

# Classification of Osteoporosis in the Lumbar Vertebrae using L2 Regularized Neural Network based on PHOG Features

Kavita Avinash Patil<sup>1</sup>

Department of Electronics & Communication Engineering  
Sri Jagadguru Balagangadhranatha Institute of Technology  
Faculty, New Horizon College of Engineering,  
affiliated to Visvesvaraya Technological University  
Bangalore, India

K. V. Mahendra Prashanth<sup>2</sup>

Department of Electronics and Communication Engineering  
Sri Jagadguru Balagangadhranatha Institute of  
Technology, affiliated to Visvesvaraya Technological  
University, Bangalore  
India

Dr. A Ramalingaiah<sup>3</sup>

Sapthagiri Institute of Medical Sciences and Research Centre, M.S (Ortho) Orthopedic Doctor, India

**Abstract**—One of the most common bone diseases in humans is osteoporosis, which is a major concern for the public health. Osteoporosis can be prevented if it is detected at an early stage. The research agenda consists of two phases: pre-processing of X-ray images of the spine and analysis of texture features from trabecular bone lumbar vertebrae L1-L4 for detecting osteoporosis. The preprocessing involves image enhancement of texture features and co-register the images in order to segment the L1-L4 regions in the lumbar spine. Range filtering and Pyramid Histogram of Orientation Gradient (PHOG) are used to analyze texture features. Input images are filtered with a range filter to adjust the local sub range intensities in a specified window to detect edges. Then a PHOG algorithm is designed to determine both the local shape of an image texture and its spatial layout. Based on texture features of lumbar vertebrae L1-L4, classify them as normal or osteoporotic using neural network (NN) models with L2 regularization. In an experiment, X-ray images and dual-energy X-ray absorptiometry (DXA) reports of individual patients are used to verify the system. DXA reports describe a statistical analysis of normal and osteoporotic results. However, the proposed work is categorized according to the texture features as normal or osteoporotic. 99.34% classification accuracy is achieved; cross-validation of these classified results is done with the DXA reports. Diagnostic accuracy of the proposed method is higher than that of the existing DXA with X-ray. Further, the area under the Receiver Operating Characteristic (ROC) curve for L1-L4 had a significantly higher sensitivity for osteoporosis.

**Keywords**—Cross validation; image enhancement; lumbar spine; lumbar vertebrae; neural network; osteoporosis; PHOG; regularization; trabecular bone; texture features; x-ray images

## I. INTRODUCTION

Osteoporosis is a disease which affects the density and strength of bones. A bone's density is the amount of bones (bone mineral density-(BMD)), while its strength (quality) is the fibers in the bones. Osteoporosis leads to weaker, more porous and more brittle bones, as well as a greater risk of fractures [1]. This paper proposed an efficient model for

assessing osteoporosis in the lumbar vertebrae L1-L4 using L2 regularized neural network. The model consists of two parts, first part of the system is a pre-processing of an X-ray spine image in order to visualize the fine textures of the trabecular micro architecture of L1 - L4 vertebrae using the following steps: initially, the X-ray image needs to be enhanced by improving the visual quality, and then, the input images need to be aligned into the same plane by using co-registration, and finally, segment the image. As part of the segmentation process, the primary objective is to determine the region of interest (ROI) in the image, namely between the L1 and L4 vertebrae, in order to improve the actual segmentation. As illustrated in Fig. 1, the second part of the system uses PHOG texture features to detect osteoporosis in the L1-L4 lumbar vertebrae using NN with L2 regularization in order to classify normal and abnormal images.

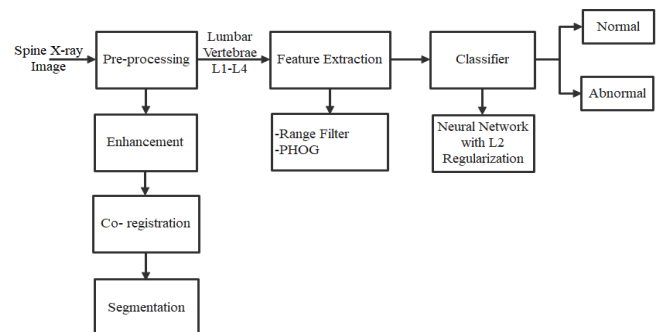


Fig. 1. An Overview of the System Model.

