



## An Ensemble-Of-Deep Learning Model with Optimally Selected Features for Osteoporosis Detection from Bone X-Ray Images

Pallavi Hallappanavar Basavaraja<sup>1\*</sup>

Shanmugarathinam Ganesarathinam<sup>1</sup>

<sup>1</sup>Presidency University, Bengaluru, India

\* Corresponding author's Email: pallavihb.7@gmail.com

---

**Abstract:** Osteoporosis is a condition wherein bone tissue deteriorates and bone strength deteriorates. Over time, the disease that can lead bone to become more permeable and fragile makes it more prone to breaking. Early recognition of osteoporosis via an X-ray approach will enhance diagnosis and treatment practices, but it will also assist in preventing national economic loss via mass screening and awareness. A novel early osteoporosis diagnosis model is developed for X-ray images in this research work. The following four primary steps are used to construct a unique Osteoporosis detection model in this study: “(a) pre-processing, (b) feature extraction, (c) optimal feature selection, and (d) Osteoporosis detection”. Gabor filtering (noise reduction) and histogram equalization are used to pre-process the obtained raw data (X-ray) (quality enhancement). Features such as “Active shape model (ASM), active appearance model (AAM), gray level co-occurrence matrix (GLCM), mean local gradient pattern (M-LGP), and improved median ternary pattern (I-MTP)” are recovered from the pre-processed data. Following that, a new hybrid optimization model chooses the best features from the retrieved features. The cat guided hummingbird foraging algorithm (CGHFA) is a conceptual combination of the basic artificial hummingbird algorithm (AHM) and the cat hunting optimization algorithm (CHOA). The deep learning classifiers in the Osteoporosis detection phase are trained using these ideally selected characteristics. The newly created ensemble-of-classifiers model is used to represent the osteoporosis diagnostic phase. “Quantum deep neural network (QDNN), improved deep convolution neural network (I-DCNN) and recurrent neural network (RNN)” are some of the deep learning classifiers that is employed here. All of these classifiers are trained using the optimal features available. The loss function of DCNN is improved via harmonic mean based cross-entropy function. The final detection performance will be calculated by combining the results obtained from all of these characteristics (by taking the mean). Finally, the effectiveness of the anticipated model is validated by a comparative examination. Accordingly, the detection accuracy attained by the proposed deep ensemble model +CGHFA at Learn\_rate=60 is 90.7%, at Learn\_rate=70 is 92.14%, at Learn\_rate=80 is 93.482% and at Learn\_rate=90 is 94.8%, which is higher than the existing models.

**Keywords:** Osteoporosis detection, X-rays, Early diagnosis, Mean local gradient pattern (M-LGP), Improved median ternary pattern (I-MTP); Cat guided hummingbird foraging algorithm (CGHFA), I-DCNN.

---

### 1. Introduction

Osteoporosis has become more common and well-known in recent years, with approximate 200 million people globally suffering from the disease [1, 2, 3]. Nonetheless, nearly 75% of such persons are misdiagnosed and untreated. The “World health organization (WHO)” defines osteoporosis as “a systemic skeletal illness deterioration of bone tissue and microarchitectural degradation of bone tissue with a corresponding increase in bone fragility and fracture susceptibility”[4, 5, 6]. Osteoporosis-related

fractures have a significant risk of the cause of death and disability, a lower standard of living, and raised medical costs dramatically.

Early diagnosis is critical for prompt diagnosis and therapy of individuals who are at risk for osteoporotic fractures. Diagnostic imaging of osteoporosis serves two main goals in this frame of reference: (1) to detect the existence of osteoporosis and (2) to obtain prognostic information on the likelihood of future bone deformation from bone density ]quantification [7]. The quantitative evaluation of BMD, which has now been regarded the

greatest indicator of osteoporotic bone fractures, is being used to diagnose osteoporosis [8]. Although being the gold standard for osteoporosis diagnosis, X-ray could be extensively employed as an osteoporosis screening technique [9].

Artificial intelligence (AI) has recently been employed in medical image interpretation. Furthermore, numerous researchers sought to use AI technologies to deliver an effective osteoporosis testing method. A few attempts have employed machine learning or deep learning algorithms to predict osteoporosis based on basic radiography data [10, 11].

The major contribution of this research work is:

- To extract the M-LGP and I-MTP and the other texture and shape-based features.
- To select the optimal features using the new CGHFA, which is indeed the combination of standard CHOA and AHM
- The improved ensemble deep learning classifiers is designed in the Osteoporosis detection with quantum deep neural network (QDNN), improved deep convolution neural network (I-DCNN) and recurrent neural network (RNN)
- The loss function of DCNN is improved via harmonic mean based cross-entropy function.

The rest of this paper is arranged as: section 2 tells the literature works done in osteoporosis detection. Section 3, section 4, section 5 and section 6 tells about proposed osteoporosis detection from bone x-ray images, feature extraction, optimal feature selection via CGHFA and detection with an improved deep ensemble model, respectively. The results acquired with the projected model are validated in Section VII. This paper is concluded in Section VIII.

## 2. Literature review

Zhang *et al.* [1] established a DCNN model using lumbar spine X-ray images in 2020, to detect osteopenia and osteoporosis. Devikanniga *et al.* [2] created an effective classification algorithm for osteoporosis diagnosis in 2020. The tried-and-true ELM has been chosen for all of this. Furthermore, a unique hybrid metaheuristic algorithm has been created by mixing two nature-inspired heuristic search techniques, such as the “artificial algae algorithm with multi-light source and the Monarch butterfly optimization method”. Mebarkia *et al.* [3] developed an osteoporosis diagnostic method shape and texture assessment in 2021. The feature extraction process implements the hierarchical multiscale local binary pattern (H-MLBP) approach. Tejaswini *et al.* [4] suggested a simple first-line approach for detecting and predicting osteoporosis in 2016. With

the aid of LabVIEW, an impulse response assessment was conducted on the tibial bone to diagnose osteoporosis. In 2017, Reshmalakshmi *et al.* [5] proposed a technique for assessing the trabecular bone architecture to diagnose osteoporosis condition. The goal of such technology has always been to aid doctors in detecting, controlling, and treating osteoporosis in its initial stages. The severity of the illness has been determined by examining the trabecular bone structure with a variety of computer vision techniques and typical visual quality criteria.

Numerous methods have been employed for osteoporosis detection. But, still there exist a common problem like no elimination in variability across images owing to various X-ray scan settings [1], low classification accuracy [2], low precision [3], disturbances due to human contact and other noises [4], and high implementation cost and time consumption [5]. Hence, in order to solve the abovementioned issues, this work proposes a new osteoporosis detection model is for X-ray images.

