to the fact that the database of recent studies is not the same, it does not seem fair to evaluate the studies one by one. However, we compared our proposed model with some well-known deep convolutional networks based on our collected database. The compared convolutional networks included ResNet17 [42], VGG19 [43], GoogleNet [44], and Inceptionv3 [45]. The results obtained after 500 iterations are presented in Table 3. As is shown, the obtained classification accuracy for the ResNet and VGG networks is almost equal to the proposed model. However, these networks need more time to reach convergence, which will increase the computational load of the algorithm. The performance of convolutional networks can be increased for the purpose of classification in the case of increasing data. According to Table 3, it has been proven that, for data with smaller dimensions, the proposed method has a higher performance in terms of speed and accuracy.

Table 3. Comparing the effectiveness of the proposed model with other networks.

Methods	ACC. (%)	F1 Score (%)	Time per Iteration
ResNet17	89.05	87.14	4.2 s
VGG19	88.90	89.1	7.1 s
GoogleNet	79.47	80.36	9 s
Inception v3	85.04	86.74	8.5 s
Proposed Model	89.08	87.84	1 s

Due to the nature of MRI, the slightest movement of the patient can cause noise in the captured images. Accordingly, the proposed method should be designed in such a way that it can be resistant to movement and environmental noises. In order to evaluate the proposed model more accurately, we artificially added Gaussian white noise in a wide range of different SNRs to the images and present the results based on the compared ResNet17, VGG19, GoogleNet, and Inceptionv3 pre-trained deep networks in Figure 12. According to Figure 12, as can be seen, the proposed model based on the selection/extraction of engineering features can be more resistant to noise than the ResNet17, VGG19, GoogleNet, and Inceptionv3 networks in a wide range of different SNRs. So, at an SNR = -4, it still remains above 80%.

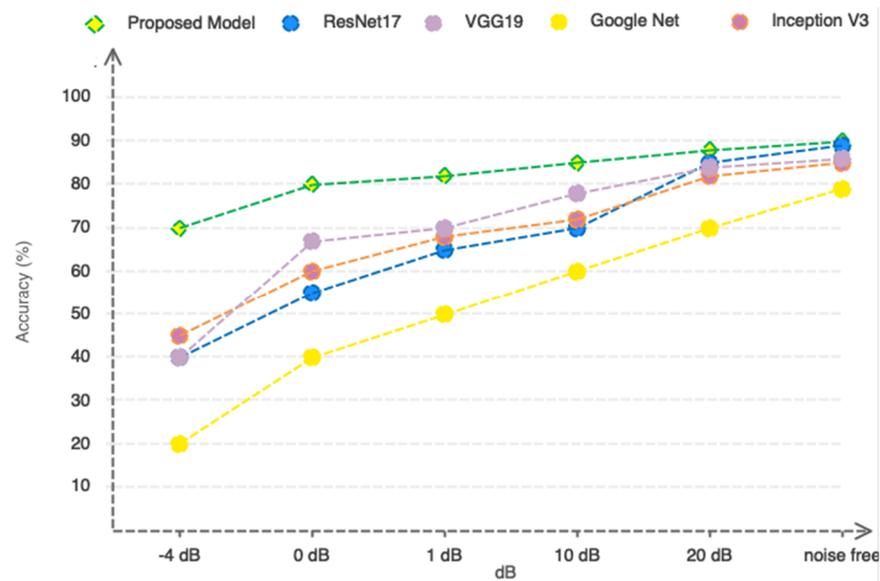


Figure 12. Comparing the performance of the proposed model with pre-trained networks in a noisy environment.

4. Discussion

Studies have shown that MRI scans and extracted parameters can be used to discriminate between healthy and unhealthy human bones quantitatively and to assess fracture risk. Deniz et al. [20] used deep convolutional neural networks (CNNs) to carry out segmentation of the PFBs using MRI automatically and compared their results with PFBs manually segmented by an expert. They used 86 subjects and their segmentation method resulted in 95% accuracy. This segmentation method may produce better results than our classification method; however, it needs a higher number of subjects and the same number of scan parameters.

