

ISSN (Print) 2794-7629

ISSN (Online) 2794- 4549

Received 24/04/2026

Accepted 15/06/2026

FULL PAPER**Hybrid Framework between Bobcat Optimization and CNN for Osteoporosis Detection*****Prepared by******Ahmed Khalid Abdullah***

Directorate of Ninawa Health
Ministry of Health, Mosul, Iraq

Raid Rafi Omar Al-Nima

Technical Engineering College
for Computer and Artificial
Intelligence/Northern Technical
University
Mosul, Iraq
Artificial Intelligence Research
Center, Northern Technical
University, Mosul, Iraq

Khalid Ghanim Majeed

Department of Medical
Physiology
College of Medicine
Nineveh University, Mosul,
Iraq

Ahmed Isam Abdulrhman

Directorate of Ninawa Health
Ministry of Health, Mosul, Iraq

Marwn Fawzi al-Bayati

Electrical engineering College
Middle Technical University
Bagdad, Iraq

Abstract

Osteoporosis is characterized by diminished bone mass and bone tissue loss, which results in weakening bones, decreasing bone strength and increasing risk of fractures. This paper exploits a Medical Lumber Spine Images (MLSI) dataset of Dual-Energy X-ray Absorptiometry (DEXA) in a clinic in Mosul / Iraq in order to be classified as either normal or having osteoporosis. Here, it is decided to explore use of Bobcat Optimization Algorithm (BOA) for selection of best Convolutional Neural Network (CNN) hyperparameters. The proposed method is called Bobcat Optimization for CNN (BOCNN) model, which systematically explores and optimizes 18 hyper-parameters. It combines the effectiveness of BOA for optimizing effective CNN hyperparameters which can further reduce the manual tuning efforts. This study uses the real clinical data set from the MLSI. The proposed model is able to diagnose Osteoporosis correctly in 95% of the cases from unseen data. The Area Under Curve (AUC) of Receiver Operating Characteristic (ROC) equals to 0.98.

Keywords: Osteoporosis, Deep learning, Bobcat Optimization Algorithm, Convolutional Neural Network

1. Introduction

Osteoporosis is a skeletal disorder characterized by reduced bone strength, leading to an increased risk of fractures. Bone strength is determined by multiple factors, including Bone Mineral Density (BMD), bone geometry, mineralization, microarchitecture, and bone turnover rate [1].

Hormone balancing and healthy lifestyle and taking into account internal and external factors are the best ways to achieve the best bone density and prevent fractures. Osteoporosis is a common disease with important clinical implications [2]. Figure 1 illustrates the contrast between healthy bones and osteoporotic bones.

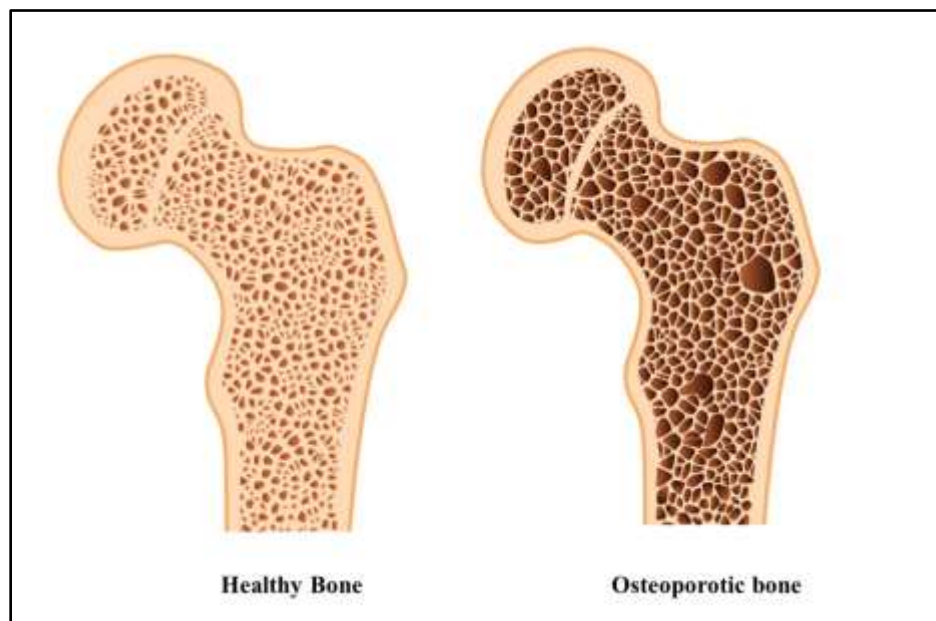


Figure 1: Deference between healthy and osteoporotic bones

DEXA is thought to be the best tool for finding osteoporosis, and it is known around the world as the gold standard for measuring BMD. DEXA measures the density of minerals in bones to find osteoporosis, check the risk of fractures, and keep track of changes over time. Figure 2 illustrates axial bone densitometry of the lumbar spine and hip, a method frequently employed in clinical practice [3].



Figure 2: Full body DEXA scan

Recent research has focused on the utilization of Artificial Intelligence (AI) in the interpretation of medical images, including the creation of a tool for osteoporosis screening. Notwithstanding various efforts to tackle osteoporosis utilizing ML and DL methods, problems persist [4].

This paper aims to optimize a deep learning model for the diagnosis of osteoporosis using real clinical and medical imaging DEXA data. This work proposes the BOCNN model, which utilize BOA to automatically tune eighteen CNN hyperparameters in order to get optimal performance. The subsequent sections will be organized as follows: related literature review is provided in Section 2, the proposed BOCNN model is elucidated in Section 3, findings and discussions are considered in Section 4 and the conclusion is provided in Section 5.

2. Related literature review

In 2019, Derkatch *et al.* proposed the use of machine learning techniques to develop an automated system for the identification and classification of osteoporosis in bone X-ray radiographss. Images are pre-processed using Contrast Limited Adaptive Histogram Equalisation (CLAHE) . Grey Level Co-occurrence Matrix (GLCM) and Principal Component Analysis (PCA) are used for feature extraction and Support Vector Machines (SVM) and Nearest Neighbours (NN) are used to classify the images. The collection consists of 58 radiographic images of osteoporotic and non-osteoporotic groups, respectively. The results demonstrate that the proposed method can successfully discriminate between healthy and osteoporotic subjects with 95% accuracy of SVM classifiers. [5].

In 2020, Ho *et al.* aimed to determine whether DL models could accurately estimate bone BMD and predict osteoporosis from conventional radiographs using DEXA as the clinical reference standard. The researchers collected DEXA images with measurements of the hip and lumbar spine, preprocessed the images by cropping and intensity normalization, and developed a CNN termed DeepDXA for both regression and classification tasks. The model acquired the ability to forecast continuous BMD values and to apply DEXA T-score thresholds to determine the presence of osteoporosis in a patient. We employed ROC metrics for classification and correlation analysis for regression to evaluate the model's performance. The proposed DeepDXA had a Pearson correlation coefficient of 0.88 with BMD measured by DEXA. The model achieved an AUC of 0.94, an accuracy of 90.3%, a sensitivity of 89.2%, and a specificity of 91.1% in the classification of osteoporosis. The results indicate that the CNN can accurately replicate DXA-derived bone density and function effectively as a screening tool utilizing solely X-ray images [6].