## 3. Proposed osteoporosis detection from bone X-ray images

### 3.1 Architectural description

In this research work, a novel Osteoporosis detection model is developed by following four major phases: “(a) pre-processing, (b) feature extraction, (c) optimal feature selection and (d) Osteoporosis detection”. The architecture of the projected model is shown in Fig. 1. Let the collected bone x-ray images are denoted as  $img^{X-ray}$ , which is further given for processing.

The steps that is followed in the projected model is furnished below:

**Step 1-** Initially, the collected raw data  $img^{X-ray}$  is pre-processed via Gabor filtering (noise removal) and histogram equalization (quality enhancement). The pre-processed images acquired from the histogram equalization is denoted as  $img_i^{Xpre}$

**Step 2-** From  $img_i^{Xpre}$ , the features like “Active shape model (ASM)  $h^{ASM}$ , active appearance model (AAM)  $h^{AAM}$ , GLCM  $h^{GLCM}$ , mean local gradient pattern (M-LGP)  $h^{M-LGP}$ , improved median ternary pattern (I-MTP)  $h^{I-MTP}$ ” are extracted. These extracted features are fused together as  $H = h^{ASM} + h^{AAM} + h^{GLCM} + h^{M-LGP} + h^{I-MTP}$ .

**Step 3-** Subsequently, the optimal features  $H^{opt}$  is selected from the extracted features  $H$  using a new hybrid optimization model.

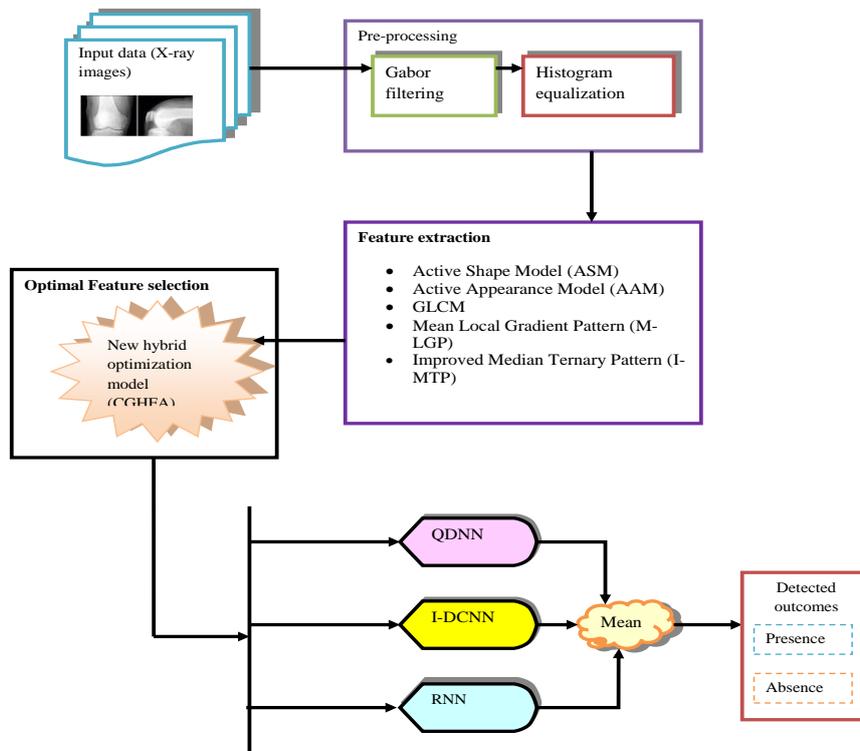


Figure. 1 Architecture of the projected osteoporosis model

The projected hybrid optimization model- CGHFA is the conceptual blend of the standard AHM and CHOA, respectively. The deep learning classifiers in the Osteoporosis detection phase are trained using these optimally selected features.

**Step 4-** The Osteoporosis detection phase is modeled with the newly constructed ensemble-of-classifiers model. Here, deep learning classifiers like QDNN, I-DCNN and RNN are used. All these classifiers are trained using the optimally selected features  $H^{opt}$ . Finally, the outcome acquired from all these features is combined (by taking mean); and this is the final detection performance.

### 3.2 Pre-processing

The collected raw X-ray images  $img^{X-ray}$  are pre-processed to enhance the quality of the image, by means of suppressing the noise and other artifacts within it. The  $img^{X-ray}$  is first filtered via Gabor filtering to remove the noises within it. Fig. 2 depicts this phase diagrammatically. Then, onto the filtered image  $img_i^{Xfilter}$ , histogram equalization is applied to enhance the quality of the image. The resultant acquired from the histogram equalization is the pre-processed image and it is denoted as  $img_i^{Xpre}$ .

**Gabor filtering** [14]: the Gabor filter is utilized to remove the noise within  $img^{X-ray}$ .

**Histogram equalization** [15]: The histogram equalization is used to improve the contrast of the

filtered image  $img_i^{Xfilter}$ . Histogram equalization, often referred as histogram flattening, is indeed a nonlinear lengthening and redistributing of image pixel values that results in a relatively equivalent amount of pixel values from the mean grey range.

### 4. Feature Extraction

From the pre-processed data  $img_i^{Xpre}$ , the features like active shape model (ASM), active appearance model (AAM), GLCM, mean local gradient pattern (M-LGP), and improved median ternary pattern (I-MTP) are extracted. These extracted features are fused together as  $H$ .

#### 4.1 ASM

The active appearance model (AAM), which is well-documented in the scientific literature, is yet another version of the ASM technique. In addition, the statistical shape model of objects, known as the active shape model (ASM), is being used to explain it. Shapes are extracted from images using this approach. The ASM algorithm is depicted below:

- (1) Align each shape of the sample at first. The landmark points are determined (manually) in a collection of  $img_i^{Xpre}$  training images. A point distribution model is created from these groupings of landmark points. The landmark points in the shape

(2)	140	30	240
	130	60	230
	135	20	235

Figure. 2 Pixel values of original image: an illustration

140	0	240
130	0	230
135	0	235
135	0	235

Computed mean value

Figure. 3 Computed mean of the sample image

vector are marked as  $G = (g_1, o_1, \dots, g_n, o_n)^T$  here refers to the  $o^{th}$  landmark in the  $g^{th}$  shape ( $g_i$  is histogram curve).

(2) Repeat step 1 until you reach a point of convergence:

(a) calculate the mean shape. The arithmetic mean of coordinates representing each element of the sample after alignment defines the mean shape. This is mathematically shown in Eq. (4).

$$\mu = \frac{1}{S} \sum_{i=1}^S img_i^{xpre} \quad (4)$$

(b) Using the obtained mean shape, do PCA on the shape vector.

