

Available online at www.qu.edu.iq/journalcm

JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS

ISSN:2521-3504(online) ISSN:2074-0204(print)



Hybrid Fuzzy Harmony Search and Multi-Modal Feature Fusion for Knee Osteoporosis Diagnosis

Saud M. Abdul Razzaq^{a*}, Baydaa I. Khaleel^b

^aDepartment of Computer Sciences, College of Computer Science and Mathematics, University of Mosul, Iraq
Email: saud.23csp48@student.uomosul.edu.iq

^bDepartment of Computer Sciences, College of Computer Science and Mathematics, University of Mosul, Iraq. Email: baydaaibraheem@uomosul.edu.iq

ARTICLE INFO

Article history:

Received: 15 /05/2026

Revised form: 14 /06/2026

Accepted : 17 /06/2026

Available online: 30 /09/2026

Keywords:

Knee

x-ray images

osteoporosis

harmony search algorithm

Gath-Geva.

ABSTRACT

Osteoporosis is a common bone disease and a major contributing factor to fractures among the elderly, directly impacting the quality of life of those affected. Therefore, early diagnosis is crucial. We proposed developing an intelligent system for detecting and classifying knee osteoporosis by analyzing knee X-ray images. We proposed a novel feature extraction technique called Hybrid Multimodal Feature Fusion (HMMFF), which combines traditional techniques such as GLCM, HOG, and ORB with deep learning models, including Squeezenet, YOLOv11, VGG-16, and ResNet50. The proposed system was applied to two databases for binary and multi-classification. To improve the classification efficiency of the proposed system, feature selection was performed using the harmony search optimization algorithm. We also proposed a new hybrid feature selection method, the Hybrid Fuzzy Harmony Search Algorithm (HFHSA), that combines harmony search with Gath-Geva fuzzy clustering. The two techniques were evaluated and compared by feeding the selected feature sets into several machine learning classifiers. The results showed the superiority of the HFHSA technique for feature selection over Harmony Search alone, as the HistGradient Boosting model achieved the highest accuracy across the binary classifiers, reaching 96.8% in binary classification while the XGBoost model achieved 95.38% in multi-classification. This performance improvement demonstrates that applying the HMMFF feature extraction and fusion technique, along with feature selection using HFHSA, confirms the effectiveness of integrating different AI techniques by extracting and selecting important features that contribute to improving the accuracy of classifying and diagnosing osteoporosis using input X-ray images.

MSC..

<https://doi.org/10.29304/jqcm.2026.18.32893>

1. Introduction

The use of new technologies in healthcare, such as the cloud and artificial intelligence-based data analysis, has led to better health outcomes and longer life expectancy. As a result, the ageing population has increased, leading to an increased incidence of age-related diseases like osteoporosis. Osteoporosis is defined by the World Health Organization as a skeletal disorder defined by low bone mass and structural deterioration of bone tissue, with a consequent increase in fracture risk [1-3]. Osteoporosis is the most common cause of 30-40% of all bone fractures in women, and the disease is expected to affect one billion people by 2050 [4, 5].

*Corresponding author: Saud M. Abdul Razzaq

Email addresses: saud.23csp48@student.uomosul.edu.iq

Communicated by 'sub etitor'

Knee osteoporosis is one of the more important forms of the disease, given that the knee is a weight bearing joint and is subject to stresses during normal activities. Osteoporosis affects bone density and damages cartilage in the knee joint, leading to pain and loss of function, the condition is associated with age, female gender (particularly postmenopausal vitamin D), lack of exercise and smoking. So early detection is important as not only does this enhance the effectiveness of treatment, it also prevents fractures, decreasing the strain on health care resources [6-8].

Traditional X-rays are the most common imaging technique used to evaluate bone diseases, providing clinicians with important information about bone density, bone changes and fractures [9]. But early visual signs of changes in bone density are subtle. Machine learning is a powerful tool that can be used to complement conventional imaging to detect complex patterns and indicators that might not be noticed by the human eye [10, 11].

By examining existing literature in depth, it was found that there is a number of important research gaps that this work fills. Existing research on knee X-ray-based osteoporosis diagnosis has focused mainly on knee X-ray features based on traditional methods or on deep learning models individually. Moreover, the combination of multiple traditional feature extraction methods and multiple deep learning models was not previously performed in a single feature vector, an issue that this proposed Hybrid Multimodal Feature Fusion (HMMFF) directly solves. Likewise, the search optimization for combinatorial search and fuzzy clustering algorithms like Gath-Geva have been explored independently in the machine learning literature, but there is no prior research on the fusion of the two in a hybrid framework for feature selection in medical imaging applications, including osteoporosis detection. This paper presents a novel multi-criteria fuzzy fitness function, which improves the classification accuracy, feature stability and feature redundancy, and thus overcomes the limitation of the traditional single-scale evaluation functions in standard metaheuristic feature selection, the proposed Hybrid Fuzzy Search Algorithm (HFHSA).

In this research, we proposed a methodology of the automated detection of knee osteoporosis and the efficacy of several machine learning methods to detect the disease correctly. It also provides a number of significant contributions:

- An innovative method of HMMFF feature extraction is proposed, that is, combining the traditional methods (GLCM, HOG, ORB) with deep learning-based features of several models (Squeezenet, YOLOv11, VGG-16, ResNet50), which allows obtaining a variety of comprehensive features.
- A novel feature selection method is proposed using a hybrid HFHSA technique that integrates harmony search optimization with Gath-Geva fuzzy clustering. This method leverages the comprehensive search capabilities of harmony search and the uncertainty management of fuzzy logic to identify optimal feature sets. This can improve the identification of important data related to bone density, thereby increasing the performance efficiency of classifiers.

2. Related Works

There is extensive research in artificial intelligence related to identifying osteoporosis and assessing the internal structure of bones through the application of various artificial intelligence techniques, these works are listed below.

In 2022, Yang presented research on detecting and classifying knee x-ray images using deep-learning CNNs, specifically the VGG16 model and Late-Fusion techniques, the researcher found that these strategies were effective in detecting the degree of knee osteoporosis in x-ray images, however, there has been no study using deep learning to identify x-ray images in the context of knee osteoporosis, therefore, the researcher claimed that the same strategy would effectively assist physicians in identifying and evaluating X-ray images, the researcher compared the results of the models, with the Late-Fusion model achieving 77% accuracy on the OA dataset and the VGG16 model reaching 82% accuracy on the OS dataset, the researcher proposed the use of deep learning methods to detect knee osteoporosis [12].

In 2023, Wani and Arora proposed a method for diagnosing osteoporosis from knee X-ray images using transfer learning in convolutional neural networks, such as ResNet-18, VggNet-16, AlexNet, and VggNet-19, the results show that AlexNet achieved the best performance with an accuracy of 91%, while VggNet-19 achieved the lowest accuracy with an accuracy of 84.2%, The study suggests that a deep learning system with transfer learning can aid doctors in detecting osteoporosis early, thereby reducing the risk of fractures [13].

In 2023, Dodamani and Ajit presented a comparative study to evaluate the accuracy and efficiency of contemporary deep learning CNN models, including VGG16, VGG19, DenseNet121, ResNet50, and InceptionV3, in detecting and classifying normal and osteoporotic conditions, the research used 830 X-ray images of the spine, hand, leg, knee, and hip, comprising 420 normal cases and 410 osteoporotic cases, different performance metrics were used to evaluate each model, the results revealed that DenseNet121 performed exceptionally well, with an accuracy of 93.4%, an error rate of 0.07, and a validation loss of only 0.57 compared to the other models evaluated in this research [14].

In 2023, Kumar et al. presented a model designed to diagnose osteoporosis, the ability to distinguish between X-rays of healthy individuals and those obtained from individuals with osteoporosis was tested using eight well-known models pre-trained on ImageNet to provide a comprehensive comparison, the database consisted of 372 X-ray images, divided into training and test sets for validation purposes, standard evaluation parameters (such as precision, recall, and f1 score) were calculated for all architectures, and the majority of designs performed remarkably well, with the highest average accuracy reaching 86.36% when distinguishing between specific classes [15].

In 2023, Kumar et al. proposed a model for classifying osteoporosis knee x-ray datasets, an ensemble-based classifier model was built using three convolutional neural network (CNN) architectures: Inception v3, Xception, and ResNet 18, taking into account two different factors in the decision scores generated by the previously described base classifiers, the proposed ensemble technique uses a rank-based fuzzy consolidation of the classifiers, unlike simple consolidation procedures documented in the literature, the proposed ensemble approach considers the confidence level of the base classifiers' recommendations when making final predictions on test samples, a publicly available benchmark dataset was used to evaluate the constructed framework using a five-fold cross-validation technique, with a loss of 0.082, the proposed model achieved an accuracy of 93.5% [16].

In 2024, Shen conducted research on osteoporosis diagnosis using knee X-ray analysis using advanced deep learning algorithms, this strategy uses artificial intelligence technology to improve diagnostic accuracy and efficiency, providing a more convenient and non-invasive method for early detection and treatment of osteoporosis, this could reduce the risk of fractures and improve overall health outcomes for individuals, the VGG 19 model, known for its ability to extract complex features from 2D images, was used, the model achieved 89% accuracy using datasets from Kaggle and Mendeley, demonstrating its effectiveness in detecting osteoporosis features in knee x-rays [17].