Our results demonstrate the superiority of the SVM algorithm. However, we used a basic function for decision tree and logistic regression, while more sophisticated models for these techniques may produce better results. In addition, we investigated the effect of extracting geometric features (see Figure 2) on osteoporosis for the first time among previous studies, and we concluded that there is a direct relationship between geometric features and the early detection of osteoporosis. In addition, we compared our proposed method, which is based on extracting engineering features from images, to deep learning methods and concluded that our proposed method is resistant to environmental noises and can be used in practice.

Makond et al. [32] used a decision tree model based on CHAID's algorithms for postmenopausal osteoporosis screening, which resulted in a sensitivity ($TP/(TP + FN)$) of 92.3% and a 82.8% positive predictive value ($TP/(TP + FP)$). Ferizi et al. [46] used different classifiers to predict fragility fractures from MRI data. The top three classifiers in their study were RUS-boosted trees, logistic regression, and linear discriminant, which outperformed linear, cubic, and quadratic SVM.

We considered the weight of a false healthy diagnosis (patients that were diagnosed falsely as healthy by the classifier) to be more than that of a false unhealthy diagnosis. The former misses the chance to be diagnosed, which is the aim of this study, and the latter after referral to a DEXA scan or further investigation will be diagnosed as healthy. Thus, even though the accuracy and F1 score of SVM-Polynomial was lower than that of SVM-RBF, it may be a better classifier for our purpose.

Sollmann et al. showed a strong correlation between $T2^*$ and the BMD [19]. These researchers showed a significant positive correlation between $R2(=1/T2)$, $R2^*(=1/T2^*)$ and

$R2' (=R2^* - R2)$ and the BMD in the coronal section of the femoral neck. The use of this parameter as a feature could potentially improve the classification accuracy.

The small influence of texture-based features in our study may be due to the different acquisition protocols used during MRI, which can affect the results of texture analysis [17]. Furthermore, the different measuring techniques used in different centers and by different manufacturers produce different patterns in the texture [17], which influences texture analysis. Thus, further investigation with a higher number of images with the same scan parameters and the same/similar machines is necessary to evaluate the importance of texture-based features in comparison to the geometric ones.

The SVM algorithm is suitable for small sample sizes. However, a small sample size, a large number of variables, and the inclusion of redundant variables, especially when samples cannot fully represent the distribution of the sample space, may lead to overfitting [17]. Borra and Di Siaccio [37,38] showed that 10-fold cross-validation and the parametric bootstrap estimator had the best performance in a simulation with varying signal-to-noise ratios and sample sizes (using regression trees, projection pursuit regression, and neural networks). The use of our validation technique (10-fold cross-validation) increases our confidence in the results.

Since it is still unclear which features are the most discriminative ones for the detection of osteoporosis from PFB images, we plan to employ deep neural network models using a higher number of images with the same scan parameters to identify the most appropriate features for this purpose in a future study. In this way, an end-to-end deep model can be utilized to perform both feature extraction and classification.

Despite its efficiency and effectiveness, this study has the following limitations.

First, when the number of databases grows, the current method will be unable to fulfill the needs of the medical community on a large scale because the use of engineering features based on machine learning techniques does not ensure the optimality of the feature vector in diverse scenarios. Second, if the database is expanded, deep learning techniques should be applied and analyzed. Third, in order to increase the amount of data, the performance of generative adversarial networks should be examined in future studies.

5. Conclusions

To our knowledge, this is the first study to carry out qualitative classification of human PFB using machine learning and genetic selection algorithms to identify subjects with unhealthy PFBs by reference to the results obtained from DEXA scans.

The SVM resulted in higher accuracy and was superior to the other algorithms. Geometric characteristics in femoral bone demonstrated the highest influence on patient discrimination. However, further investigation with a higher number of images of the same imaging protocol is suggested. Also, the effect of other segmentation techniques (such as deep convolutional neural networks) and other relevant features (such as $T2^*$) on the classification performance should be investigated in future studies.