In 2022, Zhang *et al.* Performed a multicenter retrospective study to create and evaluate a Deep Convolutional Neural Network (DCNN) for the screening of osteopenia and osteoporosis utilizing lumbar spine radiographs, with DEXA T-scores as the clinical reference standard. Images were subjected to standardized preprocessing, which included contrast normalization and extraction of the vertebral region-of-interest (ROI), followed by data augmentation to rectify class imbalance. Multiple CNN architectures, including ResNet-50 and DenseNet-121, were trained and assessed utilizing five-fold cross-validation. The evaluation of model performance was conducted by sensitivity, specificity, and the AUC in comparison to corresponding DEXA labels. The optimal model attained an AUC of 0.93 and an accuracy of 89.7%, indicating that DCNNs trained on radiographs may effectively detect persons with low bone mineral density, as validated by DEXA. [7].

Hybrid Framework between

In the same year, Jang *et al.* Created a deep learning model that predicts osteoporosis directly from regular hip radiographs, use proximal femur DEXA data as the reference standard. The research compiled a dataset of hip radiographs alongside associated DEXA derived BMD and T-score labels. Each image was cropped to separate the femoral head and neck areas, thereafter undergoing histogram normalization and geometric augmentation to enhance model robustness. Multiple pre-trained CNN architectures, such as ResNet-50, InceptionV3, and DenseNet-121, were optimized by transfer learning. Model evaluation was conducted utilizing ROC analysis, accuracy, sensitivity, and specificity in relation to DEXA derived osteoporosis classifications. The optimal model, utilizing a fine-tuned DenseNet-121, attained an AUC of 0.95 and an accuracy of 91.2. [8].

In 2023, Pan *et al.* Executed a pioneering study designed to facilitate opportunistic osteoporosis screening by amalgamating imaging and clinical data via a DCNN architecture. The primary objective was to investigate if the integration of radiographic image features with clinical factors may enhance diagnostic accuracy in comparison to image-only models when evaluated against DEXA standards. The process entailed the automatic extraction of spinal ROI from standard chest or opportunistic CT scans, succeeded by feature encoding utilizing CNN architectures like ResNet and DenseNet. The collected picture features were subsequently integrated with patient-level clinical factors (e.g., age, sex, BMI, and medical history) using a late-fusion classifier head, resulting in a multi-feature prediction model. The evaluation of model performance utilized essential parameters such as AUC, sensitivity, and specificity in relation to DEXA defined osteoporosis state. The findings indicated that the fusion model markedly surpassed imaging-only baselines, with an AUC of 0.88 and a sensitivity of 85%, in contrast to 0.81 and 77% for image-only networks. [9].

In 2024, Dhanagopal et al. Intended to enhance automated osteoporosis detection from DEXA images in 2024 by the utilization of sophisticated deep learning methodologies. They employed transfer learning with a channel-boosted CNN to extract additional features from bone areas while utilizing reduced training data. The model underwent training and evaluation using DEXA datasets, achieving a sensitivity of 0.88, a specificity of 0.90, and an overall accuracy of 89%. This surpassed both baseline machine learning and conventional CNN techniques. The results indicate that integrating transfer learning with channel-boosting significantly enhances AI's capability to detect osteoporosis in DEXA images [10].

In 2024, Siddiqua A. Designed to provide a resilient computer-aided system for the automatic diagnosis of osteoporosis exploiting DEXA pictures through a transfer-learning framework integrated with

Hybrid Framework between

a U-Net–based segmentation architecture. The process commenced with the segmentation of bone regions of interest from the DEXA scans via the U-Net network, facilitating accurate extraction of pertinent structural information while minimizing background noise and extraneous anatomical regions. The segmented features were subsequently input into a transfer-learning classifier to assess BMD and T-scores, enabling precise categorization of normal, osteopenia, and osteoporotic conditions. The model was trained and validated using gathered DEXA datasets, attaining a classification accuracy of 90%, alongside strong sensitivity and specificity metrics, demonstrating dependable performance across various patient samples. The results underscore the efficacy of combining segmentation-driven feature extraction with transfer learning, illustrating that these AI pipelines can deliver swift, precise, and clinically relevant evaluations of bone health from DEXA images, thereby aiding radiologists in diagnostic decision-making and enhancing osteoporosis screening processes [11].

In 2025, Abdullah *et al.* proposed a hybrid framework that integrates Genetic Algorithm (GA) optimization with CNN for osteoporosis diagnosis. The study addresses the limitations of conventional deep learning models by introducing an adaptive optimization mechanism for hyperparameter tuning and network configuration. The results demonstrated improved diagnostic performance, emphasizing the effectiveness of combining metaheuristic optimization techniques with deep learning models in medical imaging applications [12]

No published attempts were found where DEXA imaging datasets were used for osteoporosis diagnosis. Besides, no application of BOA exists to optimise the hyperparameters of CNN. Such gaps are filled by this study.

3. Methodology

3.1 Proposed Approach

The proposed BOCNN model for classification of MLSI dataset into normal and osteoporotic classes. The process consists of a CNN model with a given architecture and the BOA for optimising hyperparameters such as the number of filters, convolutional kernel sizes, convolutional kernel strides, pooling block sizes, pooling block strides, number of dense neurones, and dropout rates. Properly tuning

Hybrid Framework between

hyperparameters improves accuracy and reduces overfitting and improves the generalisation of the model [13][14]. Figure 3 illustrate the whole framework of the suggested method.