In 2024, Naguib et al. conducted a study that categorized knee X-ray images into three classifications: normal, osteoporosis, and osteopenia. The study employed deep learning and an overflow mechanism, which proposed two distinct types of blocks. The proposed model was trained, tested, and validated using two knee datasets, with two pre-trained models, ResNet50 and AlexNet, undergoing transfer learning. The results were compared with the proposed model, which demonstrated superior performance over the pre-trained models. Among all pre-trained models, the Superfluity DL architecture achieved the highest accuracy, recording 85.42% for dataset 1 and 79.39% for dataset 2 [18].

Muhammad et al. presented a deep learning hybrid model to detect osteoporosis in 2025 based on the KOP dataset. This method is a combination of two pre-trained models, DenseNet169 and ViT, using a custom attention model. The classification performance was also greatly enhanced since it effectively obtained spatial and channel specific information with high accuracy (0.8611), high specificity (0.9474) and high precision (0.9286). The findings indicate that the model can assist in improving the accuracy and efficiency of osteoporosis diagnosis through studying the X-ray images of the knee, which contributes to the possibility of clinical use to improve the early detection and specific interventions. This study emphasizes that the adoption of advanced models is important to enhance bone characterization and creates the prospects of a further study into more various bone-related diseases[19].

In 2025, OM, A., & Gunasundari, R. created an innovative hybrid deep learning algorithm that combines a ResNet50 network in space feature extraction with Gateway Repetition Units (GRUs) in learning time dependency, which allows analyzing X-ray images on the knee better. The model had a classification accuracy of 95.65 and was able to differentiate between osteoporosis and normal conditions and made itself as a good non-invasive screening method of the early detection of osteoporosis[20].

3. Methodology

The research aims to implement an image-based classification of knee osteoporosis by creating and applying a proposed intelligence system to analyze knee X-ray images, these images are analyzed to determine the health of the

knee bones and detect osteoporotic injuries in the knee, the classification process is carried out on common osteoporosis cases, normal ,osteopenia, osteoporosis, this is achieved by using different machine learning models to perform the classification process and then comparing the results obtained from these models.

3.1. Data Set

The X-ray image dataset was obtained from the Kaggle website, which includes (1947)[21]. X-ray images of the knee divided into the following categories: normal ,osteopenia, osteoporosis of the knee area, The data were then balanced to using Random Over sampling Random Oversampling creates new minority-class samples by randomly duplicating existing samples from the minority class, thereby increasing their representation in the dataset ,verify the distribution of categories within the dataset and correct any imbalances, achieving a largely balanced distribution. Figure 1, illustrates the categories of X-ray images of knee osteoporosis cases. Table 1, shows the number of samples in the binary classification, while Table 2, displays the number of samples for each knee osteoporosis category.



Fig. 1 categories of X-ray images of knee osteoporosis cases.

Table 1. Number of samples, binary classification.

Class	Total
Normal	780
disease	793

Table 2. Number of samples in each category of knee osteoporosis.

Class	Total
Normal	780
Osteopenia	374
Osteoporosis	793

3.2. Images Pre- processing

Image processing is crucial for improving the quality of images input into the system. It prepares the data effectively and efficiently for subsequent stages, thereby contributing to image recognition and differentiation. The image processing was carried by cropping the image, focusing on the knee area, then applying out through a series of steps applied to X-ray images to simplify the computational complexity of handling them, as follows:

- Applying Contrast Limited Adaptive Histogram Equalization (CLAHE) to improve contrast, address uneven lighting in the image, and enhance the quality of images with blurred details.
- Applying a Gaussian Blur to reduce noise in the image.
- Improving image clarity and highlighting important details through Sharpening.
- Applying the Morphological Gradient to define and highlight edges in the images to reveal the external boundaries of bones. Figure 2, shows the preprocessing process for knee x-ray images.

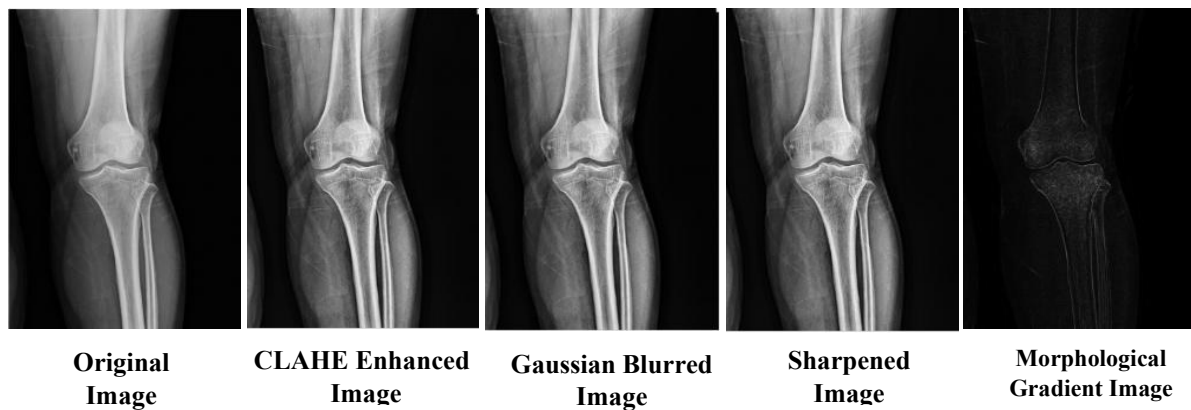


Fig. 2 The special pre-processing process for knee X-ray images.

3.3. Feature Extraction

In image classification, feature extraction plays a pivotal role in building accurate models, with the model's performance depending on the features' relevance and quality. Not only does the extraction of useless features add to the computational burden of an image processing system, it has a negative impact on classification performance, as features represent the core content of an image for various tasks, such as diagnosis and classification, an effective feature extraction approach is critical for any image processing system[22-24].

Our research introduces a new feature extraction technique called Hybrid Multi-Modal Feature Fusion (HMMFF) that combines complementary hand-crafted and deep learning features to form a robust image representation. Three conventional methods were used to extract handcrafted features. The Gray-Level Co-occurrence Matrix (GLCM) was calculated with a distance 1 and four angles (0° , 45° , 90° , 135°) using 256 gray levels, resulting in five texture features: contrast, dissimilarity, homogeneity, energy, and correlation, which can capture the changes in bone structure during the development of osteoporosis [25]. The Histogram of Oriented Gradients (HOG) was computed with 16×16 pixel cells and a 1×1 overlap in blocks, resulting in a 1,176-dimensional vector of edge orientations and structural changes (such as micro-fractures) [26, 27]. Finally, the Oriented FAST and Rotated BRIEF (ORB) descriptors were extracted, with a maximum of 100 keypoints, to produce a 500-dimensional vector of local features that were rotation and illumination invariant [28, 29].

Features were extracted from three deep learning models with their final classification layers removed to get the intermediate representation. The 16-layer VGG-16 model pre-trained on ImageNet provided a 512-dimensional vector of semantic features [30]. YOLOv11 provided a small 128-dimensional vector, taking advantage of its efficiency in visual detection. ResNet-50, a 50-layer residual network, contributed a 2,048-dimensional vector, encoding complex patterns, well suited for medical image processing were then extracted using the pre-trained Squeezenet model, which offers high accuracy and efficiency using its Fire Modules structure. This produced 512 features that were then used to obtain the final feature representation[31].

The final HMMFF representation was formed by concatenating all the feature vectors, This was done by combining all the extracted features to obtain a vector containing 4,896 features, including low-level texture features, geometric features, local keypoint descriptors and high-level semantic features, This formed a strong and discriminative feature space suitable for classification tasks, Then, a combined vector was used after feature selection for classification models. Figure 3, illustrates the steps of the HMMFF method for feature extraction.