Author Contributions: Conceptualization, M.N.; methodology, S.N.R. and S.D.; software, T.Y.R. and M.N.; validation, T.Y.R.; formal analysis, S.N.R. and S.D.; investigation, S.N.R.; resources, M.N., S.D., T.Y.R. and S.D.; writing—original draft preparation, S.N.R. and M.N.; writing—review and editing, All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: We have made our dataset available to the public at <https://github.com> (file name: Mojtab2023/Classification-of-Proximal-Femoral-Bone-Using-Geometric-Features-and-Texture-Analysis-in-MR-Images-f (1 January 2022).

Acknowledgments: The authors would like to extend their gratitude and acknowledgements to Hanieh Abdi for her tremendous support with python programming and to Nima Sahba for his support with DEXA and MRI scans.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lei, C.; Song, J.-H.; Li, S.; Zhu, Y.-N.; Liu, M.-Y.; Wan, M.-C.; Mu, Z.; Tay, F.R.; Niu, L.-N. Advances in materials-based therapeutic strategies against osteoporosis. *Biomaterials* **2023**, *296*, 122066. [[CrossRef](#)] [[PubMed](#)]
2. Yang, L.; Chen, C.; Zhang, Z.; Wei, X. Diagnosis of bone mineral density based on backscattering resonance phenomenon using coregistered functional laser photoacoustic and ultrasonic probes. *Sensors* **2021**, *21*, 8243. [[CrossRef](#)]
3. Holubiak, I.S.; Leuciuc, F.V.; Crăciun, D.M.; Dobrescu, T. Effect of strength training protocol on bone mineral density for postmenopausal women with osteopenia/osteoporosis assessed by dual-energy x-ray absorptiometry (DEXA). *Sensors* **2022**, *22*, 1904. [[CrossRef](#)] [[PubMed](#)]
4. Shahini, N.; Bahrami, Z.; Sheykhivand, S.; Marandi, S.; Danishvar, M.; Danishvar, S.; Roosta, Y. Automatically identified EEG signals of movement intention based on CNN network (End-To-End). *Electronics* **2022**, *11*, 3297. [[CrossRef](#)]
5. Yu, W.; Xie, Z.; Li, J.; Lin, J.; Su, Z.; Che, Y.; Ye, F.; Zhang, Z.; Xu, P.; Zeng, Y. Super enhancers targeting ZBTB16 in osteogenesis protect against osteoporosis. *Bone Res.* **2023**, *11*, 30. [[CrossRef](#)]
6. Wang, H.; Luo, Y.; Wang, H.; Li, F.; Yu, F.; Ye, L. Mechanistic advances in osteoporosis and anti-osteoporosis therapies. *MedComm* **2023**, *4*, e244. [[CrossRef](#)] [[PubMed](#)]
7. Marcucci, G.; Domazetovic, V.; Nediani, C.; Ruzzolini, J.; Favre, C.; Brandi, M.L. Oxidative stress and natural antioxidants in osteoporosis: Novel preventive and therapeutic approaches. *Antioxidants* **2023**, *12*, 373. [[CrossRef](#)] [[PubMed](#)]
8. Sabahi, K.; Sheykhivand, S.; Mousavi, Z.; Rajabioun, M. Recognition COVID-19 cases using deep type-2 fuzzy neural networks based on chest X-ray image. *Comput. Intell. Electr. Eng.* **2023**, *14*, 75–92.
9. Al-Saleh, Y.; Sulimani, R.; Sabico, S.; Alshahrani, F.M.; Fouda, M.A.; Almohaya, M.; Alaidarous, S.B.; Alkhawashki, H.M.; Alshaker, M.; Alrayes, H. Diagnosis and management of osteoporosis in Saudi Arabia: 2023 key updates from the Saudi Osteoporosis Society. *Arch. Osteoporos.* **2023**, *18*, 75. [[CrossRef](#)] [[PubMed](#)]
10. Tai, T.-W.; Huang, C.-F.; Huang, H.-K.; Yang, R.-S.; Chen, J.-F.; Cheng, T.-T.; Chen, F.-P.; Chen, C.-H.; Chang, Y.-F.; Hung, W.-C. Clinical practice guidelines for the prevention and treatment of osteoporosis in Taiwan: 2022 update. *J. Formos. Med. Assoc.* **2023**, *in press*. [[CrossRef](#)]