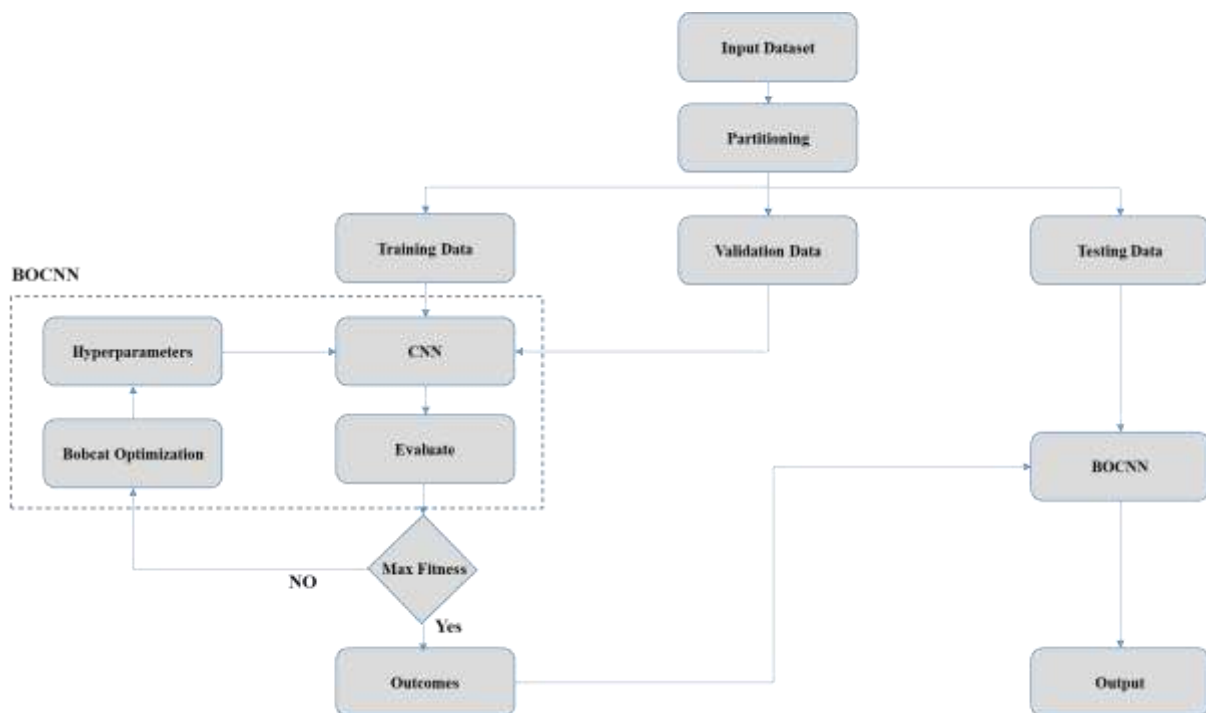


Figure 3: Block diagram of the proposed work

“The dataset of images runs as the input. Subsequently, it is partitioned into training, validation and testing groups of 60%, 20% and 20%, respectively. Consequently, we construct a CNN and initialize its hyperparameters. The BOA is employed to fine-tune hyperparameters and achieve the optimal configuration. We refer to this as the BOCNN, which subsequently tested to acquire the assessed accuracy”.

3.2 CNN Fundamentals

The input layer of a CNN receives images. Each image is of the colored type (Red, Green, and Blue (RGB)). The inputs consist of colored images, each including three RGB channels.

“The convolutional layer constitutes the essential physical mechanism in the CNN. Two-Dimensional (2D) convolution procedures must be performed for each input channel. This layer consists of a collection of image channels referred to as feature maps (or filters). The convolutional layer is crucial

Hybrid Framework between

for extracting various features from a 2D channel [15]. The process involves convolving a kernel with a channel of image. The convolution function is represented by the following equation:

$$Z(x, y) = \sum_{m=1}^M \sum_{n=1}^N R(m, n) \times W(x - m, y - n) \quad (1)$$

where $Z(x, y)$ symbolizes a computed convolution, M and N respectively refer to the width and height of the 2D input channel, $R(m, n)$ represents the input channel, and $W(x - m, y - n)$ is weight value [16].

The multiple layers also use ReLU as it is a well-known transfer function in the DL. The ReLU transfer function maintains the positive values and discards the negative values of any preceding channel (feature map). Therefore, it allows non-linear calculations. The ReLU activation function is given by the following equation:

$$ReLU(x, y) = \max(0, (x, y)) \quad (2)$$

where $ReLU(x, y)$ represents a computed ReLU, and $\max$ is the maximum operation [16].

“The pooling layer reduces the sizes of received channels, resulting in small sized feature maps. It is typically implemented in one of two types: max pooling or average pooling. Its advantage is reducing the spatial dimensions of feature maps while preserving the most prominent characteristics. The computation of maximum pooling is executed in accordance with the below equation:

$$K(i, j) = \max(\mathbf{V}) \quad (3)$$

where $K(i, j)$ is a computed pooling and $\mathbf{V}$ is a small matrix (block) in a previous channel [15].

The Fully connected layer, also known as dense layer, is carried out after the feature extraction part. It comprises of many neurons, every neuron in this layer is connected to all neurons in a previous layer. The number of neurons in FC layer can be equal to the number of required outputs of classes (C). The essential equation in FC layer can be illustrated as follows”:

$$FC(c) = \sum_{l=1}^p (FW(c, l) \times P(l)) + B(c) \quad (4)$$

where $FC(c)$ mean the calculated value in the fully connected layer, c illustrate the neuron's number in the current layer, P shows the number of neurons in the previous layer, $FW(c, l)$ represents a connection weight value between the previous and current layers and $B(c)$ represents a bias value [16].

A softmax layer provides the related probabilities for each output class accordingly to the provided input, such probabilities sum up to one. This layer basically uses the softmax activation function. The general equation in the softmax layer can be given as follows:

$$S(c) = \frac{\exp(FC(c))}{\sum_{l=1}^C \exp(FC(l))} \quad (5)$$

where $S(c)$ shows the calculated softmax value and $\exp$ is the exponential [17].

3.2 BOA Fundamentals

“The BOA is a recently established nature-inspired metaheuristic algorithm, proposed in 2024. It mimics the distinctive hunting technique of bobcats in their natural habitat. It aims to tackle intricate optimization challenges by efficiently balancing between the exploration and exploitation. The two phases of this algorithm make it unique. These are the exploration and exploitation phases. The exploration phase mimics the bobcat's strategy in stealthily tracking a prey over a wide area. The exploitation phase represents a precise and aggressive chase to catch the prey. This bio-inspired dual phases mechanism allows BOA to search globally while efficiently and rapidly converging toward the optimal solution [18]”.

“In our study we utilize the BOA to optimize the hyperparameters of CNN to achieve a maximum performance. The algorithm initiates with a set of candidate solutions (bobcats) that are randomly scattered throughout the searching space, each one of them represents a potential solution to the optimization problem. The fitness of each candidate is evaluated iteratively, increasingly guiding the population toward the optimal region of searching space. As mentioned, the algorithm process is structured into two main phases that reflects the natural hunting behavior of bobcats [18]. In more details”:

The exploration Phase (Tracking the Prey): In this phase, bobcats survey the searching area broadly, simulating to their behavior when tracking a prey. The algorithm promotes diversity by allowing bobcats' candidates to move towards randomly selected positions. This prevents premature convergence and ensures comprehensive exploration of the solution space. This phase is essential for pinpointing the prospective area inside the search space [18].

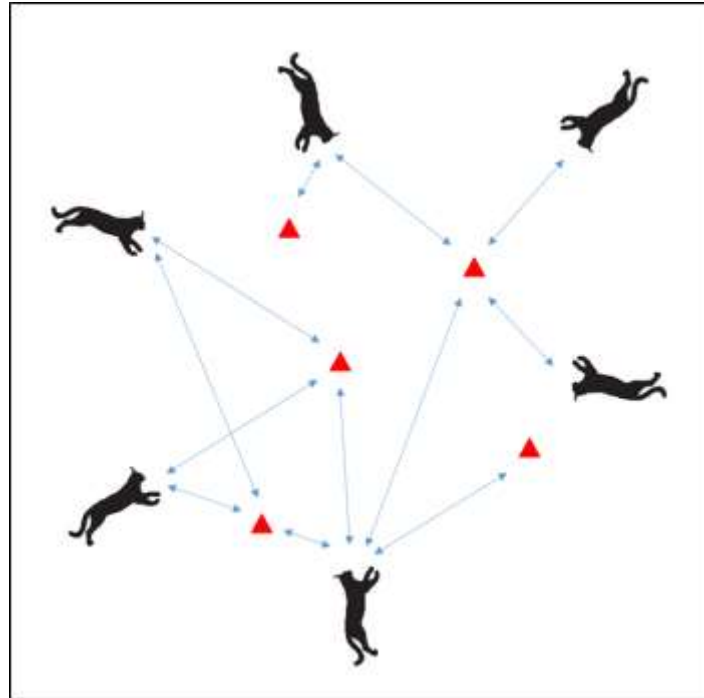


Figure 4: Conceptual depiction of the BOA exploration phase, wherein bobcats traverse randomly towards several candidate solutions, hence providing diversity throughout the search space.