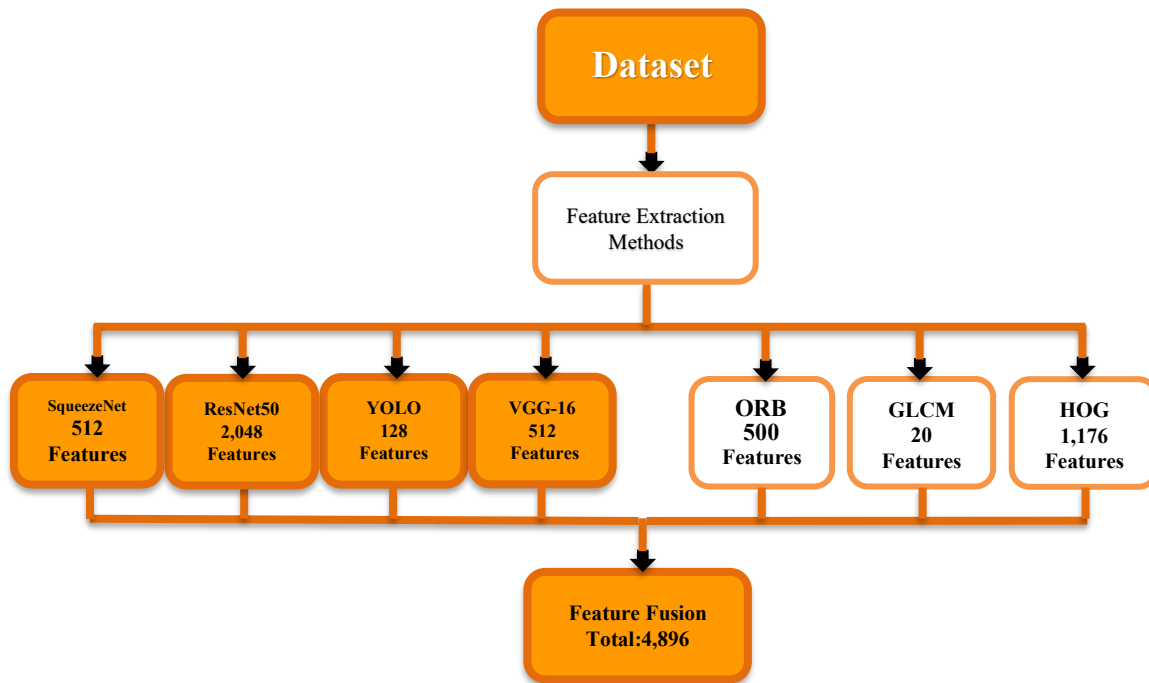


Fig. 3 HMMFF method for feature extraction.

3.4. Feature Selection

Feature selection is a major challenge for machine learning methodologies and approaches. It seeks to identify a subset of highly relevant and influential features from the original dataset, with the goal of improving the performance and accuracy of machine learning models by reducing computational complexity. Feature extraction and selection techniques are used to enhance model effectiveness [32]. Feature extraction transforms the original data into new variables that retain essential information, while feature selection discovers and represents simpler, more interpretable models of the relationships between variables and outcomes, reducing computational complexity and accelerating training time [32, 33].

The selection of the features in our paper is done in two folds: Firstly, in the first method, the harmony search algorithm is applied to run the feature selection process. In the second approach, feature selection is done under a new hybrid approach named HFHSA which integrates harmony search algorithm with Gath-Geva algorithm which is a fuzzy logic algorithm.

3.4.1. Hybrid Fuzzy Harmony Search Algorithm (HFHSA)

The new hybrid HFHSA method used in this work, combining the Gath-Geva algorithm and the harmony search algorithm, represents a new way to improve the performance of algorithms in feature selection tasks, this approach allows for broader exploration of the solution space, reduces reliance on fixed parameters, and overcomes the limitations of traditional evaluation functions in feature search algorithms, most of these functions focus on a single metric, such as accuracy, which can lead to the selection of unstable feature sets or those containing duplicate information, this reduces the risk of falling into local solutions and enhances the model's generalizability, improving its resistance to changes in data distribution.

Suggested in 2001 by a Korean scientist Zong Woo Geem, harmony search algorithm is a metaheuristic algorithm based on musical improvisation, which is based on the following basis: the harmony memory update equation, the harmony memory acceptance rate update equation, and the application of the pitch modulation rate to the determination of the optimal solution, by use of such mechanisms the algorithm tries to move efficiently through the search space, adapt to any potential modification in the optimization problem and converge to the global optimum without falling into it in local optimization [34, 35]. It is used to address complex problems through intelligent exploration of the solution space, the stochastic nature of the algorithm facilitates in-depth exploration of the search space and reduces the likelihood of falling into the trap of local minima, this property enables the harmony search

method to surpass traditional methods, this method is also characterized by its insensitivity to initial settings [36, 37].

The Harmony search algorithm is used to explore the space of possible solutions, while fuzzy logic is applied to evaluate the quality of these solutions, this is achieved by enhancing the fitness function by leveraging the strengths of each algorithm to enhance the flexibility and accuracy of the evaluation process, this is achieved by replacing traditional fitness functions with a multi-criteria fuzzy inference system, the quality of the solution (fitness) is evaluated based on three metrics: classification accuracy, feature stability, and redundancy between features, these values are converted into fuzzy metrics using the Gath-Geva model, which is based on mixed Gaussian distributions.

The evaluation process begins by calculating and finding the numerical sharp values for the specified metrics and for each solution (a set of features) in the harmony memory, membership functions derived from these are then used, the Gath-Geva algorithm converts these numeric values into fuzzy values, this is done according to the following steps:

Calculating basic metrics:

First, a clear numerical value is calculated for each of the three metrics for the proposed feature set, which represents a proposed solution in the Harmony algorithm, as follows:

1. Calculating accuracy:

This metric measures the model's ability to correctly classify using the selected set of features.

The standard mathematical equation for calculating accuracy is:

$$\text{Accuracy} = \frac{\text{Number of correct Prediction}}{\text{Total Number of samples}} \quad (1)$$

2. Stability Calculation:

Measures the stability of an individual feature in terms of its predictive ability and consistency, as it does not contain much noise or random fluctuations, stability is calculated by calculating two sub-measures:

- Variance-dependent stability: Measures the ratio of the variance of the feature to the mean, a feature is more stable when its variance is low compared to its mean.
- Entropy-dependent stability: Measures the degree of randomness in the distribution of values for a feature, a feature is more stable when its entropy is low. The following mathematical equation is used to calculate stability [38]:

$$\text{stability} = 0.6 \times \frac{1}{1 + \frac{\text{Var}(X)}{|\text{Mean}(X)| + \epsilon}} + 0.4 \times \left(1 - \frac{H(X)}{\log(N)}\right) \quad (2)$$

Where stability represents an indicator of the stability of the feature, $\text{Var}(X)$ is the variance in the values of the feature, ϵ is a small value to avoid dividing by zero, $H(X)$ is the entropy representing a measure of the randomness of the values in the feature, and N is the number of samples.

3.Redundancy:

It measures the degree of correlation between selected features. The goal is to select a set of features that provides diverse, non-redundant information.

The standard mathematical equation for average correlation coefficients is [38]:

$$\text{Redundancy} = \frac{2}{k(k-1)} \sum_{i=1}^{k-1} \sum_{j=i+1}^k |\rho_{ij}| \quad (3)$$

Where Redundancy represents an index of redundancy in the selected features, k represents the selected features, ρ_{ij} is the correlation coefficient between feature i and feature j .

B- Converting Metrics to Fuzzy Values:

To convert numerical values for solution evaluations into fuzzy values, the Gath-Geva algorithm is first trained on the training data. The model divides the feature space into a set of clusters, each with its own center, covariance matrix, and Gaussian distribution. The Gath-Geva algorithm is then used to generate a fuzzy affiliation for each metric, representing each metric as a multi-level fuzzy variable, and calculating the non-fuzzy value affiliation score for each level.

1. Calculating the Membership Score:

Each metric value is treated as a new data point. The values are normalized, and then their membership scores are calculated for each cluster.

The basic equation for calculating the membership score of point x_i in cluster C_j in the Gath-Geva algorithm is [39, 40]:

$$p_i(x_k) = \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} \exp\left(-\frac{1}{2}(x_k - \mu_i)^T \Sigma_i^{-1} (x_k - \mu_i)\right) \quad (4)$$

Then, the appropriate cluster is chosen for each of the three measures, where the highest values of belonging scores are selected for the accuracy and stability measures, while the lowest values of belonging scores are chosen for the repetition measure, given that the lower the repetition, the better the results. The measures are then combined after converting them to fuzzy values, using the following fuzzy inference system to obtain the final fitness value that guides the search process in the Harmony Search algorithm:

$$\text{Fitness} = (\text{Accuracy}_f)^{w_1} \times (\text{Stability}_f)^{w_2} \times (1 - \text{Redundancy}_f)^{w_3} \quad (5)$$

Where w_1, w_2, w_3 are weights used to determine the importance of each metric in the final evaluation.

The pseudo-codes better illustrates the steps of the HFHSA algorithm, and figure 4, shows the scheme of the new hybrid HFHSA method. The proposed hybrid method for feature selection demonstrates that incorporating fuzzy logic into the fitness calculation stages enhances the algorithm's ability to handle non-linear complexities in the data, reduces reliance on sharp values, and improves the efficiency of feature selection by increasing stability and reducing redundancy while maintaining predictive performance, thus, the approach followed in the study helps in discovering an ideal subset of features by reducing the complexity of the resulting set of features after merging, which will be used in machine learning models for classification.