The covariance matrices  $S$  are indicated as Eq. (5).

$$S = \frac{1}{s-1} (g_i - \bar{g})(g_i - \bar{g})^T \quad (5)$$

The Eigen vectors corresponding to the largest Eigen values can be given in the form of matrix as  $\partial$ . Now the shape can be given as:  $G \approx \bar{G} + \partial b$ . Here,  $b$  is a vector of elements with parameters, and it can be computed as  $b = \partial^T \cdot (G - \bar{G})$ .

(b) adjust the mean shape:

(i) By default, to a size, orientation, and origin,

(ii) To the first shape,

(c) Align each shape on the mean shape.

The extracted feature is pointed as  $h^{ASM}$

## 4.2 AAM

The AAM [16, 17] model's algorithm is made up of three steps:

(a) In the training set, linking shape and texture vectors towards each AAM, shape and texture vectors are coupled in the very same vector:  $C_i = (T_i \cdot img_i^{xpre})^T$ . For  $m^t$  pixel, the texture vector is denoted as  $T_i$

(b) In the training set, calculate the correlation coefficient matrix for related shape and texture vectors.

(c) Using PCA analyzes the correlation coefficient matrix. The hybrid parameters are the eigenvectors  $P$  of  $C$ , and they are required for monitoring the form and texture of the AAM model.  $C_i = \bar{C} + A \cdot (P \cdot b)$  Can be proposed for each pattern in the training program. Here,  $A$  is the diagonal matrix and diametrical contrast correspondent to  $i^{th}$  coefficient is pointed as  $C$ . The extracted feature is pointed as  $h^{AAM}$ .

## 4.3 GLCM

In general, this GLCM [18] in image data is used to create a matrix co-occurrence, wherein the features matrices function may be obtained. The extracted GLCM features are variance, sum variance, sum entropy, sum average, mathews correlation coefficient (MCC) (2nd higher eigen value of), IMC1, IMC 2, homogeneity, entropy, energy, difference variance, difference entropy, correlation, contrast. The extracted feature is pointed as  $h^{GLCM}$ .

## 4.4 Mean local gradient pattern (M-LGP)

The LGP operator calculates the gradient values of a specified pixel's eight surrounding pixels that are computed as absolute values of intensity differences between given pixel and its neighbors. It has drawbacks, such as (a) producing large histograms, which reduces identification speed, and (b) missing the local feature in certain instances since the influence of the centre pixel also isn't taken into account. A unique I-LGP model has been proposed in this research work to address all of the aforesaid issues. I-steps LGP's are as follows:

Step 1- The gradient value of the pixels is computed by taking the mean of each column (in LGP, the gradient pixel value is computed by comparing the centre pixel and its neighbors). Let the pixel values of  $img_i^{xpre}$  be, as per shown in Fig. 2. The computed mean value is shown in Fig. 3.

Mathematically, the mean computation can be given as per Eq. (6) and Eq. (7).

$$\overline{img_j^{xpre}} = \frac{N}{\sum_{k=0}^{N-1} \frac{1}{img_k^{xpre}}}; j - row; k - column \quad (6)$$

Step 2- After computing the mean of the columns, find the difference between computed mean

140-135	30-30	240-235
130-135	60-30	230-235
135-135	20-30	235-235

Figure. 4 Difference computation

5	0	5
5	30	5
0	10	0

Figure. 5 Computed difference value of the sample image

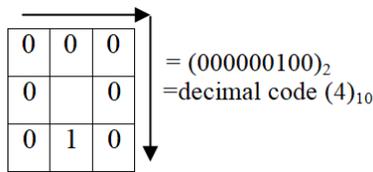


Figure. 6 Computed feature values

0	120	96
0	95	105
0	98	101

Pixel value of original image

Figure. 7 Original image: an illustration

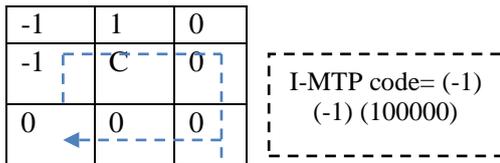


Figure. 8 I-MTP computation

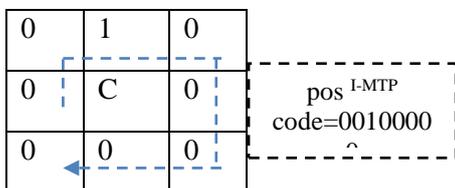


Figure. 9 Positive parts

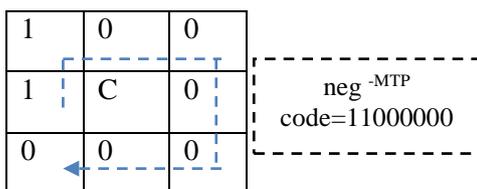


Figure. 10 Negative parts

values and the column values (pixels values in the column). This phase is diagrammatically shown in Fig. 4 and Fig. 5.

$$\bar{g} = \left| \overline{img_{i(k)}^{xpre}} - img_{i(j,k)}^{xpre} \right| \quad (7)$$

Step 3- Compute the median and replace the center pixel. The final outcome is shown in Fig. 6.

The extracted feature is mathematically shown in Eq. (9).

$$f^{I-LGP} = \sum_{j=0}^{N-1} S(img_{i(k)}^{xpre} - \bar{g}). 2^j \quad (9)$$

The extracted feature is pointed as  $h^{M-LGP}$

### 4.5 I-MTP

The MTP involves the integration of the median with the quantification of pixel intensities. MTP overlooks critical texture features that do not consider the intensity variation between adjacent pixels, which could significantly influence environments with non-monotonic lighting change. As a result, a novel I-MTP model is introduced in this study.

The feature representation in the three-level coding scheme includes a certain level of magnitude variation data, rendering it resilient in the presence of non-monotonic illumination variation. The approach is more resistant to random noise because using the median and the three-level coding method makes it simpler to generate texture patterns compatible with the local image characteristic (smooth or high-textured).

First, a 3 × 3 neighbor is defined around each pixel (as shown in Fig. 7) the harmonic mean intensity of the nine pixels is calculated. This is mathematically shown in Eq. (10).

$$f^{I-MTP} = \begin{cases} 1 & V > \frac{MC+GM}{2} + t \\ 0 & \frac{MC+GM}{2} - t \leq V \leq \frac{MC+GM}{2} + t \\ -1 & V < \frac{MC+GM}{2} - t \end{cases} \quad (10)$$

Here,  $MC, GM, t, V$  denotes the local median, harmonic mean, and user-specified threshold and neighbor gray level, respectively.