The Exploitation Phase (Chasing the Prey): Upon identifying the promising area, the algorithm transitions to this phase, designated for pursuit behavior. In this case, bobcats change their positions based on the best solutions they've got so far, making their search more focused on finding the best solution. This targeted enhancement guarantees that the algorithm efficiently advances candidate solutions and approaches the optimal solution [18].

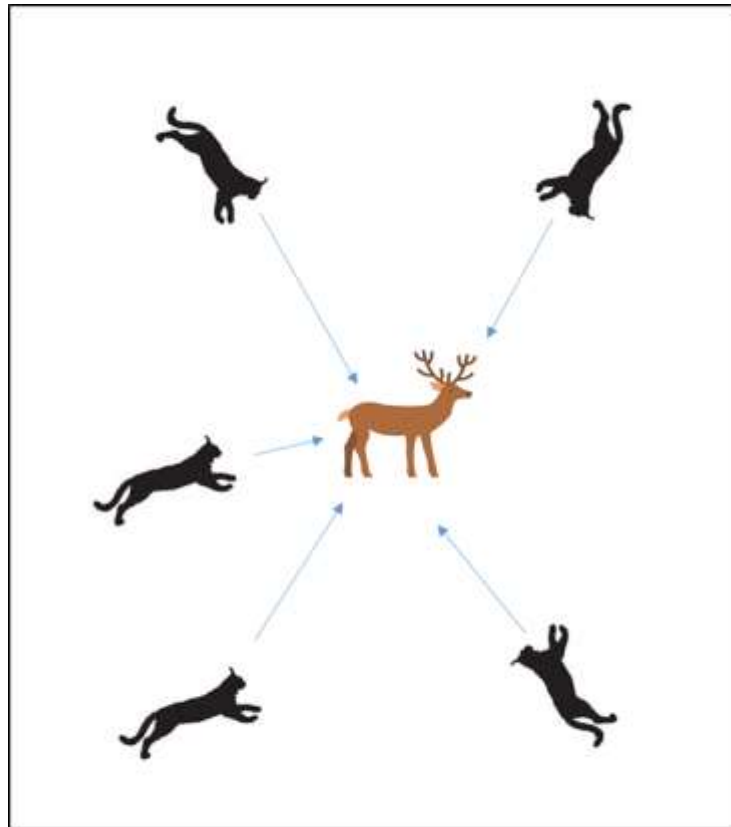


Figure 5: Demonstration for the convergence of bobcats toward the optimal solution

“BOA is different from other metaheuristic algorithms accordingly to how it was inspired and how it works. The following main points illustrate the distinctive aspects of BOA's biological modelling, search technique, and adaptive behavior in comparison to other algorithms [18]:

- Exceptional biological inspiration: BOA is different from other metaheuristics like WOA and GA because it is based on its uniqueness in goal-driven hunting behavior. It focusses on how predators of bobcats interact with each other toward the prey. Each candidate solution mimics a bobcat that explores its environment independently, then switches to exploitation by simulating a targeted chase toward the best-found prey. This biologically faithful modeling introduces a novel structure for balancing exploration and exploitation in optimization problems.
 - Dual phases search mechanism: BOA delineates a definite separation between the periods of exploration and exploitation. It entails a systematic search procedure, improved capacity to evade local optima, and a maintained equilibrium between global and local searches.

- Adaptive behavior: BOA demonstrates adaptive search dynamics by modifying its movement according on the prevailing distribution and fitness of candidate solutions. This capability to respond enables the algorithm to alternate between exploration and exploitation, facilitating proximity to optimal solutions while mitigating the risk of premature stagnation.

It has been cited that empirical assessments have confirmed BOA capability. Particularly, in solving complex optimization problems. That is, it surpasses many well-known metaheuristic algorithms, while maintaining a good balance between exploration and exploitation”.

3.3 Integrating between CNN and BOA

The BOCNN model is proposed as discussed earlier. BOA is used in the two phases of tracking for prey then chasing it for the best solution which means the best optimised hyperparameters set of CNN. “For BOA ranking the quality of possible solutions, it critically depends on the provided objective function. The fitness function could vary depending on the specific application, such as maximizing accuracy or minimizing error in AI tasks. The fitness function is used to evaluate the model's performance during the optimization process. In this study, it is suggested to combine both training and validation accuracies to facilitate the balance of the model's generalization. It is constructed as a weighted aggregation of training accuracy and validation accuracy. This weighting system emphasizes the model's proficiency in learning from training data while also accounting for its generalization abilities on validation data. The inclusion of validation accuracy aids in reducing the likelihood of overfitting. The fitness function of the WOCNN can be illustrated by the following equation:

$$F = (\alpha \times TA) + (\beta \times VA) \quad (6)$$

where F represents the fitness function, α and β represent variable factors as $(\alpha + \beta = 1)$, TA represents the training accuracy, and VA represents the validation accuracy.

The training accuracy metric evaluates the model's learning from the training data, whereas validation accuracy measures the model's generalization to unseen data throughout the training process [19]. To find the best CNN-optimizing hyperparameters, BOA looks at 18 of them, such as convolutional filter numbers, convolutional kernel size, convolutional kernel stride, pooling size, pooling block stride, dense neuron numbers, and dropping rates. Hyperparameters optimization can

Hybrid Framework between

enhance convergence speed, reduce training duration and lower computing expenses [20]". Such proposed optimization can improve predictive performance and provide more efficient model for osteoporosis diagnosis.

4. Results and Discussions

4.1 Employed Dataset

This dataset is usually known as MLSI. Diagnosis of osteoporosis is acquired by the DEXA scan. Contains 211 lumbar spine pictures. The dataset includes images cropped from the DEXA display monitor, with the approximate size of $453 \times 519 \times 3$ pixels. All images are in the Joint Photographic Experts Group (JPEG) format. The images were collected and validated at Dr. Khalid Ghanim Majeed's private clinic in Mosul, Iraq. The dataset is divided into osteoporotic and normal cases according to the BMD measurement.