11. De Sire, A.; Lippi, L.; Venetis, K.; Morganti, S.; Sajjadi, E.; Curci, C.; Ammendolia, A.; Criscitiello, C.; Fusco, N.; Invernizzi, M. Efficacy of antiresorptive drugs on bone mineral density in post-menopausal women with early breast cancer receiving adjuvant aromatase inhibitors: A systematic review of randomized controlled trials. *Front. Oncol.* **2022**, *11*, 829875. [[CrossRef](#)]
12. Ciancia, S.; Högler, W.; Sackers, R.J.; Appelman-Dijkstra, N.M.; Boot, A.M.; Sas, T.C.; Renes, J.S. Osteoporosis in children and adolescents: How to treat and monitor? *Eur. J. Pediatr.* **2023**, *182*, 501–511. [[CrossRef](#)] [[PubMed](#)]
13. Khare, M.R.; Havaladar, R.H. Non-invasive methodological techniques to determine health of a bone. In *Techno-Societal 2020: Proceedings of the 3rd International Conference on Advanced Technologies for Societal Applications*; Springer: Cham, Switzerland, 2021; Volume 1, pp. 343–350.
14. Guerri, S.; Mercatelli, D.; Gómez, M.P.A.; Napoli, A.; Battista, G.; Guglielmi, G.; Bazzocchi, A. Quantitative imaging techniques for the assessment of osteoporosis and sarcopenia. *Quant. Imaging Med. Surg.* **2018**, *8*, 60. [[CrossRef](#)] [[PubMed](#)]
15. Kanis, J.A.; McCloskey, E.V.; Harvey, N.C.; Cooper, C.; Rizzoli, R.; Dawson-Hughes, B.; Maggi, S.; Reginster, J.-Y. The need to distinguish intervention thresholds and diagnostic thresholds in the management of osteoporosis. *Osteoporos. Int.* **2023**, *34*, 1–9. [[CrossRef](#)] [[PubMed](#)]
16. Lungaro, L.; Manza, F.; Costanzini, A.; Barbalinardo, M.; Gentili, D.; Caputo, F.; Guarino, M.; Zoli, G.; Volta, U.; De Giorgio, R. Osteoporosis and celiac disease: Updates and hidden pitfalls. *Nutrients* **2023**, *15*, 1089. [[CrossRef](#)]
17. Han, M.; Chiba, K.; Banerjee, S.; Carballido-Gamio, J.; Krug, R. Variable flip angle three-dimensional fast spin-echo sequence combined with outer volume suppression for imaging trabecular bone structure of the proximal femur. *J. Magn. Reson. Imaging* **2015**, *41*, 1300–1310. [[CrossRef](#)]
18. Dittrich, A.T.; Janssen, E.J.; Geelen, J.; Bouman, K.; Ward, L.M.; Draaisma, J.M. Diagnosis, Follow-Up and Therapy for Secondary Osteoporosis in Vulnerable Children: A Narrative Review. *Appl. Sci.* **2023**, *13*, 4491. [[CrossRef](#)]
19. Sollmann, N.; Löffler, M.T.; Kronthaler, S.; Böhm, C.; Dieckmeyer, M.; Ruschke, S.; Kirschke, J.S.; Carballido-Gamio, J.; Karampinos, D.C.; Krug, R. MRI-based quantitative osteoporosis imaging at the spine and femur. *J. Magn. Reson. Imaging* **2021**, *54*, 12–35. [[CrossRef](#)]
20. Deniz, C.M.; Xiang, S.; Hallyburton, R.S.; Welbeck, A.; Babb, J.S.; Honig, S.; Cho, K.; Chang, G. Segmentation of the proximal femur from MR images using deep convolutional neural networks. *Sci. Rep.* **2018**, *8*, 16485. [[CrossRef](#)]
21. Sheykhivand, S.; Yousefi Rezaei, T.; Mousavi, Z.; Meshini, S. Automatic stage scoring of single-channel sleep EEG using CEEMD of genetic algorithm and neural network. *Computational Intelligence in Electrical Engineering* **2018**, *9*, 15–28.
22. Dhanaji Kale, K.; Ainapure, B.; Nagulapati, S.; Sankpal, L.; Sambhajirao Satpute, B. Chronological-hybrid optimization enabled deep learning for boundary segmentation and osteoporosis classification using femur bone. *Imaging Sci. J.* **2023**. [[CrossRef](#)]