This paper makes the following contributions: The proposed technique for the detection of osteoporosis in the lumbar spine (L1-L4) is described in Section I. Section II outlines the related work of other methods that have been developed so far. The Section III Database detail description is used to test the system. A preprocessing stage was described in Section IV to improve image quality. An in-depth discussion of image texture feature extraction was presented in

Section V. NN networks are regularized by L2 in Section VI, used as a classifier. The Section VII presents a detailed simulation study and discussion. The final Sections VIII and IX of the report summarize the proposed system and the future work required to improve its performance, respectively.

## II. RELATED WORK

Osteoporosis is a disorder in which there is a loss of bone mass and abnormally degenerated bone architecture, especially, hips, spine, and wrist, known as osteoporosis [2]. Among osteoporotic fractures, spine fractures are the most common and are a major health concern among the elderly. Consequently, osteoporotic fractures require early diagnosis of patients at high risk. Typically bone density scans are used to detect osteoporosis, among which the dual energy X-ray absorptiometry (DXA) is considered as a common technique to measure bone size and bone mineral density (BMD). Various studies that deal with BMD measurements have been conceded by analyzing image texture features using an X-ray's. An easy and inexpensive method for diagnosing osteoporosis has been proposed based on analysis of image texture features using X-rays. In conjunction with a machine-learning algorithm, a fractal model was utilized to develop the software, using pixel variations for grey levels[3]. An osteoporosis patient's bone structure value (BSV) is estimated from BMD using spine radiograph images. It is necessary to conduct further studies to determine BSV's potential to be a reliable assessment of treatment effects and future fracture risk in individuals with osteoporosis and those without it. The first group had 83 patients treated for osteoporosis alone, while the second group had 76 patients treated for both osteoporosis and lumbar spinal stenosis (LSS) [4]. The T-scores over the first year as well as after a year, two years, and three years were confirmed. Two groups were compared on mean BMD and changes in BMD over three years. In addition, three-year BMD improvements were evaluated along with their relationship to initial BMD change and related factors. Study participants were given ibandronate dose for newly diagnosed osteoporosis, so LSS was examined for its effect on BMD. The study was focused only on whether LSS could improve bone mineral density in the treatment of osteoporosis, clinical outcomes related to osteoporosis treatment, such as osteoporotic fractures, were not assessed in the follow up. An alternative approach to determining the extent of bone loss caused by osteoporosis is fractal analysis when studying spinal CT (computed tomography) images [5]. Based on the results of the study and the K-NN (K-Nearest Neighbor) classifier, a computerized system based on CT images could assist physicians in making initial diagnoses in difficult cases. Overall the system provides 81% classification accuracy; an alternative method is required to improve the system performance. By using the Picture Archiving and Communication System (PACS), the lumbar vertebra was measured in QCT (Quantitative computed tomography) and the HU (Hounsfield unit) of its vertebral body in conventional CT [6,7]. The correlation between the T-score of conventional CT and the T-score of QCT was estimated using a multiple regression algorithm. Further, a logistic regression algorithm was applied to predict osteoporotic and non-osteoporotic vertebrae. With QCT data, the predictor modeling algorithm