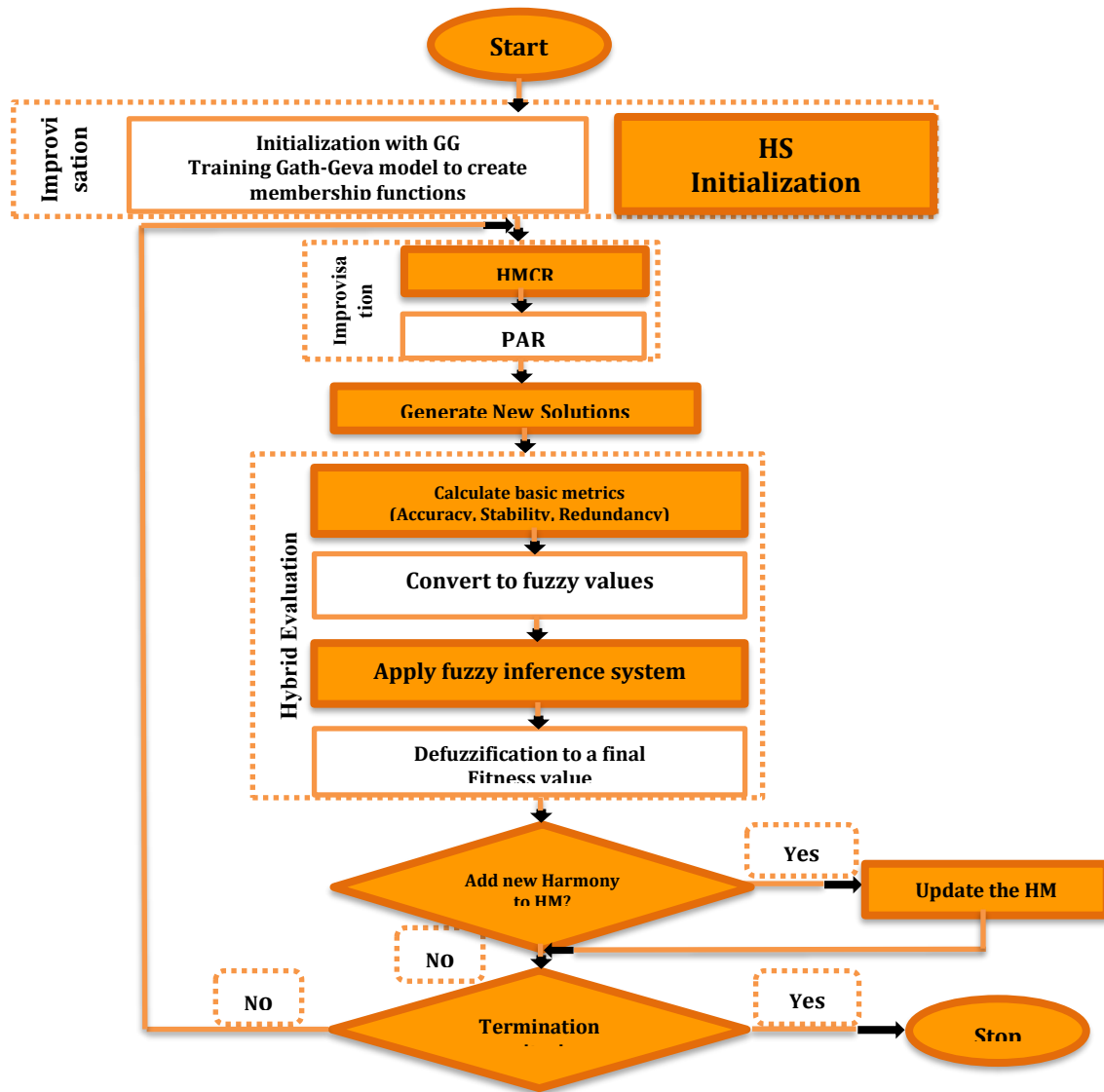


Fig. 4 Schematic diagram of the new hybrid HFHSA method.

Algorithm: Pseudo-codes of the (HFHSA) algorithm

Input: Combined Feature Space from HMMFF, Harmony Memory Size (HMS), HMCR, PAR, (w1, w2, w3),

Max_Iterations, and Fuzzy Weights

Output: Optimized Feature Subset

1. Initialize Harmony Memory (HM) and train Gath-Geva clustering models
2. While termination criterion is not satisfied do
3. Improvise a new harmony vector using HMCR and PAR
4. Calculate basic crisp metrics for the vector:
 - Accuracy using Equation (1)
 - Stability using Equation (2)
 - Redundancy using Equation (3)
5. Transform these metrics to fuzzy membership values using the Gath-Geva model via Equation (4)
6. Evaluate the final Fitness value using the Fuzzy Inference System via Equation (5)
7. If the new harmony vector's fitness is better than the worst in HM then
8. Update HM by replacing the worst solution with the new one
9. End If
10. End While
11. Return the optimal feature subset from HM

3.5. Performance Metrics

To evaluate the proposed system for detecting knee osteoporosis using knee X-ray images, this research used several benchmarks, including Recall, Precision, and F1-score, to evaluate the effectiveness of the system using the import library (confusion matrix) in Python to calculate specific performance evaluation metrics defined below [41, 42]:

Accuracy: It shows what proportion of the system's predictions are accurate:

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{True Positives} + \text{False Negatives} + \text{False Positives} + \text{True Negatives}} \quad (6)$$

Recall: Recall expresses the ratio of correct optimistic predictions to the total number of actual positive samples.

$$\text{Sensitivity} \text{ -(Recall)} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (7)$$

Precision: Accuracy expresses the ratio of correct optimistic predictions to the total number of positive predictions.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (8)$$

F1-score: The harmony mean of precision and recall used to evaluate the overall performance of a model.

$$\text{F1 - score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (9)$$

3.6. Proposed Method

The classification and detection processes were implemented in a set of systematic steps, as shown in Figure 5. These processes involve image preprocesses, e.g., noise cancellation and contrast amplification to enhance image quality and allow extracting image features. This is then preceded by data balancing which aims to eliminate disproportionate category distributions and ensure that the model results are less biased. The proposed new HMMFF technique is then used in the feature extraction process; it comprises of the extraction features with the use of traditional feature extraction algorithms (GLCM, HOG, ORB) and deep learning architectures (YOLOv11, VGG-16,

ResNet50, SqueezeNet). The attributes were obtained by turning off the classification layer in the models. The last outcome is a complete feature set of 4,896 features per image.

In order to enhance the classification performance of the proposed models by reducing the dimensions of the proposed models, two distinct feature selection strategies were carried out:

-Harmony Search that set the number of features down to 2,507 when searching binary classification and 2,494 when searching multi-classification.

-Hybrid HFHSA that turned out to be more efficient in detecting the most influential features, to 445 binary and 464 multi-classification.

Lastly, the phase of classification involved the introduction of the selected features to various machine learning models such as the Random Forest, ExtraTrees, XGBoost, MLPClassifier, LightGBM, HistGradientBoosting to binary and multi-classification. An evaluation of the performance of the various models was then done by comparing them.

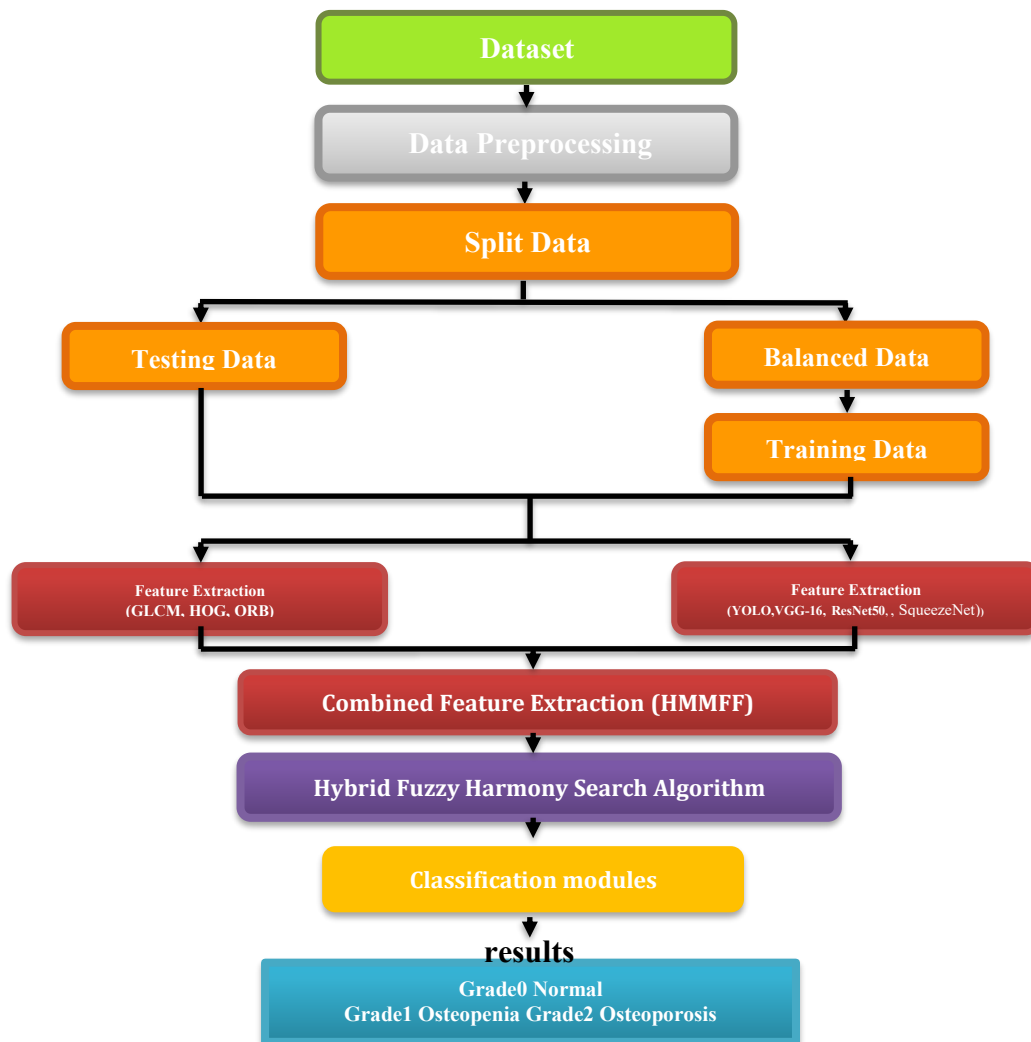


Fig. 5 Flowchart of the proposed method of work.