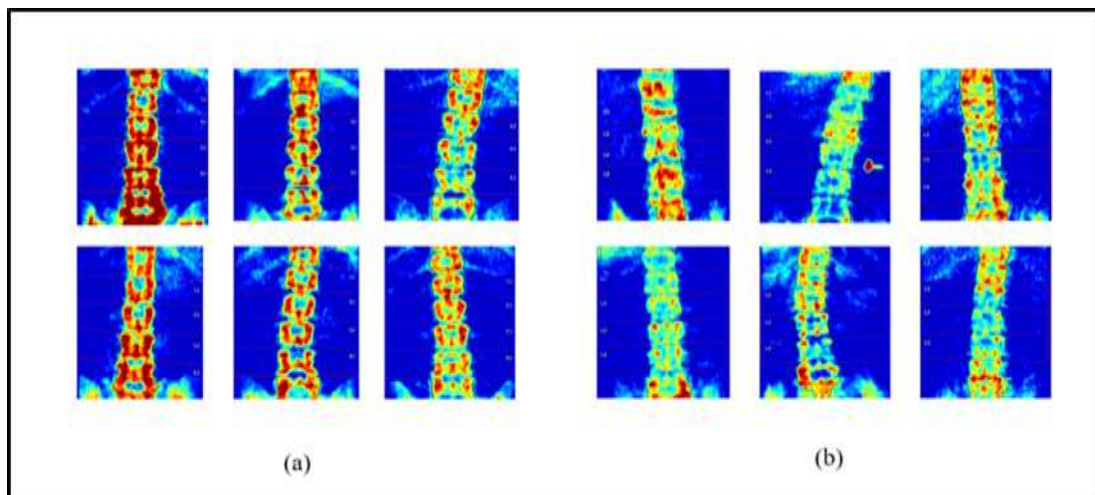


Figure 6: MLS Diverse Images from dataset (a) Normal (b) Osteoporotic

4.2 Standardizations

The MLSI dataset has 126 training images, 42 validation images, and 43 testing images, for a total of 211 images in three categories. All images are normalised to the size of $128 \times 128 \times 3$. To improve the network's performance and to avoid overfitting, we increased the number of training and validation

images. Augmentation operations: rescaling, shearing by a factor of 0.2, zooming by a factor of 0.2, and horizontal flipping. They have been used to provide insight into the diversity between the images offered.

All experiments were performed on a laptop computer with the following specifications: The Lenovo laptop features an Intel core i5 processor running at 2.20 GHz and comes with 12 GB of RAM. Code has been executed online via google Colab platform by activating Pro Pulse version with Python programming language version 3.9. and a Central processing unit (CPU) with a maximum of 51 gigabytes of Random Access Memory (RAM).

4.3 BOA Optimization

As previously mentioned, the proposed BOCNN model utilize BOA to optimize CNN hyperparameters. Figure 7 shows the development of fitness score curve over iterations.

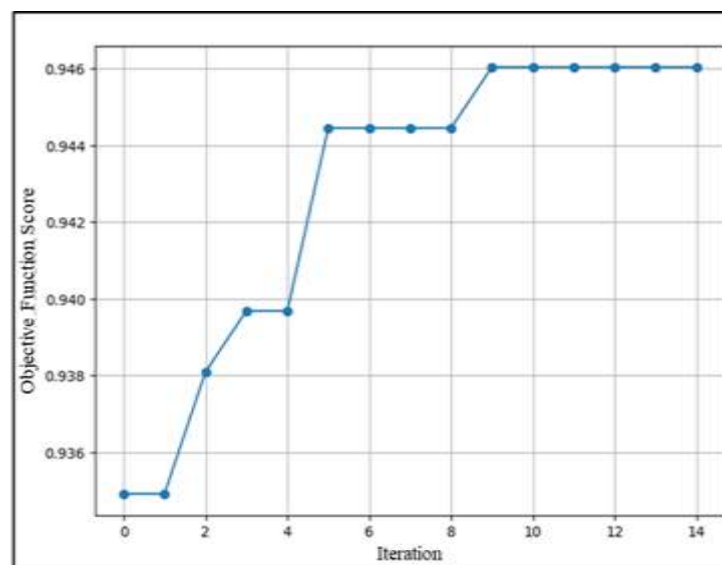


Figure 7: Development of fitness score curve over iterations

This figure demonstrates the progression of the best fitness score over 15 generations during the BOA hyperparameter tuning process. Initially, the fitness score is relatively low, starting at approximately 0.935. Nonetheless, it advances swiftly inside the initial generations, demonstrating the algorithm's capacity to promptly discern more advantageous choices. From the 2nd to the 5th generations, the enhancement progresses consistently, indicating that BOA is effectively optimizing its search within a

Hybrid Framework between

constrained solution space. Between the 6th and 8th generations, the rate of enhancement diminishes, signifying that the algorithm is shifting from exploration to fine-tuning, implementing incremental modifications to optimize performance. The curve reaches its apex at the 9th generation with a fitness score about 0.946 and maintains stability until the 14th generation. This plateau indicates the approach towards an ideal or near-optimal solution. The graphic illustrates BOA's ability to reconcile global search with local exploitation, showcasing its efficacy in optimizing hyperparameters to improve model performance.

4.4 BOCNN Specifications

After several experiments the proposed CNN model for classification of the MLIS images into normal and osteoporosis cases. It has 11 layers architecture that extracts and classifies osteoporosis features systematically. The layers are input layer, first convolution layer, first pooling layer, second convolution layer, second pooling layer, third convolution layer, third pooling layer, flatten layer, first dense layer, second dense layer and softmax layer. BOA optimises 18 hyperparameters for the BOCNN, which are located in the CNN main layers of convolution, pooling and dense, where all such layers are considered. Each convolution layer has the parameters such as number of filters, kernel size and kernel stride. The parameters of any pooling layer are: block size, block stride. The parameters in each dense layer are the number of neurones and the dropout rate. Table 1 summarises the 18 hyperparameters and their initial setup BOA values.

“Table 1: The examined CNN hyperparameters along with their initial value configurations (search space)

Index	Hyperparameter	Initial Value	Type
1	Number of filters in layer 1	Between 32 to 256	Random Integer
2	Kernel size in layer 1	3, 4 or 5	Random choice
3	Kernel stride in layer 1	1 or 2	Random choice
4	Block size in layer 2	2, 3 or 4	Random choice
5	Block stride in layer 2	1 or 2	Random choice
6	Number of filters in layer 3	Between 32 to 256	Random Integer
7	Kernel size in layer 3	3, 4 or 5	Random choice
8	Kernel stride in layer 3	1 or 2	Random choice

Hybrid Framework between

9	Block size in layer 4	2, 3 or 4	Random choice
10	Block stride in layer 4	1 or 2	Random choice
11	Number of filters in layer 5	Between 32 to 256	Random Integer
12	Kernel size in layer 5	3, 4 or 5	Random choice
13	Kernel stride in layer 5	1 or 2	Random choice
14	Block size in layer 6	2, 3 or 4	Random choice
15	Block stride in layer 6	1 or 2	Random choice
16	Number of dense neurons in layer 7	Between 64 to 512	Random Integer
17	Number of dense neurons in layer 8	Between 64 to 512	Random Integer
18	Dropout rate in layers 7 and 8	Between 0.1 to 0.5	Random float"

The best hyperparameters are selected by the maximum fitness value. Fig. 8 shows the final BOCNN design, it has the best values achieved for the 18 hyperparameters. For the first convolutional layer, the number of filters is 79 filters, the kernel size is 5×5 pixels and the stride is 2×2 pixels. First pooling layer: block size = 3×3 pixels, stride = 2×2 pixels. For second convolution layer, number of filters = 163 filters, kernel size = 5×5 pixels and stride = 1×1 pixel. The block size of the second pooling layer is 4×4 and the stride is 1×1 pixels. The number of filters for the third convolution layer is 125 filters with kernel size 5×5 pixels and stride 2×2 pixels. For the third pooling layer, block size = 3×3 pixels, stride = 1×1 pixels. Dense layer 1: no. of neurones = 226 nodes, dropout = 0.3. For second dense layer # of neurones = 286 nodes and dropout = 0.3.