Each MTP code is further split into its corresponding positive and negative parts and are treated as two separate binary patterns called  $pos^{I-MTP}$  and  $neg^{I-MTP}$ , respectively. This phase is clearly shown in Fig. 8 to Fig. 10, respectively. This is mathematically shown in Eqs. (11-14), respectively.

$$pos^{I-MTP} = \sum_{pix=0}^7 f_{pos}(f_{I-MTP}(i_{pix})) \times 2^{pix} \quad (11)$$

$$f_{pos}(V) = \begin{cases} 1 & V = 1 \\ 0 & otherwise \end{cases} \quad (12)$$

$$neg^{I-MTP} = \sum_{pix=0}^7 f_{neg}(f_{I-MTP}(i_{pix})) \times 2^{pix} \quad (13)$$

$$f_{neg}(V) = \begin{cases} 1 & V = -1 \\ 0 & otherwise \end{cases} \quad (14)$$

The extracted feature is pointed as  $h^{I-MTP}$ .

These extracted features are fused together as  $H = h^{ASM} + h^{AAM} + h^{GLCM} + h^{M-LGP} + h^{I-MTP}$ .

## 5. Optimal Feature Selection via CGHFA

### 5.1 Hybrid optimization Model-CGHFA

A new hybrid optimization approach-CGHFA is proposed in this research work for optimum feature selection (i.e. picking the relevant features by separating them from the irrelevant features).

The following are the stages used in the newly proposed hybrid optimization model:

**Step 1- Initialization** Phase: The  $P$  search agents (hummingbird and cats [13]) are initialized. The position of the search agents  $H_i$  are initialized.

**Step 2-** Generate the solution using the opposition learning (proposed) to enhance the convergence of the solutions and to limit the solution from getting trapped into local optima.

**Step 3- Cat guiding foraging:** As per the proposed concept, each hummingbird (search agents) has an inherent wish to travel the food source (feature) with the most nectar quantity, implying that a target supply has to have a high nectar-refilling rate as well as a lengthy duration between visits by that hummingbird (search agents). According to AHA, a hummingbird should identify the sources of food with the greatest visitation frequency for directed foraging behavior, then select one with the maximum nectar-refilling frequency as its targeted food source (optimal feature or relevant feature). By incorporating a heading switch vector into the AHA algorithm, three flying talents, encompassing "omnidirectional, diagonal, and axial flights", be appropriately utilized and represented during foraging. These flight patterns may be extended to a d-D space by defining the axial flight as shown in Eqs. (17-21), respectively.

$$D^{(i)} = \begin{cases} 1 & \text{if } i = rand([1, d]) \\ 0 & \text{otherwise} \end{cases} \quad (17)$$

The following is the definition of diagonal flight:

$$D^{(i)} = \begin{cases} 1 & \text{if } i = P(j), j \in [1, k], \\ & P = rndperm(k); k = \\ & [2, [rand1.(d - 2) + 1]] \\ 0 & \text{otherwise} \end{cases} \quad (18)$$

The following is the definition of omnidirectional flight:

$$D^{(i)} = 1; i = 1, 2, \dots, d \quad (19)$$

Here,  $rndperm(k)$  creates a random permutation of integers from 1 to  $k$ , and  $rand1$  is a random number generated between  $[0, 1]$ .

The following is the mathematical equation for replicating guided foraging behavior using the cats tracking model and a suitable food source (proposed):  $V_i(itr + 1) = H_d(itr) + rand2.C \times (H_d^{best} - H_d(itr))$

$$rand2 = \frac{1 - \sqrt{C_k + 1}.rand(itr)}{C_k.max^{itr}}; k = 1, 2 \quad (20)$$

Here,  $max^{itr}$  is the maximal iteration and  $itr$  is the current iteration.  $rand2$  is a random number and  $H_d^{best}$  is the search agent with best fitness.  $C1, C2$  are set relatively small.

The  $i^{th}$  food source's location (feature) is updated as follows:

$$H_i(itr + 1) = \begin{cases} H_i(itr) & f(H_i(itr)) \leq f(V_i(itr + 1)) \\ V_i(itr + 1) & f(H_i(itr)) > f(V_i(itr + 1)) \end{cases} \quad (21)$$

$f(.)$  is the function fitness value. This function fitness value  $f(.)$  of this research work is the minimization of the detection errors  $d^{error}$ , and this is the overall objective of this research work. This is mathematically shown in Eq. (22).

$$f = \min(d^{error}) \quad (22)$$

**Step 4-** circle map territorial foraging: A hummingbird is more highly likely to hunt for a new food source after visiting its target food source (optimal feature), where the flower nectar has indeed been consumed, rather than visiting other existing sources of food. As a result, a hummingbird could indeed easily migrate to a close vicinity inside its territory (do not get trapped into local optima), where a new source of food might well be discovered as a

feasible solution that is potentially preferable to the current one. The following is the mathematical equation (Eq. (23)) for modeling the local search of hummingbirds in their circle map territorial foraging strategy and a possible food source

$$V_i(itr + 1) = H_i(itr) + b \cdot D \cdot H_i(itr) \quad (23)$$

Here,  $b$  is the a territorial factor that is computed using the circle chaotic map to enhance the convergence speed of solutions (proposed)

**Step 5- Migration Foraging:** When a hummingbird's favorite foraging spot runs out of food, the hummingbird will travel towards a more faraway source of food. A migration coefficient is specified inside the AHA method. The hummingbird finding there at prey species with worst nectar-refilling rate would travel toward a food source created random in the whole solution space if the number of iterations exceeds the threshold value of the migration factor. This hummingbird then will forsake the old source and eat at the new one, as well as the visitation table would be modified automatically. A hummingbird's migratory foraging from the fountain with the lowest nectar refilling rate toward a new one created at randomness can be defined as per Eq. (24).

$$H_{worst}(itr + 1) = LB + rand3 \cdot (UB - LB) \quad (24)$$

$H_{worst}$  Is the food source with the worst nectar-refilling rate in the population.

**Step 6-** Return the best position of the solutions (i.e. optimal features  $H^{best}$ ).

## 6. Detection with improved deep ensemble model

### 6.1 Improved deep ensemble model

The Osteoporosis detection phase is modeled with the newly constructed ensemble-of-classifiers model. Here, deep learning classifiers like quantum deep neural network (QDNN), improved deep convolution neural network (I-DCNN) and recurrent neural network (RNN) are used. All these classifiers are trained using the optimally selected features  $H^{opt}$ . Finally, the outcome acquired from all these features is combined (by taking mean); and this is the final detection performance.