estimated similar T-scores. In HU, similar results are observed as with QCT, with the exception of one osteoporotic vertebra that did not demonstrate discordance with an accuracy rate of 92.5% was recorded. The predictive accuracy will improve with more collected data. The purpose of study is to find out whether recurrent neural networks are capable of predicting osteoporotic fractures by analyzing spine images [8]. It explores the best design directions for such prediction models by experimenting with various network architectures. Transfer learning gives the advertised benefits, such as faster training speeds and greater suitability for larger datasets. By segmenting and finding vertebral edges, can diagnose the compression and locate the anomaly using Morphometric features and measurements using CT images [9,10]with 88.3% accuracy. There are challenges involved in finding the midpoint in the vertebral body and passing it to the next closest midpoint on the vertebral body boundary to analyse 3D textures, extracted the gray-level co-occurrence matrix Haralick, Wavelets (WL), local binary patterns (LBP), histogram of gradients (HoG), and harmonic alternator patterns (HAR) [11]. The texture-features and vBMD data extraction, fractured vertebrae were excluded. An assessment of prevalence of osteoporosis fourfold cross-validation was conducted to evaluate vertebral fractures. There is a correlation between vertebral level parameters and classification results. Mark-Segmentation-Network and Bone-Conditions-Classification-Network are used to analyze diagnostic CT images to automatically detect bone conditions [12]. The system achieves receiver operating characteristic curve area of 0.9167 and accuracy of 76.62%. Feature extraction from lumbar vertebra CT images as well as other clinical characteristics might be relevant to the diagnosis of the bone condition to improve the system performance. As compared with the traditional Osteoporosis Self-Assessment Tool for Asians (OSTA) model, the ANN, SVM, RF, and LoR models performed significantly better in both men and women [13].

## III. DATABASE DETAILS

According to Table I, 80 numbers of samples are used in the process of developing the proposed system. Initially Digital X-ray images of the lumbar spine (lateral view) are taken for processing. Dr. A Ramalingaiah, Orthopedic, provided the images from Abhilasha orthopedic hospital in Banashankari at No.271,3rd Stage, 5th Block, 100ft Road, Bangalore; the lumbar spine of 80 subjects in 2D, JPEG format. An experiment has been conducted to test the system by collecting 20 control subjects (normal) as well as 60 pathological (osteoporotic or abnormal) X-ray images, as well as DXA reports on the same people. In the DXA report, the L1-L4 lumbar spine statistical analysis status details are provided for each person.

TABLE I. RAW DATA BASE

Lumbar spine X-ray image	No. of Subjects (80)	DXA reports (80)
Controlled Subjects (Normal)	20	20
Osteoporotic Patients	60	60

Among 80 subjects, the segmentation algorithm correctly considers 18 (normal) out of 20 and 58 out of 60 (abnormal), which means  $18+58=76$ . A total of 304 ROI images of lumbar spine images (L1-L4) are extracted from each subject, and  $18 \times 4 = 72$  and  $58 \times 4 = 232$  results in 304 ROI images.

#### IV. PREPROCESSING STAGE

In the preprocessing stage, an adaptive histogram equalization (AHE) and a co-registration are used to obtain L1-L4 region of interested sub images with enhancement.

##### A. AHE

An image enhancement technique is adapted in this proposed work to provide a more interpretable image, for better quality input to the next phases of the work. A contrast-limited (CL) AHE process is used to enhance the contrast of grey scale image. Rather than processing the entire image, CLAHE focuses on small regions ( $8 \times 8$ ) of it, called tiles. Divides an image into a number of rectangular contextual regions. It calculates for each region independently by AHE. By enhancing the contrast of each region, the output histogram approximates by the distribution of its parameters. The Rayleigh distribution is used to create a contrast transform function, which depends on the input image type. Using this distribution, imagery appears more natural. Bilinear interpolation [14] is then used to combine the neighbouring blocks, eliminating artificial boundaries. Contrast can be reduced in homogeneous areas to prevent noise from being amplified. With clip-limit (0.01), further prevent the oversaturation of homogeneous areas of an image. In such areas, the histogram of an image is dominated by a high peak because many pixels are within the same gray scale range. The spine lateral view enhanced image (b) and the input image (a) as shown in Fig. 2.