3.7. Results and Discussion

3.7.1 Results

The performance of the proposed machine learning models in classification was evaluated based on four sets of selected features that are fed into these models for the purpose of classification, which mainly depend on the type of technology used in implementing the feature selection.

A- Performance Results in Binary Classification

Binary classification aims to differentiate between normal and diseased states. The results showed outstanding performance across both the Harmony Search algorithm and HFHSA techniques. As shown in Table 3, HistGradientBoosting demonstrated clear superiority in the Harmony Search algorithm results, achieving an overall classification accuracy of 95.87% with accuracy, sensitivity, and F1 values of 95.98%, 95.85%, and 95.87%, respectively. XGBoost came in second with an accuracy of 95.56% and accuracy, sensitivity, and F1 values of 95.69%, 95.53%, and 95.55%, respectively. ExtraTrees and LightGBM shared third place with an accuracy of 95.24%, followed by RandomForest at 94.92%, and finally MLPClassifier with the lowest at 91.11%.

Table 3. Performance results in binary classification by harmony search algorithm.

Classifier	Accuracy	Precision	Recall	F1
RandomForest	0.949206	0.95126	0.94889	0.949120
ExtraTrees	0.952381	0.95408	0.95210	0.952312
XGBoost	0.955556	0.95693	0.95531	0.955501
MLPClassifier	0.911111	0.91113	0.91116	0.911110
LightGBM	0.952381	0.95408	0.95210	0.952312
HistGradientBoosting	0.958730	0.95981	0.95851	0.958689

The results of the HFHSA binary classification, overall performance improved significantly across most algorithms. HistGradientBoosting topped the list with the highest accuracy of 96.83%, outperforming the standard Harmony Search algorithm by 0.96%, with precision, sensitivity, and F1 values of 96.91%, 96.81%, and 96.82%, respectively. ExtraTrees came in second with 96.51% accuracy and 96.50% F1, while RandomForest, XGBoost, and LightGBM shared third place with an equal accuracy of 96.19%. MLPClassifier ranked last with 92.06% accuracy. This can be explained by the sensitivity of neural networks to the number of input features and the size of the training data. With a limited feature set, the hidden layers lose their ability to learn deep abstract representations, thus reducing their competitive advantage over tree models. Table 4, shows the performance results in the binary classification of features selected using HFHSA.

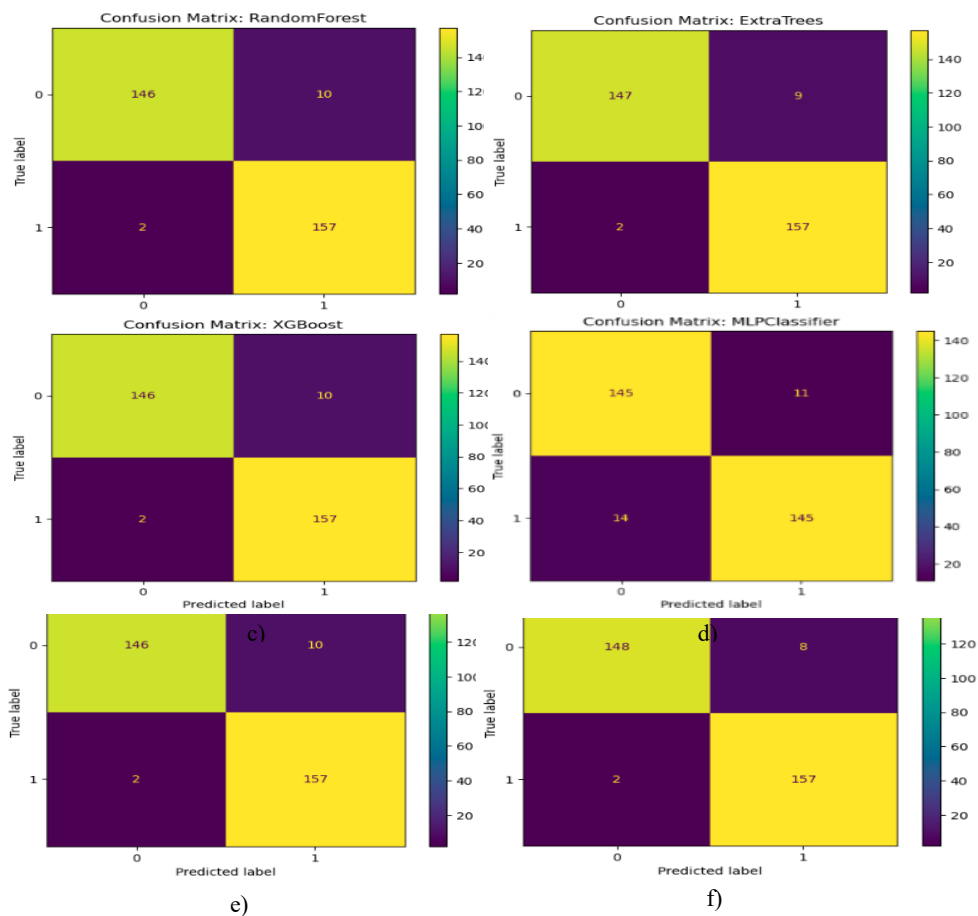
Table 4, Performance results in binary classification using HFHSA technology.

Classifier	Accuracy	Precision	Recall	F1
RandomForest	0.96190	0.963303	0.96165	0.961858
ExtraTrees	0.96507	0.966180	0.96486	0.965044
XGBoost	0.96190	0.963303	0.96165	0.961858
MLPClassifier	0.92063	0.920718	0.92071	0.920635
LightGBM	0.96190	0.963303	0.96165	0.961858
HistGradientBoosting	0.96825	0.969091	0.96807	0.968228

At the level of detailed model performance results for each disease category in binary classification using HFHSA technology, HistGradientBoosting achieved balanced and outstanding performance, with an F1 coefficient of 0.97 for both the normal and disease categories and a recall rate of 95% for normal cases and 99% for the disease category, reflecting its superior ability to detect disease cases. ExtraTrees also demonstrated strong performance with an F1 coefficient of 0.96 for both categories, while RandomForest, XGBoost, and LightGBM shared similar performance values with an F1 coefficient of 0.96 for both categories and a 99% recall rate for the disease category. Table 5, provides a detailed overview of the performance for each disease category, while Figure 6, visually illustrates these results through confusion matrices.

Table 5. Detailed performance results of classification models using HFHSA technology in binary classification.

Proposed model	Disease Class	Precision	Recall	F1-score
Random Forest	normal	0.99	0.94	0.96
	disease	0.94	0.99	0.96
Extra Trees	normal	0.99	0.94	0.96
	disease	0.95	0.99	0.97
XGBoost	normal	0.99	0.94	0.96
	disease	0.94	0.99	0.96
MLP Classifier	normal	0.91	0.93	0.92
	disease	0.93	0.91	0.92
LightGBM	normal	0.99	0.94	0.96
	disease	0.94	0.99	0.96
HistGradient Boosting	normal	0.99	0.95	0.97
	disease	0.95	0.99	0.97

**Fig. 6** Confusion matrices in binary classification of classification models using HFHSA: (a) Random Forest, (b) ExtraTrees, (c) XGBoost, (d) MLP Classifier, (e) LightGBM, (f) HistGradientBoosting.

B- Performance Results in Multi-Classification

The implementation of multi-category classification aims to categorize osteoporosis into three groups: normal, osteopenia, and osteoporosis. However, this presents a more complex challenge as the radiographic boundaries between the groups overlap, particularly between normal and osteopenia in the early transitional stages. As shown in Table 6, the harmony search algorithm exhibited varying results among the models. XGBoost achieved the highest accuracy at 93.59% with an average F1 coefficient of 93.21%. HistGradientBoosting and LightGBM tied for second place with an accuracy of 93.33%, followed by RandomForest and ExtraTrees with an accuracy of 93.08% each. In contrast, MLPClassifier recorded the lowest accuracy at only 80.26%, a sharp decline reflecting the significant difficulty in learning the boundaries between the three groups in the absence of a sufficiently deep representational structure. This indicates its limited ability to handle the complexities of this task.

Table 6. Performance results in multi-classification using the Harmony Search algorithm.