#### QDNN

The quantum deep neural network [19] includes  $N_i$  inputs  $N_H$  hidden layer and  $N_o$  output layers. The

weight between the output unit  $z$  and hidden unit  $v$  is denoted as  $W_{vz}$ . The weight between the hidden  $v$  and input unit  $u$  is denoted as  $W_{uz}$ . The input to  $v^{th}$  hidden unit from the  $i^{th}$  pattern of  $H^{best}$ , and  $i^{th}$  pattern of  $H^{best}$ , the response of  $v^{th}$  hidden unit is described as per Eq. (25) and Eq. (26), respectively. It is trained with  $H^{best}$ .

$$\overline{hid}_{v,i} = \sum_{z=0}^{N_i} W_{vz} \cdot H_{z,i}^{best} \quad (25)$$

$$hid_{v,i} = \frac{1}{N_s} \sum_{r=1}^{N_s} hid_{v,i}^r = \frac{1}{N_s} sgm(\partial_{hid} (\overline{hid}_{v,i} - \theta_v^r)) \quad (26)$$

Here,  $\partial_{hid}$  is a slope factor and  $N_s$  is the count of levels and  $\theta_v^r$  is the transfer function.

From  $i^{th}$  pattern vector to the output unit, and response of output unit  $u$  to the  $i^{th}$  pattern of  $H^{best}$  is given as per Eq. (27) and Eq. (28), respectively.

$$\overline{out}_{u,i} = \sum_{v=0}^{N_h} W_{uv} \cdot hid_{v,i} \quad (27)$$

$$out_{u,i} = sgm(\partial_{out}(\overline{out}_{u,i})) \quad (28)$$

Here,  $\partial_{out}$  is the slope factor of output transfer functions. During the training phase, the variables of the QNN are modified. There are two phases in the QNN learning process. The synaptic weights are established initially in order to divide the feature space uniformly. The unpredictability of the feature space must still be learnt in the following stage by optimizing the parameters. The outcome is pointed as  $out^{QDNN}$ .

### 6.2 IDCNN

IDCNN is a deep learning model with diverse layers. It is trained with  $H^{best}$ .

**Convolutional layer:** The convolution layer being the first layer is used to extract specific maps from the input image (while keeping relationships between pixels).

**Pooling layer:** This layer down samples feature maps to lower their dimension and enhance feature extraction. To reduce the dimension, the pooling layer combines a local receptive field into a single neuron.

**Batch normalization layer:** Each batch's prior layer activations are normalized to keep the mean activation value near to 0 as well as the standard deviation activation rate close to 1. It may substantially speed up convergence, decrease over fitting, minimize initial weight insensitivity, as well as allow us to employ a faster information gain.

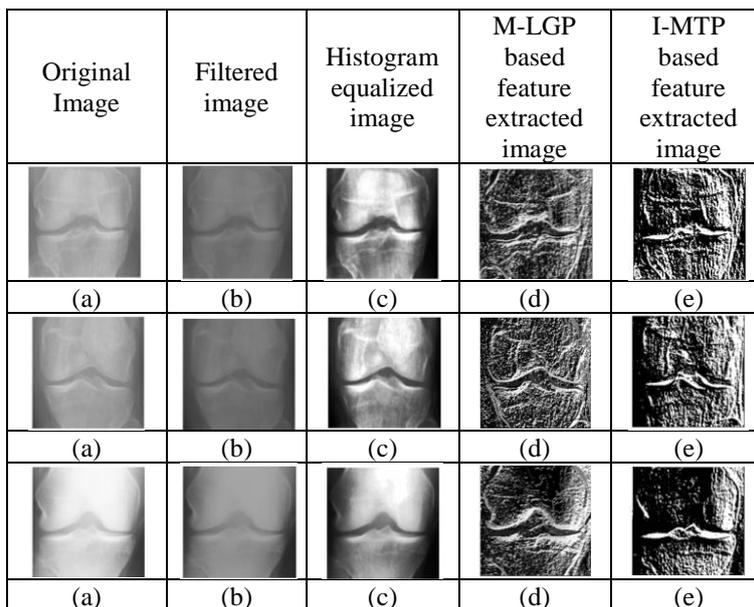


Figure. 11 Sample normal images and their corresponding outcomes

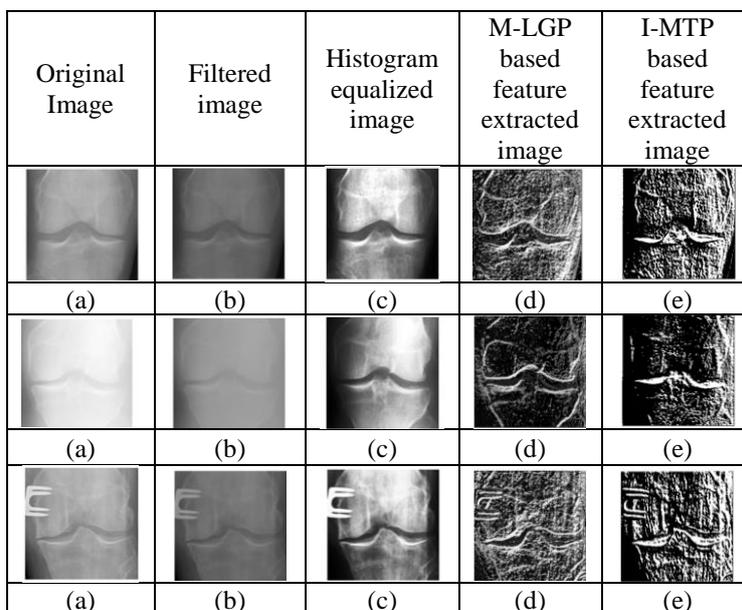


Figure. 12 Sample abnormal images and their corresponding outcomes

Flatten layer: By transforming the input from multi-dimensional space to one-dimensional space, the flattening layer is commonly used to transition from a convolutional layer to a fully connected layer.

Dropout Layer: The dropout process sets the neuron value to 0 at arbitrary with a chance of 50% within every training batch. As a consequence of this process, the CNN will become less receptive to specific groupings of neurons, which help in decreasing the contact among hidden layer neurons, avoid overfitting, and enhance the model's generalization capacity [20].

Proposed loss function- improved cross-entropy: The consistency of the network's output is quantified by the loss function. Instead of using the cross-

entropy, an improved cross-entropy function is used in I-DCNN. This is mathematically shown in Eq. (30).

$$error = - \frac{HM}{\sum_{i=1}^N p_i \cdot \log q_i + (1-p_i) \log(1-q_i)} \quad (30)$$

Here, *HM* denotes the harmonic mean. The outcome is pointed as  $out^{I-DCNN}$

### 6.3 RNN

The output from the previous step is provided as input to the next step in a recurrent neural network (RNN) [20]. RNNs have a "memory" that stores all of the data about the computation. It utilizes the very

same variables for every input since it produces the outcome by performing the same function on all inputs or hidden layers. In contrast to other neural networks, this reduces the complexity of the parameters. It is trained with  $H^{best}$ .