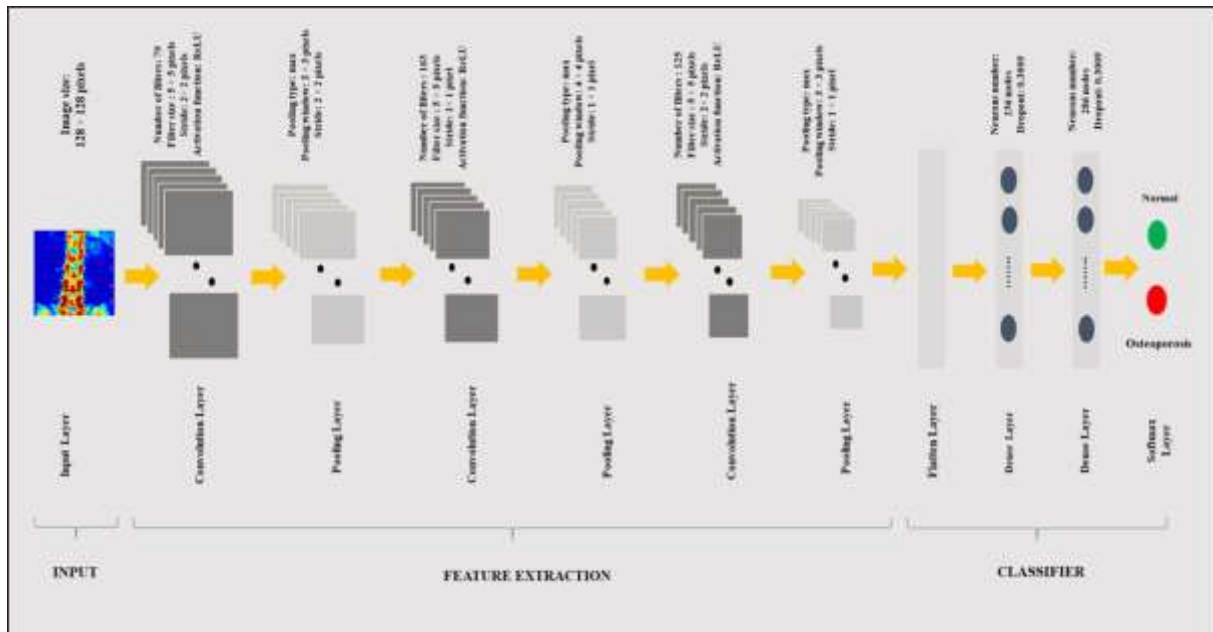


Figure 8: full architecture of the proposed BOCNN model

4.5 BOCNN Training and Validation

The BOCNN is then trained with the optimal hyperparameters values found. The training settings are as following: ADAM optimiser, learning rate 0.001, batch size 32, max pooling window, ReLU activation function, same padding, maximum 60 epochs, early termination if validation loss does not improve over 5 epochs. Watch the validation loss to save time and avoid overfitting.

Figure 9, part A, presents the training and validation accuracy curves of the model over 40 epochs. During the early stages (epoch 0 to 5), there is a significant rise in training accuracy, suggesting that the model quickly starts to grasp patterns from the training data. The validation accuracy shows a similar rising trend, although it is characterized by notable oscillations that reflect heterogeneity in the model's generalization performance. As training progresses beyond epoch 10, both curves generally ascend, with training accuracy stabilizing above 0.90. The validation accuracy exhibits oscillations during training but remains close to the training accuracy, indicating a well-balanced learning process marked by adequate generalization and minimal overfitting.

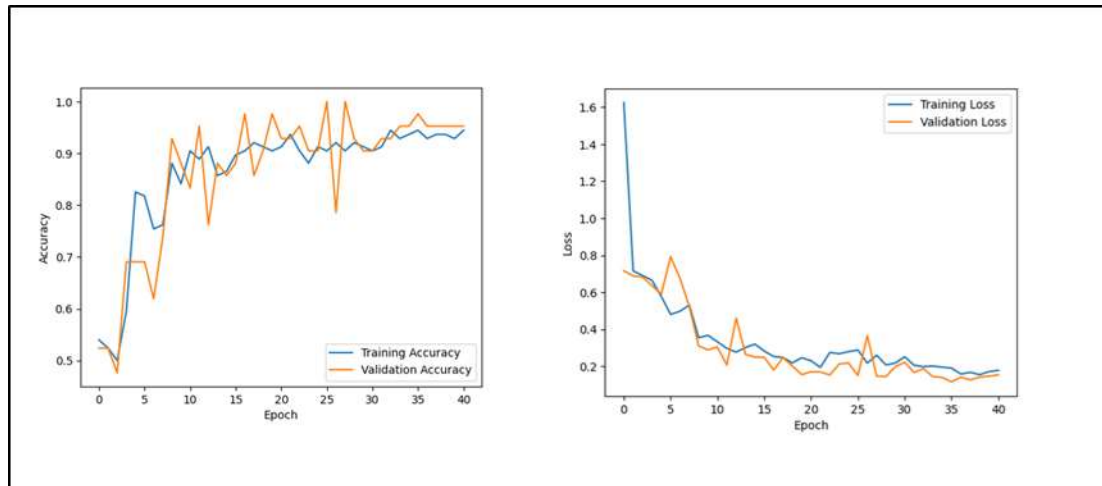


Figure 4.9: Graphs depicting the training and validation phases of the BOCNN model. The accuracies' training and validation curves are displayed on the left, while the losses' training and validation curves are presented on the right.

Part B of the Figure 4.9 depicts the paths of training and validation loss. The training loss commences at a significantly high level and declines rapidly in the initial epochs, signifying effective early learning. The validation loss demonstrates a significant decline during this phase and maintains a consistent downward trend, ultimately stabilizing at a low range in the subsequent epochs. Intermittent small increases in validation loss could indicate a temporary responsiveness to validation samples or slight overfitting; however, the general trend stays consistent. Collectively, these curves illustrate the model's ability to learn efficiently from the data while preserving generalization performance over the course of the training period.

4.6 GOCNN Testing

Finally, the testing phase of BOCNN model has exploited unexplored data of MLSI dataset. Figure 9 shows the confusion matrix for the model's performance on test data for the normal and osteoporotic classes. The model properly identifies 20 instances as normal and 21 cases as osteoporosis. However, it has misclassified two normal cases. This results in a recall of 91%. The accuracy for osteoporosis detection has achieved the exact proportion of 100%, which means every anticipated osteoporosis is right. The overall accuracy has been benchmarked to be 95%, which shows that our proposed BOCNN model can identify well between normal and osteoporotic cases for real clinical data. The high findings suggest that

Hybrid Framework between

the BOCNN is learning important patterns from the training data and is able to confidently detect the presence or absence of osteoporosis.