Classifier	Accuracy	Precision	Recall	F1
RandomForest	0.93076	0.937303	0.928744	0.932333
ExtraTrees	0.93076	0.926499	0.924169	0.925243
XGBoost	0.93589	0.931314	0.933058	0.932092
MLPClassifier	0.80256	0.817956	0.828860	0.800945
LightGBM	0.93333	0.929256	0.930961	0.930015
HistGradientBoosting	0.93333	0.933093	0.928694	0.930826

Performance results for models in multi-classification based on feature selection using HFHSA technology showed that XGBoost achieved the highest accuracy of 95.38%, surpassing its performance with the traditional harmony search algorithm by 1.79%, with an average F1 coefficient of 95.16%, representing the best overall performance in multi-classification. RandomForest and HistGradientBoosting came in second with an accuracy of 94.87%, followed by LightGBM with an accuracy of 94.62%, then Extra Trees with an accuracy of 94.4% and an average F1 coefficient of 94.27% with an improvement of 1.51%, while MLPClassifier remained the lowest performing with an accuracy of 80.51%. These significant improvements achieved by applying HFHSA technology confirm that hybridization is more effective in more complex tasks. Table 7. shows the performance results in multi-classification selected by HFHSA technology.

Table 7. Performance results in the multi-classification using HFHSA technology.

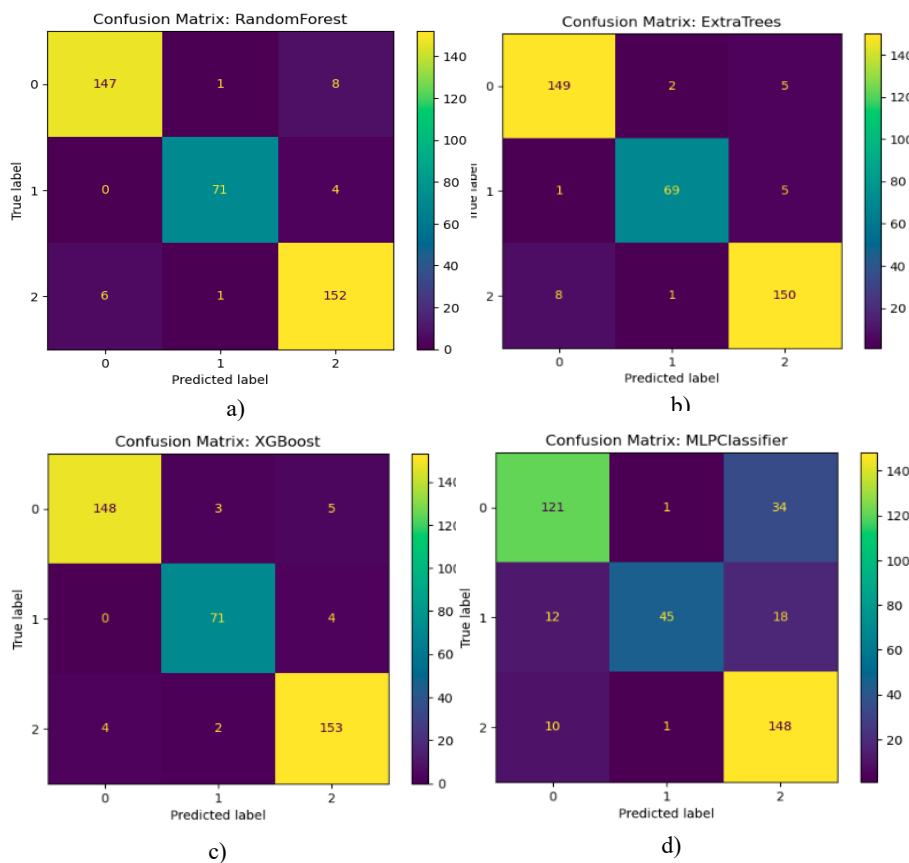
Classifier	Accuracy	Precision	Recall	F1
RandomForest	0.94871	0.95340	0.948316	0.950697
ExtraTrees	0.94359	0.94629	0.939508	0.942753
XGBoost	0.95384	0.95078	0.952550	0.951569
MLPClassifier	0.80512	0.84786	0.768820	0.790527
LightGBM	0.94615	0.94453	0.946180	0.945212
HistGradientBoosting	0.94871	0.94909	0.948276	0.948453

Detailed performance results across disease categories in the multi-classification using HFHSA technology, which provides important insights into the strengths of each model in this classification, showed that the XGBoost model achieved comprehensive and balanced performance across all categories, with an F1 coefficient of 96% for the normal category, 0.94 for the osteopenia category, and 0.95 for the osteoporosis category, with a 96% recall rate for the osteoporosis category, reflecting its good ability to detect advanced disease cases. HistGradientBoosting also demonstrated reliable performance across the three categories, with an F1 coefficient of 0.95 for the normal and

osteopenia categories and 0.94 for the osteoporosis category. RandomForest demonstrated strong performance with an F1 coefficient of 0.95 for the normal category, 0.96 for the osteopenia category, and 0.94 for the osteoporosis category. In contrast, MLPClassifier remained the weakest performer in the osteopenia category with an F1 coefficient of only 0.74. Table 8, details the model performance for each disease category in multi-classification, and Figure 7, shows the confusion matrices in multi-classification using HFHSA technology. Figure 8, compares the performance of machine learning models in binary and multi-classification. while Table 9 shows a comparison between the results of the proposed method and the results of similar research in the same field.

Table 8. Detailed performance results of classification models using HFHSA technology in multi-classification.

Proposed model	Disease Class	Precision	Recall	F1-score
Random Forest	Normal	0.96	0.94	0.95
	Osteopenia	0.97	0.95	0.96
	Osteoporosis	0.93	0.96	0.94
Extra Trees	Normal	0.94	0.96	0.95
	Osteopenia	0.96	0.92	0.94
	Osteoporosis	0.94	0.94	0.94
XGBoost	Normal	0.97	0.95	0.96
	Osteopenia	0.93	0.95	0.94
	Osteoporosis	0.94	0.96	0.95
MLP Classifier	Normal	0.85	0.78	0.81
	Osteopenia	0.96	0.60	0.74
	Osteoporosis	0.74	0.93	0.82
LightGBM	Normal	0.97	0.94	0.95
	Osteopenia	0.93	0.95	0.94
	Osteoporosis	0.93	0.96	0.94
HistGradientBoosting	Normal	0.97	0.94	0.95
	Osteopenia	0.95	0.95	0.95
	Osteoporosis	0.93	0.96	0.94



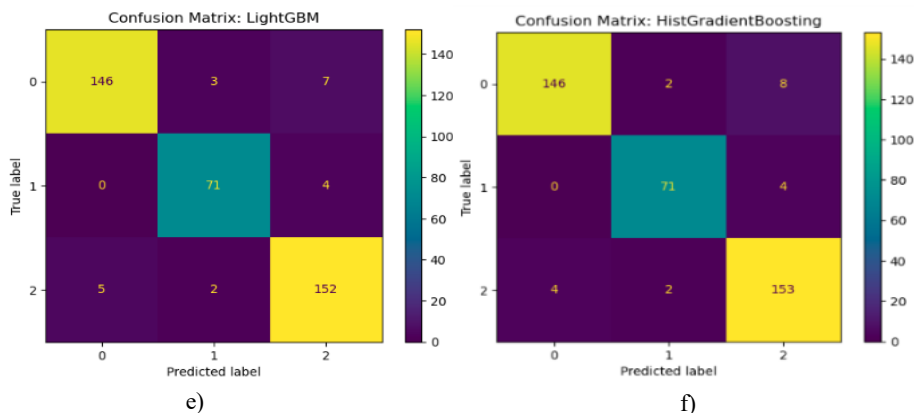


Fig. 7 Confusion matrices for classification models using HFHSA technology in multi-classification: (a) Random Forest, (b) ExtraTrees, (c) XGBoost, (d) MLP Classifier, (e) LightGBM, (f) HistGradientBoosting.

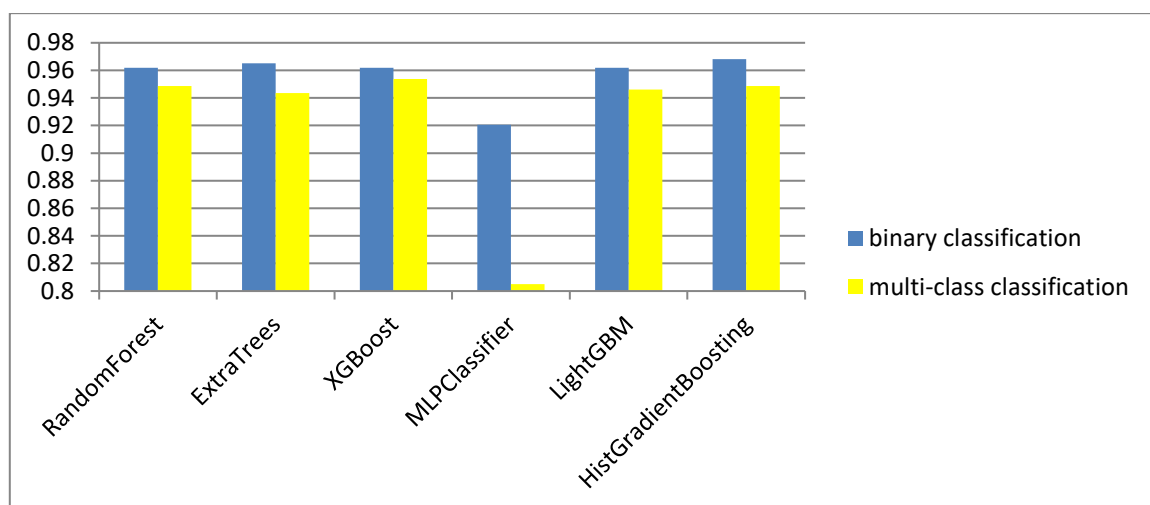


Fig. 8 compares the performance of machine learning models using HFHSA techniques.