The current state  $c_t$  is computed as per Eq. (31).

$$c_t = f(c_{t-1}, X_t) \quad (31)$$

Here,  $c_{t-1}$  is a previous state and  $X_t$  is input state. The activation function ( $\tanh$ ) is computed as per Eq. (32).

$$c_t = \tanh(W_{hh}c_{t-1} + W_{xc} \cdot X_t) \quad (32)$$

Here,  $W_{hh}$  and  $W_{xc}$  weight of recurrent neuron as input neuron, respectively.

The output is computed as per Eq. (33).

$$y_t = W_{cy} \cdot c_t \quad (33)$$

Here,  $y_t$  and  $W_{cy}$  denoted the output and weight of output layer, respectively. The outcome is pointed as  $out^{RNN}$ .

The mean of the outcome from quantum deep neural network (QDNN), improved deep convolution neural network (I-DCNN) and recurrent neural network (RNN) are computed, and it is the final outcome.

$$out = \frac{out^{RNN} + out^{I-DCNN} + out^{QDNN}}{3} \quad (34)$$

The final outcome is acquired from  $out$ .

## 7. Result and discussion

### 7.1 Simulation procedure

The proposed model has been implemented in python. The improved deep ensemble model +CGHFA is evaluated with the data collected from: <https://www.kaggle.com/parv619/data-sprint-35-osteoarthritis-knee-xray>. The samples images are shown in Fig. 11 and Fig. 12, respectively. Among the collected data, 70% of the information is used for training and the rest 30% is used for testing purposes. The assessments were carried out at different learning rates of 60, 70, 80, and 90, respectively. Positive performance (specificity, sensitivity, precision, and accuracy) and negative performance (FPR, FNR) and other metrics are used to assess the suggested classifier and method (F1-score and MCC). Positive parameters of the proposed osteoporosis detection

from X-ray, such as specificity, sensitivity, precision, and accuracy, should be kept as high as feasible, while negative measurements, such as FPR and FNR, should be kept as low as possible.

### 7.2 Performance analysis

On analyzing the acquired outcomes the improved deep ensemble model +CGHFA has recorded the highest accuracy for every variation in the learn\_rate. The results acquired are shown in Table 1- Table 5, respectively. The detection accuracy recorded by the improved deep ensemble model +CGHFA at learn\_rate=60 is 90.7%, at learn\_rate=70 is 92.14%, at learn\_rate=80 is 93.482% and at learn\_rate=90 is 94.8%. The accuracy of the improved deep ensemble model +CGHFA has attained higher value when compared with the conventional models SVM, RF, DCNN[1], ELM+HMBA[2], BI-LSTM, BI-GRU, DBN, CSOA+ improved deep ensemble model, SSA+ improved deep ensemble model, BOA+ improved deep ensemble model, AHM+ improved deep ensemble model, and CSA+ improved deep ensemble model. The major reason behind the enhancement in the detection performance with the improved deep ensemble model +CGHFA is due to the inclusion of the improved deep ensemble model for making the ultimate decisions regarding the presence/ absence of osteoporosis. Improvements with the improved deep ensemble model +CGHFA are due to the selection of the optimal features for training the classifier that makes the detections. In addition, the improved deep ensemble model +CGHFA have recorded the least error value for every variation in the learn\_rate. Thus, the objective of minimization of detection errors is said to have been achieved.

### 7.2 Convergence analysis

In this research work, a new optimization model referred as CGHFA (AHM+CSA) is introduced to select the optimal features for training the model. The objective function of this research work is the minimization of detection accuracy (i.e. classification accuracy). The model that recorded the minimal cost function is hence said to be the optimal one. On analysing the acquired outcomes, the projected model has recorded the minima cost function that the existing models at the highest iteration count. Thus, it is said to have achieved the objective function. At the 50<sup>th</sup> iteration, the cost function recorded by the projected model is 1.028, which is the least value and the most favourable value than the existing models (CSAO=1.035, SSA=1.034, BOA=1.0343,

Table 1. Performance analysis of the projected osteoporosis detection model in terms of accuracy

Learn_rate (%)	SVM	RF	DCNN[1]	ELM+HMBA[2]	BI-LSTM	BI-GRU	DBN	CSOA+ improved deep ensemble model	SSA+ improved deep ensemble model	BOA+ improved deep ensemble model	AHM+ improved deep ensemble model	CSA+ improved deep ensemble model	improved deep ensemble model +CGHFA
60	78.99	77.73	86.67	86.45	78.58	83.99	85.01	78.08	81.79	85.70	83.37	88.15	90.72
70	82.48	79.06	88.66	87.54	79.64	86.08	85.70	79.07	83.55	85.95	84.60	90.65	92.42
80	86.36	79.69	90.23	88.78	79.96	87.45	87.57	80.90	85.84	87.31	85.15	91.98	93.48
90	87.67	80.96	91.93	90.34	82.05	89.10	88.92	83.01	86.96	88.99	86.85	93.29	94.86

Table 2. Performance analysis of the projected osteoporosis detection model in terms of sensitivity

Learn_rate (%)	SVM	RF	DCNN[1]	ELM+HMBA[2]	BI-LSTM	BI-GRU	DBN	CSOA+ improved deep ensemble model	SSA+ improved deep ensemble model	BOA+ improved deep ensemble model	AHM+ improved deep ensemble model	CSA+ improved deep ensemble model	improved deep ensemble model +CGHFA
60	78.87	80.78	83.00	85.36	73.06	85.24	83.05	72.18	81.89	80.48	77.86	88.29	89.67
70	80.55	85.89	87.78	86.43	74.87	87.03	85.20	73.91	83.84	80.90	79.93	88.68	90.78
80	82.17	86.01	89.81	88.30	75.40	88.68	87.23	76.93	85.86	83.15	80.85	90.62	92.56
90	88.31	87.47	91.62	89.99	78.73	88.78	88.16	80.19	87.55	85.79	83.53	92.19	94.55

Table 3. Performance analysis of the projected osteoporosis detection model in terms of specificity

Learn_rate (%)	SVM	RF	DCNN[1]	ELM+HMBA[2]	BI-LSTM	BI-GRU	DBN	CSOA+ improved deep ensemble model	SSA+ improved deep ensemble model	BOA+ improved deep ensemble model	AHM+ improved deep ensemble model	CSA+ improved deep ensemble model	improved deep ensemble model +CGHFA
60	78.87	80.78	83.00	85.36	73.06	85.24	83.05	72.18	81.89	80.48	77.86	88.29	89.67
70	80.55	85.89	87.78	86.43	74.87	87.03	85.20	73.91	83.84	80.90	79.93	88.68	90.78
80	82.17	86.01	89.81	88.30	75.40	88.68	87.23	76.93	85.86	83.15	80.85	90.62	92.56
90	88.31	87.47	91.62	89.99	78.73	88.78	88.16	80.19	87.55	85.79	83.53	92.19	94.55