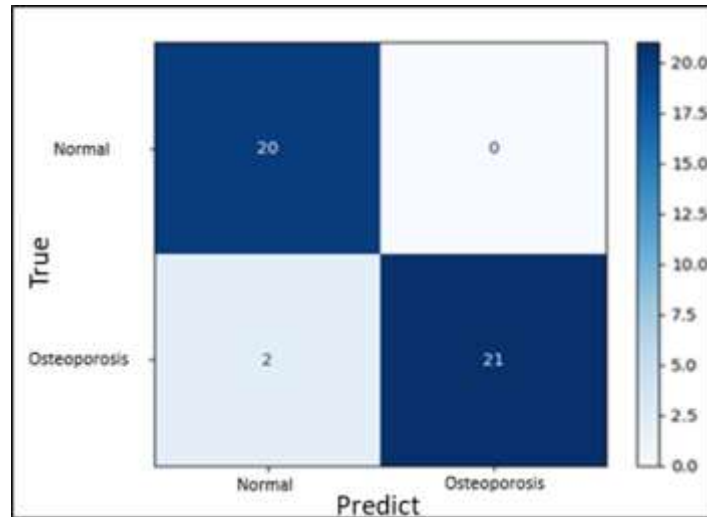


Figure 10: Confusion matrix for the WOCNN testing phase

The ROC curve shows the trade-off between the true positive rate (sensitivity) and the false positive rate across different classification thresholds. Both classes achieve an impressive AUC score of 0.98, indicating outstanding discriminatory power.

The ROC curve for the "normal" class, shown in blue, rises sharply towards the upper-left corner, demonstrating that the model correctly identifies a high proportion of true negatives and true positives with minimal false positives. Similarly, the red curve for the "osteoporosis" class follows a comparable trajectory, also approaching the ideal point (0,1), which denotes perfect classification

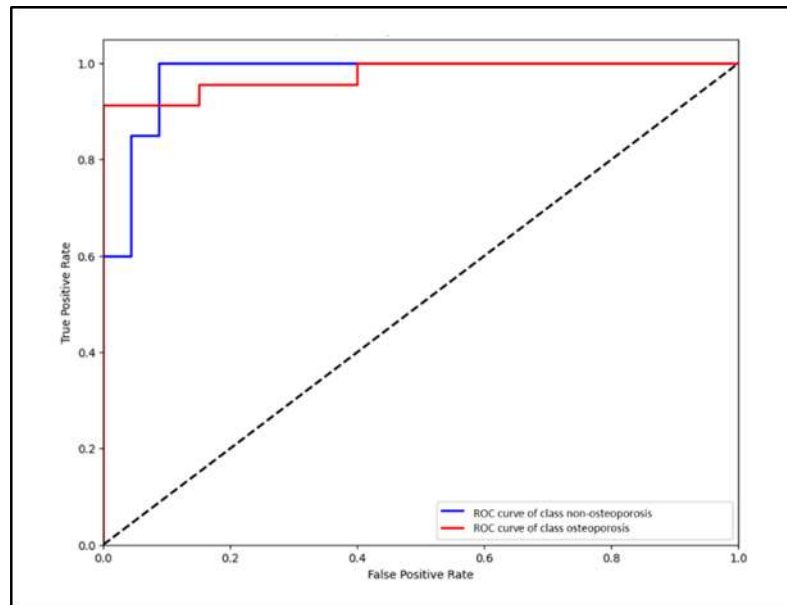


Figure 10: The ROC curve for normal and osteoporosis classifications

This result shows that the BOCNN model has achieved very high accuracy in the diagnosis of osteoporosis, which may contribute much to the early detection and intervention strategies of osteoporosis. Moreover, it can help evaluate the progress of treatment for patients, thus decreasing the need for manual and time-consuming evaluations.

4.7 Contrasts with alternative Models

Comparisons with other DL models are also considered in this study, they include: EfficientNet, InceptionV3, GoogleNet, MobileNet, ResNet and VGG16. All of these DL models have been trained, validated and tested for the same task and by using the same data, this ensures providing fair comparisons. Figure 11 shows the confusion metrics of testing the compared models, they can be considered as the bases of calculating the performance evaluations such as the accuracies and AUCs.

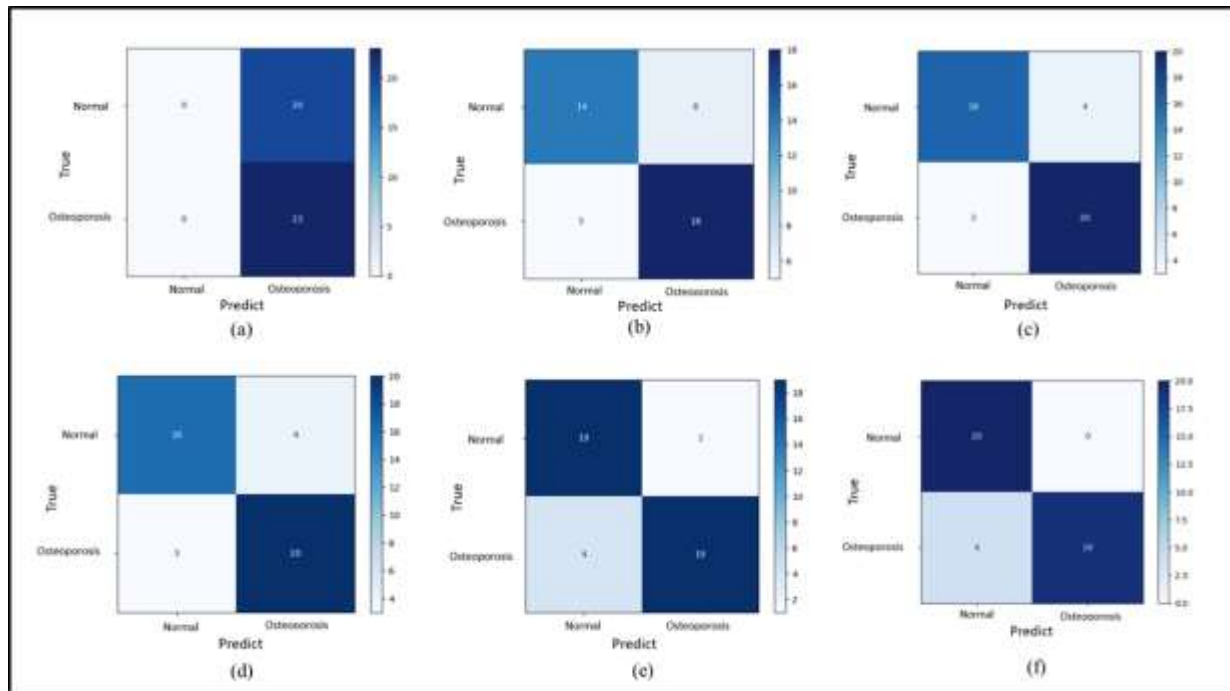


Figure 4.18: Confusion matrices of testing the compared DL models: (a) EfficientNet, (b) InceptionV3, (c) GoogleNet, (d) MobileNet, (e) ResNet and (f) VGG16

These results allow observing significant comparisons. EfficientNet attains the worst performance, with the accuracy of 53.4% and AUC of 0.60. These statistics indicate that this network has no suitable hyperparameters. InceptionV3 obtains better results, with the accuracy of 74.4% and AUC of 0.88, which can be attributed to its multi-branch architecture and slightly appropriate hyperparameters. GoogleNet and MobileNet both produce similar results, with the accuracy of 83.7% and AUC of 0.94. Such outcomes reveal that their hyperparameters are better but remain limited. ResNet achieves even better performance, with the accuracy of 88.3% and AUC of 0.93, reflecting the advantage of residual connections in mitigating vanishing gradients and improving the effects of its hyperparameters. VGG16 reaches the highest results among all previously compared DL models, with the accuracy of 90.7% and AUC of 0.94. These results confirm the capability of this DL model, however, it still lacks for the hyperparameters optimization to further enhance its performance.