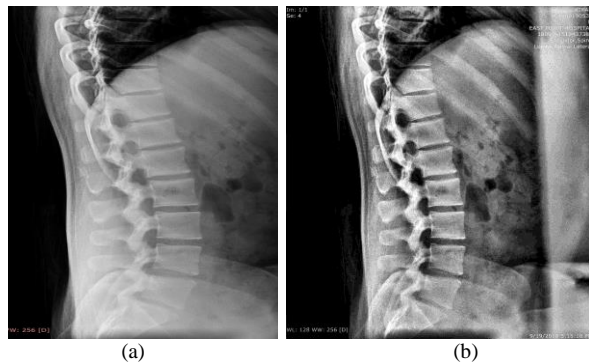


Fig. 2. (a) Input Image (b) Enhanced Image.

The algorithm-1 comprises into three major components in CLAHE: tile generation, histogram equalization, and bilinear interpolation.

---

##### Algorithm 1

Input:  $M \times N$  image, output: Enhanced image

1. Sections (sub image) are divided in the input image. These sections are referred to as tiles of  $N \times N$ .
  2. Obtain the histogram distribution for each tile.
  3. Minimum and maximum values are found through binary search based clipping [15] in the distribution and clipped portions are redistributed.
  4. In the next phase, bi-linear interpolation is carried out to combine the neighbouring blocks. The resultant one is an
- 

enhanced image with greater contrast.

5. The entropy [16]  $H$  is calculated using (1)

$$H = -\sum (h_c \times \log(h_c)) \quad (1)$$

In this case,  $h_c$  is the histogram counts.

6. Ensure  $N \leftarrow N + 2$ .

7. In order to measure the entropy  $H_{i+1}$  of the contrast-enhanced image, Where  $i = 1, \dots, N$  repeat steps 2 through 5 with a new window size.

8. If  $H_{i+1} < H_i$  stop, otherwise repeat  $N < 8$ .

9. It is certain that the enhanced image with the largest  $H$  value will also have better quality, and enhanced image is obtained at the output.

10. According to the experimental results, binary search has been shown to be an effective algorithm in terms of clipping and redistributing the pixels.
- 

##### B. Co-registration

Several biomedical imaging applications require the co-registration of images. The images may be obtained from the same sensor or from different sensors, and the spatial resolution might be the same or different. During co-registration, all images in the series are aligned spatially so that any feature in one image overlaps as well as its footprint in every other image [17,18]. An image is typically selected as the reference to which all other images are aligned during co-registration. The best reference image can be selected from the database, which contains all the features of lumbar spine lateral view. The co-registration process involves identifying common features in the reference and warping the other images that is the ones to be co-registered [19]. In the process, tie points are used to determine the locations of common features. The warped image is aligned to the reference using a polynomial function after enough tie points have been generated.

In Fig. 3, an example of the output of a co-registration process as illustrated. In part (a), the reference image is the normal image that all the target images are aligned to the same feature plane; the target image (b) does not resemble the reference image, as it represents osteoporosis; c) is overlaid image of R and co-registered, the co-registered image is completely aligned with the reference image using Geoscience extended flow optical lucas- kanade iterative (Gefolki) is a co-registration software [20,21]; overlays R and C in different colour bands to form a composite RGB image. A grey region indicates the same intensity of the two images in the composite image. Magenta and green colour regions indicate differences in intensities. (d) An image co-registered with the R image that shows the exact alignment of the T features.