Table 4. Performance analysis of the projected osteoporosis detection model in terms of precision

Learn_rate (%)	SVM	RF	DCNN[1]	ELM+HMBA[2]	BI-LSTM	BI-GRU	DBN	CSOA+ improved deep ensemble model	SSA+ improved deep ensemble model	BOA+ improved deep ensemble model	AHM+ improved deep ensemble model	CSA+ improved deep ensemble model	improved deep ensemble model +CGHFA
60	82.71	78.34	87.34	87.11	76.72	84.64	83.61	75.94	83.29	84.06	81.51	88.84	93.82
70	84.46	78.61	89.28	88.18	78.31	86.68	85.34	77.47	85.68	84.42	83.34	91.28	94.11
80	86.41	80.25	90.72	89.31	78.77	87.92	86.25	80.11	85.79	86.39	84.13	92.49	94.11
90	88.92	81.40	92.19	90.69	81.66	89.34	87.90	82.91	88.97	88.67	86.46	93.07	95.13

AHM=1.033 and CSA=1.032). The major reason behind the enhancement in convergence perform of CGHFA is the utilization of the circle map and cat-guided solution updating. The convergence speed of the solutions. The results acquired are shown in Fig. 13.

### 7.3 Analysis of projected model: with Vs without feature selection

The feature selection has a significant impact on the model's computational complexity. When the features are picked by PCA in the projected model, the detection accuracy is just 90.5 percent. Furthermore, the detection accuracy of 88.6 percent and 90.9

Table 5. Performance analysis of the projected osteoporosis detection model in terms of FNR

Learn_rate (%)	SV M	RF	DCNN[1]	ELM+HMBA[2]	BI-LSTM	BI-GRU	DBN	CSOA+ improved deep ensemble model	SSA+ improved deep ensemble model	BOA+ improved deep ensemble model	AHM+ improved deep ensemble model	CSA+ improved deep ensemble model	improved deep ensemble model+CGHFA
60	21.13	19.22	17.00	14.64	26.94	14.76	16.95	27.82	18.11	19.52	22.14	11.71	10.33
70	19.45	14.11	12.22	13.57	25.13	12.97	14.80	26.09	16.16	19.10	20.07	11.32	9.22
80	17.83	13.99	10.19	11.70	25.13	11.32	12.77	23.07	14.14	16.85	19.15	9.38	7.44
90	11.69	12.53	8.38	10.01	25.13	11.22	11.84	19.81	12.45	14.21	16.47	7.81	5.45

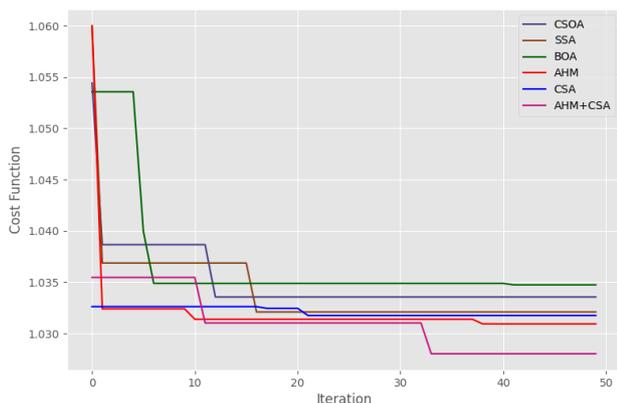


Figure. 13 Convergence analysis

Table 6. Analysis on the projected model: feature extraction and selection

Metrics	Proposed work with PCA	Proposed work with LDA	Proposed work with chi-square	Proposed work without feature selection	Proposed work with optimal feature selection
Specificity	0.913	0.843	0.904	0.911	0.936
Sensitivity	0.899	0.702	0.866	0.908	0.921
Precision	0.913	0.747	0.890	0.915	0.935
NPV	0.899	0.811	0.883	0.904	0.921
MCC	0.839	0.552	0.771	0.878	0.856
FPR	0.087	0.157	0.096	0.089	0.064
FNR	0.101	0.298	0.134	0.092	0.079
F1-score	0.906	0.724	0.877	0.911	0.928
Accuracy	0.906	0.787	0.886	0.910	0.928

Table 7. Analysis on feature size

Approach	Feature size
Before Feature Selection	760
After Feature Selection with PCA	557
After Feature Selection LDA	4
After Feature Selection with Chi-square	76
After Feature Selection with CGHFA	423

Table 8. Statistical analysis

Method	Standard Deviation	Mean	Median	Max	Min
CSOA+ improved deep ensemble model	0.004	1.035	1.034	1.054	1.034
SSA+ improved deep ensemble model	0.004	1.034	1.032	1.060	1.032
BOA+ improved deep ensemble model	0.006	1.037	1.035	1.054	1.035
AHM+ improved deep ensemble model	0.004	1.032	1.031	1.060	1.031
CSA+ improved deep ensemble model	0.000	1.032	1.032	1.033	1.032
CGHFA+ improved deep ensemble model	0.003	1.031	1.031	1.035	1.028

percent can only be obtained when the features are picked using LDA and the Chi-square model. When the best characteristics are chosen, however, the projected model has the best detection accuracy of 92.8 percent. This also helps to reduce the model's computational complexity. Furthermore, the smallest mistakes have been captured using appropriate feature selection. This is, in fact, one of the reasons behind the decrease in categorization errors (as only relevant features have been selected for training the classifier that makes the detection process). The results acquired are shown in Table 6.

**7.4 Analysis on projected model: size of selected features**

The optimal feature selection is suggested as an approach for reduction in the detection errors. The size of optimal features selected for training the classifier model is manifested in Table 7.

## 7.5 Statistical analysis

The statistical performance recorded by the projected model is shown in Table 8. On analyzing the outcomes, the projected model has recorded the minimal mean value (since, objective function is minimization function). The mean value recorded by the projected model is 1.03097996, which is indeed the least value while compared to CSOA+ improved deep ensemble model= 1.0350964, SSA+ improved deep ensemble model= 1.03408544684, BOA+ improved deep ensemble model= 1.0368292, AHM+ improved deep ensemble model= 1.0320322783 and CSA+ improved deep ensemble model= 1.032102898. The selection of the improved features, and the improved deep learning model for classification, has brought this improvement.