Table 9. Comparison between the results of the proposed method and the results of related works.

Authors	Year	Imaging Modality	Human Site	Methods	Dataset	Result
Yang, Tsai Shih	2022	X-ray images	knee	VGG16, Late-Fusion	1650 images	VGG16 ACC: 82% Late-Fusion ACC: 77%
Wani, I. M., & Arora, S.	2023	X-ray images	knee	Transfer learning is utilized with CNN architectures like AlexNet, VggNet-16, ResNet, and VggNet-19	381 images	AlexNet ACC: 91.1%
Dodamani, Pooja S., and Ajit Danti	2023	X-ray images	knee	compares various CNN models VGG-16 ,	830 images	DenseNet121 ACC: 93.4%

				VGG-19 ,DenseNet121 ,Resnet50 ,InceptionV3 VGG16, VGG19, Inception V3, Xception, DenseNet169, DenseNet201, ResNet101, and ResNet152	372 images	ACC: 86.36
Kumar et al	2023	X-ray	knee	Inception v3 ,Xception ,ResNet 18	240 images	accuracy rate of 93.5%
Kumar et al.	2023	X-ray images	knee	VGG19	1400 images	ACC: 89%
Shen et al.	2024	X-ray images	knee	superfluity and Transfer learning AlexNet and ResNet50 DenseNet169 and Vision Transformer (ViT)	240 images	Superfluity DL model ACC: 85.42% for dataset1 and ACC: 79.39% for dataset2
Naguib et al.	2024	X-ray	-- knee			
Muhammad et al.	2025	X-ray			744 images	accuracy of 0.8611,
OM, A., & Gunasundari, R.	2025	X-ray	knee	ResNet-50, Gated Recurrent Units (GRUs) VGG- 16,YOLOv11, ResNet-50, Squeezenet GLCM,HOG,ORB, harmony search, Gath-Geva Random Forest, Extra Trees, XGBoost, MLPClassifier, LightGBM, HistGradientBoo sting	389 images	ACC: 95.65%
In this research	2026	X-ray	knee		1947 image	Binary classification ACC: 96.8% and Multi- classification ACC: 95.38%

The results achieved in the research demonstrated the efficiency of the proposed system in classifying knee X-ray images with high accuracy compared to previous work mentioned in Table 9. The proposed model achieved an exceptional accuracy of 96.8% , 95.38%, in binary and multi-classification respectively outperforming all previous studies mentioned. The comparison demonstrated that the proposed system was evaluated on a large dataset of 1,947 images, the largest dataset used in previous studies. This increases the reliability and generalizability of the results compared to some previous studies that relied on smaller datasets. It surpassed recent studies such as Dodamani and Danti (2023), who achieved 93.4% using DenseNet121 on 830 images, and Kumar et al. (2023), who achieved 93.5% using Inception v3 on only 240 images.

The clear superiority of the proposed system is due to several innovative technical factors, including the adoption of the HMMFF method to combine features extracted by traditional techniques with higher-level features for deep learning which helps provide a comprehensive and effective representation of bone features in X-ray images. The

use of the hybrid feature selection technique (HFHSA) combines the power of the consensus search algorithm in global exploration with the flexibility of fuzzy logic in dealing with uncertainty, ensuring the selection of the best set of features and removing redundant features that may negatively impact system performance.

3.7.2 Discussion

The findings show the combination of the Harmony Search algorithm with machine learning has achieved excellent results in the classification of knee osteoporosis. The application of the hybrid HFHSA technique showed great improvement in the performance of the system when compared with the original harmony search algorithm across most classification scenarios .

For the binary classification problem, the HistGradientBoosting model has the highest accuracy for both techniques, with a 0.95% improvement of switching from the traditional Harmony Search algorithm to the HFHSA technique. XGBoost achieved (from 95.5% to 96.2%) an excellent gain in a medical application where each percentage point can make the difference between life and death, or between a misdiagnosis and a correct diagnosis. This might not sound much, but it is in the medical context. An increase in percentage points in the accuracy of a diagnosis means fewer misdiagnoses (false positives or false negatives), which in turn leads to better patient outcomes.

When it came to multi-classification, these improvements were even greater. XGBoost saw a 1.79% improvement (from 93.59% to 95.38%), RandomForest's accuracy rose from 93.08% to 94.87% (also a 1.79% improvement), and ExtraTrees improved from 93.08% to 94.36% (a 1.28% increase). This improvement in all methods highlights the success of hybridization in more complex problems, where varying feature selection allows the model to learn more accurate and comprehensive data patterns. Additionally, all ensemble algorithms (RandomForest, ExtraTrees, XGBoost, LightGBM and HistGradientBoosting) tested outperformed the MLPClassifier model with both techniques. This can be attributed to the underlying principles of ensemble algorithms. They use multiple sub-models (trees) and combine their predictions to reduce model variance by voting or averaging, thus increasing model generalisation. Each tree may capture different patterns in the data and when these patterns are averaged, random noise is reduced and true patterns are reinforced.

Another important finding from our results is the difficulty in classifying the osteopenia class. The MLPClassifier model showed a low F1 coefficient of 0.74 for this class, the lowest of all models and classes. This can be attributed to the visual resemblance of the radiographic features in the transitional state (osteopenia) to the healthy and severely injured states, even for expert doctors. At the beginning, the anatomical modifications may not be sufficiently clear, with some normal knees resembling early pathological stage.

4. Conclusion

It has been noted that the detection of knee osteoporosis has gained importance over time due to its related complications especially with the emergence of artificial intelligence. The hybrid HFHSA and hybrid feature fusion (HMMFF) methods have made it possible to detect osteoporosis using X-ray images and these techniques have become the promising and effective technologies to enhance the accuracy and efficiency of detecting osteoporosis by the use of AI systems.

This research shows that integrating features obtained in various methods is valuable and efficient. Such integration provides an opportunity to use the diversity of patterns and structural information in images producing a complete and informative feature-vector. Moreover, it boosts the feature selection with the HFHSA technology that uses the efficient search of the Harmony Search algorithm and the capability of the GathGeva algorithm to deal with overlapping data points. This leads to the high classification accuracy which allows the accurate identification of the disease and classification of the affected area. The proposed system is an efficient tool to assist clinicians in making better and faster decisions and help to achieve better patient outcomes and lower treatment expenses. It should be pointed out though that these systems are not intended to take the place of the clinicians they are meant to aid them. The role of human clinical experience, medical judgment and capacity to synthesize various information about patients is also critical in the diagnostic and treatment process.

It is also recommended to continue developing osteoporosis detection techniques and to test the model on larger and more diverse datasets.