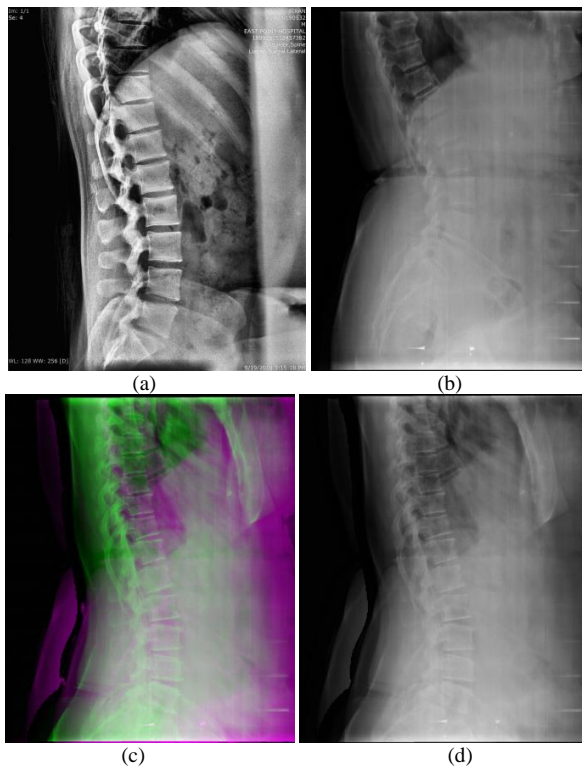


Fig. 3. (a) Reference (R) Image (b) Target (T) Image (c) Overlaid R and Co-Registered (C) Image (d) Co-registered Image.

#### Algorithm 2

A pixel alignment of two images can be divided into three steps: initialization of input image R and T of the same size, calculation of GeFolki flow, and resampling.

1. Homothety, rotation and scaling is used to make size of T to size R, denoted as T'.
2. GeFolki flow, the transfer from coordinates of T' to R by a matrix W.
  - W is composed of the y-displacement component (column) and the x-displacement component (row) for applying on the T' image.
  - In the software, GeFolki function takes as arguments T', R, and parameters, with order as radius = 64; -8:16, level = 4, iteration = 5, contrast\_adapt=false, and rank=4 as parameters.
  - A variety of radius sizes can be tested iteratively; it is a decreasing vector starts from biggest (64) radius to the smallest (16) one in steps of 8. The algorithm is more robust when the radius is large. Radii must be selected as small as possible when the flow on the image changes rapidly.
  - For the purpose of finding large displacements, create a pyramidal structure of the down sampled images in different levels. In a pyramid, the number of levels affects the size of the movements.
  - The total number of iterations required to run the gradient method for the minimum search.
  - Set contrast\_adapt is false for homogeneous images to look contrast inversion.
  - Changing the intensity value of T' with lower rank of pixel in the neighborhood within the specified window size (9×9).
3. Finally, transform the T' image to superimpose it with the R image. By resampling the image on a new coordinate grid, this operation creates the flow matrix W. Resampling is done using bilinear interpolation.

#### C. Region L1-L4 Sub Image Selection

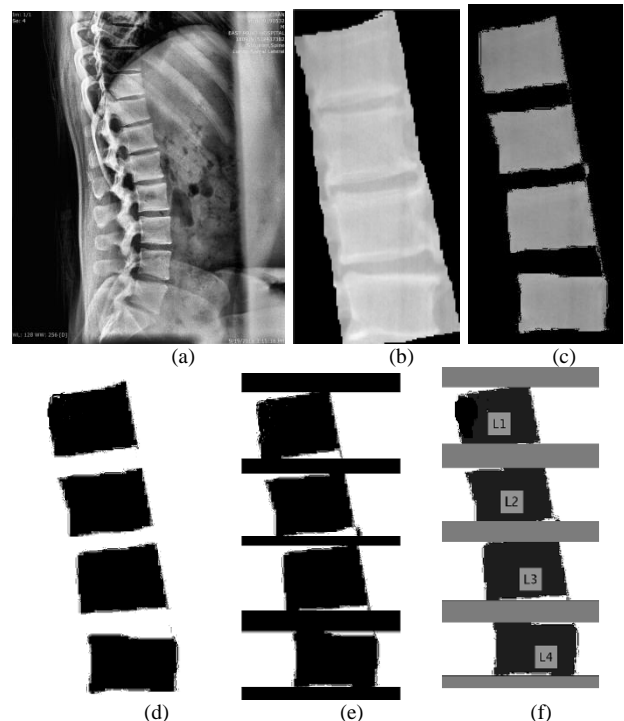
L1-L5 are the five vertebrae that make up the Lumbar Spine. As the largest vertebrae in the body [22][23], the lumbar spine bears the most weight. In contrast to the thoracic spine, Lumbar spines region have more range of motion.