23. Franco-Gonçalo, P.; Pereira, A.I.; Loureiro, C.; Alves-Pimenta, S.; Filipe, V.; Gonçalves, L.; Colaço, B.; Leite, P.; McEvoy, F.; Ginja, M. Femoral Neck Thickness Index as an Indicator of Proximal Femur Bone Modeling. *Vet. Sci.* **2023**, *10*, 371. [[CrossRef](#)] [[PubMed](#)]
24. Dolz, J.; Desrosiers, C.; Ayed, I.B. 3D fully convolutional networks for subcortical segmentation in MRI: A large-scale study. *NeuroImage* **2018**, *170*, 456–470. [[CrossRef](#)] [[PubMed](#)]
25. Ong, W.; Zhu, L.; Tan, Y.L.; Teo, E.C.; Tan, J.H.; Kumar, N.; Vellayappan, B.A.; Ooi, B.C.; Quek, S.T.; Makmur, A. Application of Machine Learning for Differentiating Bone Malignancy on Imaging: A Systematic Review. *Cancers* **2023**, *15*, 1837. [[CrossRef](#)]
26. Chay, Z.E.; Lee, C.H.; Lee, K.C.; Oon, J.S.; Ling, M.H. Russel and Rao coefficient is a suitable substitute for Dice coefficient in studying restriction mapped genetic distances of *Escherichia coli*. *arXiv* **2023**, arXiv:2302.12714.
27. Qingyun, F.; Zhaokui, W. Fusion Detection via Distance-Decay Intersection over Union and Weighted Dempster–Shafer Evidence Theory. *J. Aerosp. Inf. Syst.* **2023**, *20*, 114–125. [[CrossRef](#)]
28. Xu, Y.; Li, S.; Wang, Z.; Zhang, H.; Li, Z.; Xiao, B.; Guo, W.; Liu, L.; Bai, P. Design of multi-DC overdriving waveform of electrowetting displays for gray scale consistency. *Micromachines* **2023**, *14*, 684. [[CrossRef](#)]
29. Lv, Z.; Li, J.; Li, X.; Wang, H.; Wang, P.; Li, L.; Shu, L.; Li, X. Two adaptive enhancement algorithms for high gray-scale RAW infrared images based on multi-scale fusion and chromatographic remapping. *Infrared Phys. Technol.* **2023**, *133*, 104774. [[CrossRef](#)]
30. Tayefi, M.; Tajfard, M.; Saffar, S.; Hanachi, P.; Amirabadizadeh, A.R.; Esmaily, H.; Taghipour, A.; Ferns, G.A.; Moohebbati, M.; Ghayour-Mobarhan, M. hs-CRP is strongly associated with coronary heart disease (CHD): A data mining approach using decision tree algorithm. *Comput. Methods Programs Biomed.* **2017**, *141*, 105–109. [[CrossRef](#)]
31. Yeo, B.; Grant, D. Predicting service industry performance using decision tree analysis. *Int. J. Inf. Manag.* **2018**, *38*, 288–300. [[CrossRef](#)]
32. Makond, B.; Pornsawad, P.; Thawnashom, K. Decision Tree Modeling for Osteoporosis Screening in Postmenopausal Thai Women. *Informatics* **2022**, *9*, 83. [[CrossRef](#)]
33. Louk, M.H.L.; Tama, B.A. Dual-IDS: A bagging-based gradient boosting decision tree model for network anomaly intrusion detection system. *Expert Syst. Appl.* **2023**, *213*, 119030. [[CrossRef](#)]
34. Ambrish, G.; Ganesh, B.; Ganesh, A.; Srinivas, C.; Mensinkal, K. Logistic regression technique for prediction of cardiovascular disease. *Glob. Transit. Proc.* **2022**, *3*, 127–130.
35. Hu, Y.; Fan, Y.; Song, Y.; Li, M. A general robust low-rank multinomial logistic regression for corrupted matrix data classification. *Appl. Intell.* **2023**, *53*, 18564–18580. [[CrossRef](#)]