5. Conclusion

In this study, we used a dataset of DEXA clinic in Mosul, Iraq from MLSI to classify the cases as normal or having osteoporosis. We introduce the BOCNN model. It shows the capability of BOA to optimise 18 hyperparameters of a CNN to achieve the maximum feasible accuracy. The results highlight

the importance of the choice of optimal values of CNN hyperparameters, since they might have a considerable impact on the acquired performance.

The MLSI dataset consists of 211 genuine clinical photos. They are separated in training, validation and testing sets. The different hyperparameters studied include the number of filters, convolutional kernel sizes, convolutional kernel strides, pooling block sizes, pooling block strides, number of dense neurones, and dropout rate. The best configuration of the suggested BOCNN model attained a very high testing accuracy of 95%. Furthermore, 0.98 AUC value is attained for each case of normal class and osteoporosis class. Accuracy is at a flawless 100%. The recall achieved an impressive 91% success rate.

References:

- [1] P. C. Kristie N. Tu, PharmD, BCPS, CGP; Janette D. Lie, PharmD, BCACP; Chew King Victoria Wan, PharmD Candidate; Madison Cameron, PharmD Candidate; Alaina G. Austel, PharmD Candidate, and Jenny K. Nguyen, "Osteoporosis: A review of treatment options," *Pharmacy and Therapeutics*, vol. 43, no. 2, pp. 92–104, 2018.
- [2] E. Michael Lewiecki, Deborah Cotton, Darren Taichman, and Sankey Williams, "In the Clinical Osteoporosis," *Annals of Internal Medicine*, vol. 155, no. 1, 2011, doi: 10.7326/0003-4819-155-1-201107050-01001.
- [3] J. A. Kanis, E. V. McCloskey, H. Johansson, N. C. Harvey, A. Oden, and F. Borgström, "European guidance for the diagnosis and management of osteoporosis in postmenopausal women," *Osteoporosis International*, vol. 31, no. 1, pp. 1–15, 2020, doi: 10.1007/s00198-019-05176-3.
- [4] R. Jang, J. H. Choi, N. Kim, J. S. Chang, P. W. Yoon, and C. H. Kim, "Prediction of osteoporosis from simple hip radiography using deep learning algorithm," *Scientific Reports*, vol. 11, no. 1, 2021, doi: 10.1038/s41598-021-99549-6.
- [5] S. Bhattacharya, D. Nair, A. Bhan, and A. Goyal, "Computer Based Automatic Detection and Classification of Osteoporosis in Bone Radiographs," in *2019 6th International Conference on Signal Processing and Integrated Networks, SPIN 2019*, 2019, pp. 1047–1052, doi:

10.1109/SPIN.2019.8711616.

- [6] T. T. Ho, J. G. Lee, S. J. Lee, H. Kim, and J. H. Kim, "Deep learning-based prediction of bone mineral density and osteoporosis from plain radiographs," *IEEE Access*, vol. 8, pp. 142402–142412, 2020, doi: 10.1109/ACCESS.2020.3013795..
- [7] Y. Zhang, H. Wang, L. Xu, and Q. Li, "Deep learning-based classification of osteoporosis using lumbar spine radiographs," *Osteoporosis International*, vol. 33, no. 5, pp. 1125–1134, 2022, doi: 10.1007/s00198-021-06171-7.
- [8] R. Jang, J. H. Choi, N. Kim, J. S. Chang, P. W. Yoon, and C. H. Kim, "Prediction of osteoporosis from simple hip radiography using deep learning algorithm," *Scientific Reports*, vol. 11, no. 1, 2021, doi: 10.1038/s41598-021-99549-6.
- [9] J. Pan, P.-C. Lin, S.-C. Gong, Z. Wang, R. Cao, Y. Lv, K. Zhang, and L. Wang, "Effectiveness of opportunistic osteoporosis screening on chest CT using a deep convolutional neural network model," *Scientific Reports*, vol. 14, 2024.
- [10] R. Dhanagopal, R. Menaka, R. Suresh Kumar, P. T. Vasanth Raj, E. L. Debrah, and K. Pradeep, "Channel-boosted and transfer learning convolutional neural network-based osteoporosis detection from CT scan, dual X-ray, and X-ray images," *Computational Intelligence and Neuroscience*, vol. 2024, Article ID 3733705, 2024, doi: 10.1155/2024/3733705.
- [11] A. Siddiqua, "Automated osteoporosis diagnosis from DEXA images using U-Net and transfer learning," *arXiv preprint arXiv:2412.09330*, 2024.
- [12] A. K. Abdullah, R. R. O. Al-Nima, and K. G. Majeed, "Accurate osteoporosis diagnosis model proposing genetic algorithm optimization for convolutional neural network," *Proceedings of the 2nd International Conference on Sustainable Development Techniques (ICSDT2025)*, vol. 5, no. 1, 2025, doi: 10.56286/jc6qnd62.
- [13] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [14] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *Journal of Machine Learning Research*, vol. 13, pp. 281–305, 2012.

Hybrid Framework between

- [15] Ahmed Saeed Ibrahim Al-Obaidi, Raid Rafi Omar Al-Nima, and Tingting Han, “Interpreting arabic signal phabet by using the deep learning,” in *AIP Confonfrence Proceedings*, vol. 2862, 2023, doi: 10.1063/5.0171428.
- [16] Saad Albawi, Tareq Abed Mohammed, and Saad Al-Zawi, “Understanding of a convolutional neural network,” *Proceedings of 2017 International Conference on Engineering and Technology*,. pp. 1–6, 2017, doi: 10.1109/ICEngTechnol.2017.8308186.
- [17] H. Wu and J. Zhao, “Deep convolutional neural network model based chemical process fault diagnosis,” *Computers and Chemical Engineering*, vol. 115. pp. 185–197, 2018, doi: 10.1016/j.compchemeng.2018.04.009.
- [18] Z. Benmamoun, K. Khlie, and B. Gulnara, “Bobcat optimization algorithm: An effective bio-inspired metaheuristic algorithm for solving supply chain optimization problems,” *Scientific Reports*, vol. 14, no. 1, 2024, doi: 10.1038/s41598-024-70497-1.
- [19] Ian Goodfellow, Yoshua Bengio, and Aaron Courville, , *Deep Learning*. MIT Press, 2016.
- [20] Lisha Li, Kevin Jamieson, Giulia DeSalvo, Afshin Rostamizadeh, and Ameet Talwalkar, “Hyperband: A novel bandit-based approach to hyperparameter optimization,” *Journal of Machine Learning Research*, vol. 18, no. 1, pp. 1–52, 2018.