References

- [1] S. L. Kulkarni and H. Kour, "Narrative Review on Osteoporosis: A Silent Killer," *Journal of Clinical & Diagnostic Research*, vol. 18(4), (2024). <https://doi.org/10.7860/JCDR/2024/69058.19248>
- [2] P. Falaschi, A. Marques, and S. Giordano, "Osteoporosis and fragility in elderly patients," *Orthogeriatrics: The Management of Older Patients with Fragility Fractures*, (2021), pp. 35-52. https://doi.org/10.1007/978-3-030-48126-1_3
- [3] M. S. LeBoff, S. L. Greenspan, K. L. Insogna, E. M. Lewiecki, K. G. Saag, A. J. Singer, and E. S. Siris, "The clinician's guide to prevention and treatment of osteoporosis," *Osteoporosis international*, vol. 33(10), (2022), pp. 2049-2102.
- [4] A. M. Sarhan, M. Gobara, S. Yasser, Z. Elsayed, G. Sherif, N. Moataz, Y. Yasir, E. Moustafa, S. Ibrahim, and H. A. Ali, "Knee osteoporosis diagnosis based on deep learning," *International Journal of Computational Intelligence Systems*, vol. 17(1), (2024), p. 241.
- [5] S. S. Enitan, B. I. G. Adejumo, E. N. Adejumo, T. O. Olusanya, E. O. Osakue, O. A. Ladipo, and C. B. Enitan, "Empowering women to combat osteoporosis: Unveiling the causes, consequences, and control strategies," *Aging Commun*, vol. 5(3), (2023), p. 13.
- [6] N. Chaiyavech, S. Thiengwittayaporn, and N. Hongku, "Prevalence of Common Metabolic Bone Diseases Diagnosed by Dual-Energy X-Ray Absorptiometry Scanning and Blood Test in Outpatients With Osteoarthritis the Knee," *Geriatric Orthopaedic Surgery & Rehabilitation*, vol. 15, (2024).
- [7] V. Duong, C. Shaheed, M. Ferreira, S. Narayan, V. Venkatesha, D. Hunter, et al., "Risk factors for the development of knee osteoarthritis across the lifespan: a systematic review and meta-analysis," *Osteoarthritis and cartilage*, (2025).
- [8] M. Zamzam, M. Alamri, F. Aldarsouni, H. Zaid, and A. Ofair, "Impact of Osteoporosis in Postmenopausal Women With Primary Knee Osteoarthritis," *Cureus*, vol. 15, (2023).
- [9] M. Qureshi, M. Sani, A. Raza, M. Qureshi, W. Alghamdi, S. Akbar, R. Kadir, and M. Sarker, "Deep learning based osteoporosis classification in knee X rays using transfer learning approach," *Scientific Reports*, vol. 15, (2025).
- [10] C. Hsieh, K. Zheng, C. Lin, L. Mei, L. Lu, W. Li, F. Chen, et al., "Automated bone mineral density prediction and fracture risk assessment using plain radiographs via deep learning," *Nature Communications*, vol. 12, (2021).
- [11] R. Gaudin, W. Otto, I. Ghanad, G. Kewenig, C. Rendenbach, V. Alevizakos, P. Grün, et al., "Enhanced Osteoporosis Detection Using Artificial Intelligence: A Deep Learning Approach to Panoramic Radiographs with an Emphasis on the Mental Foramen," *Medical Sciences*, vol. 12, (2024).
- [12] T. S. Yang, "Recognition and classification of knee osteoporosis and osteoarthritis severity using deep learning techniques," *Dublin, National College of Ireland*, (2023).
- [13] I. M. Wani and S. Arora, "Osteoporosis diagnosis in knee X-rays by transfer learning based on convolution neural network," *Multimedia Tools and Applications*, vol. 82(9), (2023), pp. 14193-14217.
- [14] P. S. Dodamani and A. Danti, "Transfer learning-based osteoporosis classification using simple radiographs," *IJOE*, vol. 19(08), (2023), p. 67.
- [15] S. Kumar, P. Goswami, and S. Batra, "Enriched diagnosis of osteoporosis using deep learning models," *International Journal of Performability Engineering*, vol. 19(12), (2023), p. 824.
- [16] S. Kumar, P. Goswami, and S. Batra, "Fuzzy rank-based ensemble model for accurate diagnosis of osteoporosis in knee radiographs," *International Journal of Advanced Computer Science and Applications*, vol. 14(4), (2023), pp. 262-270.
- [17] M. Shen, "Utilizing Deep Learning for Osteoporosis Diagnosis through Knee X-Ray Analysis," *2024 International Conference on Artificial Intelligence and Communication (ICAIC 2024)*, (2024), pp. 553-560.
- [18] S. M. Naguib, M. K. Saleh, H. M. Hamza, K. M. Hosny, and M. A. Kassem, "A new superfluity deep learning model for detecting knee osteoporosis and osteopenia in X-ray images," *Scientific Reports*, vol. 14(1), (2024).
- [19] R. Muhammad, S. Rao, and B. Lee, "BONE-Net: A novel hybrid deep-learning model for effective osteoporosis detection," *PLoS One*, vol. 20(10), (2025), p. e0334664.
- [20] A. OM and R. Gunasundari, "Advanced Osteoporosis Prediction from Knee X-rays via Residual Convolution and Recurrent Networks," *International Journal of Intelligent Engineering & Systems*, vol. 18(6), (2025).
- [21] <https://www.kaggle.com/datasets/mohamedgobara/multi-class-knee-osteoporosis-x-ray-dataset>
- [22] K. U. Ahamed, M. Islam, A. Uddin, A. Akhter, B. K. Paul, M. A. Yousuf, S. Uddin, et al., "A deep learning approach using effective preprocessing techniques to detect COVID-19 from chest CT-scan and X-ray images," *Computers in Biology and Medicine*, vol. 139, (2021), p. 105014.
- [23] Y. Zhang, M. Jin, and G. Huang, "Medical image fusion based on improved multi-scale morphology gradient-weighted local energy and visual saliency map," *Biomedical Signal Processing and Control*, vol. 74, (2022), p. 103535.
- [24] T. Zhou, X. Ye, H. Lu, X. Zheng, S. Qiu, and Y. Liu, "Dense convolutional network and its application in medical image analysis," *BioMed Research International*, vol. 2022(1), (2022), p. 2384830.
- [25] N. Manakitsa, G. S. Maraslidis, L. Moysis, and G. F. Fragulis, "A review of machine learning and deep learning for object detection, semantic segmentation, and human action recognition in machine and robotic vision," *Technologies*, vol. 12(2), (2024), p. 15.
- [26] K. A. Patil, K. M. Prashanth, and A. Ramalingaiah, "Classification of osteoporosis in the lumbar vertebrae using L2 regularized neural network based on PHOG features," *International Journal of Advanced Computer Science and Applications*, vol. 13(4), (2022).
- [27] A. Behura, "The cluster analysis and feature selection: Perspective of machine learning and image processing," *Data Analytics in Bioinformatics: A Machine Learning Perspective*, (2021), pp. 249-280.
- [28] Y. Bo, G. Chen, L. Li, X. Tao, and R. Zhao, "Detection of osteoporosis using image processing methods," *Journal of Optics*, vol. 53(4), (2024), pp. 2898-2908.
- [29] A. Oad, K. Kumari, I. Hussain, F. Dong, B. Hammad, and R. Oad, "Performance comparison of ORB, SURF and SIFT using Intracranial Haemorrhage CTScan Brain images," *International Journal of Artificial Intelligence & Mathematical Sciences*, vol. 1(2), (2022), pp. 26-34.
- [30] P. Krakowski, A. Rejniak, J. Sobczyk, and R. Karpiński, "Cartilage integrity: A review of mechanical and frictional properties and repair approaches in osteoarthritis," *Healthcare*, vol. 12(16), (2024), p. 1648.
- [31] K. Wu, "Creating panoramic images using ORB feature detection and RANSAC-based image alignment," *Advances in Computer and Communication*, vol. 4(4), (2023), pp. 220-224.
- [32] D. Theng and K. K. Bhoyar, "Feature selection techniques for machine learning: a survey of more than two decades of research," *Knowledge and Information Systems*, vol. 66(3), (2024), pp. 1575-1637.
- [33] N. Pudjihartono, T. Fadason, A. W. Kempa-Liehr, and J. M. O'Sullivan, "A review of feature selection methods for machine learning-based disease risk prediction," *Frontiers in Bioinformatics*, vol. 2, (2022), p. 927312.
- [34] M. Nasir, A. Sadollah, P. Grzegorzewski, J. H. Yoon, and Z. W. Geem, "Harmony search algorithm and fuzzy logic theory: An extensive review from theory to applications," *Mathematics*, vol. 9(21), (2021), p. 2665.
- [35] A. S. Balleh, H.-C. Soong, S. K. G. Singh, and N. A. Jalil, "Automated DJ Pad Audio Mashups Playback Compositions in Computer Music Utilizing Harmony Search Algorithm," *2021 IEEE 19th Student Conference on Research and Development (SCORED)*, (2021), pp. 388-393.
- [36] M. Dubey, V. Kumar, M. Kaur, and T.-P. Dao, "A systematic review on harmony search algorithm: theory, literature, and applications," *Mathematical Problems in Engineering*, vol. 2021(1), (2021), p. 5594267.

- [37] E. Uray, S. Carbas, Z. W. Geem, and S. Kim, "Parameters optimization of taguchi method integrated hybrid harmony search algorithm for engineering design problems," *Mathematics*, vol. 10(3), (2022), p. 327.
- [38] C. Huang, "Feature Selection and Feature Stability Measurement Method for High-Dimensional Small Sample Data Based on Big Data Technology," *Computational Intelligence and Neuroscience*, vol. 2021(1), (2021), p. 3597051.
- [39] N. G. Peso'a and N. F. Gamayanti, "Implementation of the Gath-Geva Clustering Algorithm in the Clustering Districts/Cities in Central Sulawesi Based on Public Health Development Indicators," 4th International Seminar on Science and Technology (ISST 2022), (2023), pp. 320-328.
- [40] X. Wu, H. Zhou, B. Wu, and T. Zhang, "A possibilistic fuzzy Gath-Geva clustering algorithm using the exponential distance," *Expert Systems with Applications*, vol. 184, (2021), p. 115550.
- [41] J. H. Cabot and E. G. Ross, "Evaluating prediction model performance," *Surgery*, vol. 174(3), (2023), pp. 723-726.
- [42] S. Sathyanarayanan and B. R. Tantri, "Confusion matrix-based performance evaluation metrics," *African Journal of Biomedical Research*, vol. 27(4S), (2024), pp. 4023-4031.