## 7. Conclusion

In this research work, a novel osteoporosis detection model has been developed. Initially, the collected raw data was pre-processed via Gabor filtering (noise removal) and histogram equalization (quality enhancement). From the pre-processed data, the features like active shape model (ASM), active appearance model (AAM), GLCM, improved local gradient pattern (I-LGP), and improved median ternary pattern (I-MTP) has been extracted. Subsequently, the optimal features has been selected from the extracted features using a new hybrid optimization model. The projected hybrid optimization model has been the conceptual blend of the standard artificial hummingbird algorithm (AHM) [34] and cat hunting optimization algorithm (CHOA) [35], respectively. Using these optimally selected features, the deep learning classifiers in the Osteoporosis detection phase has been trained. The Osteoporosis detection phase has been modeled with the newly constructed ensemble-of-classifiers model. Here, deep learning classifiers like quantum deep neural network (QDNN), improved deep convolution neural network (I-DCNN) and recurrent neural network (RNN) has been used. All these classifiers has been trained using the optimally selected features.

## Conflict of interest

"The authors declare no conflict of interest".

## Author contributions

The paper conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing,

visualization, have been done by 1<sup>st</sup> author. The supervision and project administration, have been done by 2<sup>nd</sup> author

## References

- [1] B. Zhang, K. Yu, and S. Zhang, "Deep learning of lumbar spine X-ray for osteopenia and osteoporosis screening: A multicenter retrospective cohort study", *Bone*, Vol. 140, p. 115561, 2020.
- [2] D. Devikanniga, "Diagnosis of osteoporosis using intelligence of optimized extreme learning machine with improved artificial algae algorithm", *International Journal of Intelligent Networks*, Vol. 1, pp. 43-51, 2020.
- [3] M. Mebarkia, A. Meraoumia, L. Houam, S. Khemaissia, B. Amel, and R. Jennane, "Hierarchical Multiscale Local Binary Pattern For Better Osteoporosis Detection", In: *Proc. of 2021 IEEE 1st International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering MI-STA*, pp. 387-392, 2021.
- [4] E. Tejaswini, P. Vaishnavi, and R. Sunitha, "Detection and prediction of osteoporosis using impulse response technique and artificial neural network", In: *Proc. of 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 1571-1575, 2016.
- [5] C. Reshmalakshmi and M. Sasikumar, "Trabecular bone quality metric from X-ray images for osteoporosis detection", In: *Proc. of 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT)*, pp. 1694-1697, 2017.
- [6] L. Darlea, C. Vertan, M. Ciuc, and I. Stefan, "On the influence of image enhancement on fractal-based automatic osteoporosis detection from calcaneum x-rays", In: *Proc. of International Symposium on Signals, Circuits and Systems, ISSCS*, Vol. 1, pp. 43-46, 2010.
- [7] Y. Nasser, M. E. Hassouni, A. Brahim, H. Toumi, E. Lespessailles, and R. Jennane, "Diagnosis of osteoporosis disease from bone X-ray images with stacked sparse autoencoder and SVM classifier", In: *Proc. of 2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pp. 1-5, 2017.
- [8] R. Riandini, and M. K. Delimayanti, "Feature extraction and classification of thorax x-ray image in the assessment of osteoporosis", In: *Proc. of 2017 4th International Conference on*

*Electrical Engineering, Computer Science and Informatics (EECSI)*, pp. 1-5, 2017.

- [9] K. Ramaraj, G. Amiya, P. R. Murugan, V. Govindaraj, M. Vasudevan, and A. Thiyagarajan, "Sensors for Bone Mineral Density Measurement to Identify the Level of Osteoporosis: A Study", In: *Proc. of 2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, Tirunelveli, India, pp. 326-333, 2022.
- [10] S. Nakanislii, Y. Kinoshita, and M. Matsukawa, "Effect of medullary cavity on the two wave phenomenon in the distal part of long bone", In: *Proc. of 2017 IEEE International Ultrasonics Symposium (IUS)*, Washington, DC, USA, pp. 1-1, 2017.
- [11] N. R. Hernandez, M. C. H. Escareno, and J. R. M. Rendon, "Image Analysis Tool with Laws' Masks to Bone Texture", In: *Proc. of 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Cancun, Mexico, pp. 691-694, 2017.
- [12] W. Zhao, L. Wang, and S. Mirjalili, "Artificial hummingbird algorithm: A new bio-inspired optimizer with its engineering applications", *Computer Methods in Applied Mechanics and Engineering*, Vol. 388, p. 114194, 2022.
- [13] A. Ghaedi, A. K. Bardsiri, and M. J. Shahbazzadeh, "Cat hunting optimization algorithm: a novel optimization algorithm", *Evolutionary Intelligence*, pp.1-22, 2021.
- [14] D. Fan, M. Cao, N. Sun, and K. Liu, "Filtering ultrasonic echo of bottom sediment in drilling hole by parameters constrained gabor filter", In: *Proc. of 2009 9th International Conference on Electronic Measurement & Instruments*, pp. 4-230-4-233, 2009.
- [15] D. Fan, M. Cao, N. Sun, and K. Liu, "Filtering ultrasonic echo of bottom sediment in drilling hole by parameters constrained gabor filter", In: *Proc. of 2009 9th International Conference on Electronic Measurement & Instruments*, pp. 4-230-4-233, 2009.
- [16] S. Schreiber, A. Stormer, and G. Rigoll, "A Hierarchical ASM/AAM Approach in a Stochastic Framework for Fully Automatic Tracking and Recognition", In: *Proc. of 2006 International Conference on Image Processing*, pp. 1773-1776, 2006.
- [17] A. Kos, T. Zielinski, A. Skalski, P. Kukolowicz, P. Kedzierawski, and T. Kuszewski, "Comparison of ASM and AAM-based segmentation of prostate image in the CT scans for radiotherapy planning", In: *Proc. of 2012 Joint Conference New Trends In Audio & Video And Signal Processing: Algorithms, Architectures, Arrangements And Applications (NTAV/SPA)*, pp. 53-57, 2012.
- [18] W. Songpan, "Improved Skin Lesion Image Classification Using Clustering with Local-GLCM Normalization", In: *Proc. of 2018 2nd European Conference on Electrical Engineering and Computer Science (EECS)*, pp. 206-210, 2018.
- [19] Y. Boo, S. Shin, and W. Sung, "Quantized Neural Networks: Characterization and Holistic Optimization", In: *Proc. of 2020 IEEE Workshop on Signal Processing Systems (SiPS)*, pp. 1-6, 2020.
- [20] Y. Lu, Y. Shi, G. Jia, and J. Yang, "A new method for semantic consistency verification of aviation radiotelephony communication based on LSTM-RNN", In: *Proc. of 2016 IEEE International Conf. on Digital Signal Processing (DSP)*, pp. 422-426, 2016.