36. Gulati, K.; Kumar, S.S.; Boddu, R.S.K.; Sarvakar, K.; Sharma, D.K.; Nomani, M. Comparative analysis of machine learning-based classification models using sentiment classification of tweets related to COVID-19 pandemic. *Mater. Today Proc.* **2022**, *51*, 38–41. [[CrossRef](#)]
37. Pitchai, R.; Supraja, P.; Sulthana, A.R.; Veeramakali, T.; Babu, C.M. MRI image analysis for cerebrum tumor detection and feature extraction using 2D U-ConvNet and SVM classification. *Pers. Ubiquitous Comput.* **2023**, *27*, 931–940. [[CrossRef](#)]
38. Meng, X.; Wei, Q.; Meng, L.; Liu, J.; Wu, Y.; Liu, W. Feature fusion and detection in Alzheimer’s disease using a novel genetic multi-kernel SVM based on MRI imaging and gene data. *Genes* **2022**, *13*, 837. [[CrossRef](#)]
39. Yoon, J.K.; Choi, J.-Y.; Rhee, H.; Park, Y.N. MRI features of histologic subtypes of hepatocellular carcinoma: Correlation with histologic, genetic, and molecular biologic classification. *Eur. Radiol.* **2022**, *32*, 5119–5133. [[CrossRef](#)]
40. Liang, T.-O.; Koh, Y.H.; Qiu, T.; Li, E.; Yu, W.; Huang, S.Y. High-performance permanent magnet array design by a fast genetic algorithm (GA)-based optimization for low-field portable MRI. *J. Magn. Reson.* **2022**, *345*, 107309. [[CrossRef](#)]
41. Do, D.T.; Yang, M.-R.; Lam, L.H.T.; Le, N.Q.K.; Wu, Y.-W. Improving MGMT methylation status prediction of glioblastoma through optimizing radiomics features using genetic algorithm-based machine learning approach. *Sci. Rep.* **2022**, *12*, 13412. [[CrossRef](#)]
42. Bharati, S.; Podder, P.; Mondal, M.; Prasath, V. CO-ResNet: Optimized ResNet model for COVID-19 diagnosis from X-ray images. *Int. J. Hybrid Intell. Syst.* **2021**, *17*, 71–85. [[CrossRef](#)]
43. Khaleghi, N.; Rezaei, T.Y.; Beheshti, S.; Meshgini, S.; Sheykhivand, S.; Danishvar, S. Visual saliency and image reconstruction from EEG signals via an effective geometric deep network-based generative adversarial network. *Electronics* **2022**, *11*, 3637. [[CrossRef](#)]
44. Hajipour Khire Masjidi, B.; Bahmani, S.; Sharifi, F.; Peivandi, M.; Khosravani, M.; Hussein Mohammed, A. CT-ML: Diagnosis of breast cancer based on ultrasound images and time-dependent feature extraction methods using contourlet transformation and machine learning. *Comput. Intell. Neurosci.* **2022**, *2022*, 1493847. [[CrossRef](#)] [[PubMed](#)]
45. Madanan, M.; Sayed, B.T. Designing a deep learning hybrid using CNN and Inception V3 transfer learning to detect the aggression level of deep obsessive compulsive disorder in children. *Int. J. Biol. Biomed. Eng* **2022**, *16*, 207–220. [[CrossRef](#)]
46. Ferizi, U.; Besser, H.; Hysi, P.; Jacobs, J.; Rajapakse, C.S.; Chen, C.; Saha, P.K.; Honig, S.; Chang, G. Artificial intelligence applied to osteoporosis: A performance comparison of machine learning algorithms in predicting fragility fractures from MRI data. *J. Magn. Reson. Imaging* **2019**, *49*, 1029–1038. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.