There is limited rotational movement in the facet joints of the lower back. In this work L1-L4 is the region of interested subimage. An illustration of the interested region selection process is presented in Fig. 4 of (a)–(g) and explained in Algorithm 3.

#### Algorithm 3

1. The lateral spine view X-ray images are read from the input.
2. Obtain the border coordinates of the freehand drawn region from L1-L4, and divide the result by an array of 2-dimensional matrices that is subimage
3. Using the subimage object, produce a binary image mask.
  - It returns a mask that is associated with the subimage object B over the target image.
  - The target image must be contained within the same axes as the subimage.
  - Mask is a logical image the same size as the target image.
  - Mask is false outside the region of interest and true inside.
  - Multiplied mask with the input image to produce segmented L1-L4 region.
4. Create the inverted binary image from segmented image  
 segmented image pixels >255;  
 Seg\_inv=0;
 

```
else
Seg_inv= segmented image;
make it into a proper binary image for foreground
Inv_binary_im = Seg_inv < max (Seg_inv (:))/2;
```
5. Create the isolated L1-L4 image with equal distance and equal image size
  - find the bar rows  
 bar\_rows = sum (Inv\_binary\_im, 2) > 0.9 \* size (Inv\_binary\_im, 2);
  - make a bar image  
 im\_bars = false(size(Inv\_binary\_im));  
 im\_bars (bar\_rows, :) = true;
  - remove the bars from the image  
 im\_nobars = Inv\_binary\_im & ~im\_bars;
6. Label each lumbar vertebrae by finding its centroid.
7. Select only the connecting pixels on the ROI images for L1, L2, L3, L4.



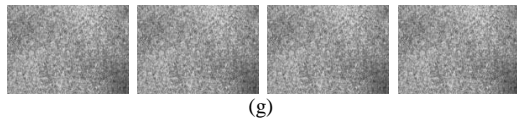


Fig. 4. Interested Region Selection Process (a) Original LS view Image (b) Cropped Image (c) Segmented Image (d) Inverted Binary Image (e) Isolated L1-L4 Image (f) Labelled Image (g) ROI Image of L1, L2, L3, L4.

## V. FEATURE EXTRACTION

The feature extraction process involves two stages; Range filter and PHOG. The range filter is used to adjust the local sub range intensities within the specified window size of  $3 \times 3$  neighborhood pixels in the input image. In a local sub-range the pixels are analyzed according to their statistical range to detect edges.

The PHOG objective is to determine the local shape of an image texture and its spatial layout. By considering the orientations of edge distribution within the sub-image as well as the spatial arrangement, the image is tiled into regions at multiple resolutions to obtain texture features.

### A. Range Filter

It operates on morphological utilities such as dilation and erosion of the image to regulate the maximum and minimum pixel values in the specified window [24]. Subsequently, it utilizes the padding operation on these morphological utilities. In Fig. 5, (a) the input and its corresponding histogram are clearly visible, showing the right-hand side of the histogram occupied by the input image and (b) the filter output and its histogram shows that the pixel ranges are narrowing by connecting local neighbors as edges, which reveals more defined texture in the image. In Fig. 6(a) and (b), the range filter response is shown for an abnormal image. A normal and abnormal image can't be distinguished by looking at them. Range filtering connects local neighbors as edges to narrow the histogram of the input.

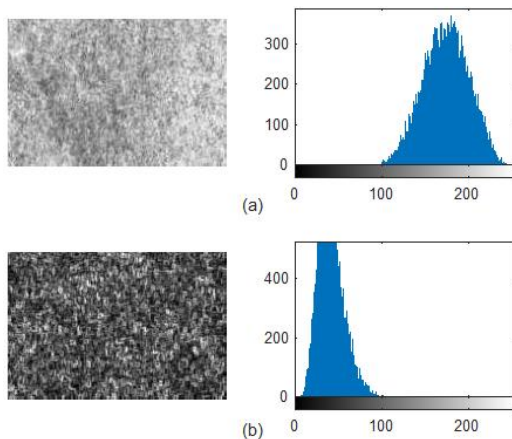


Fig. 5. Range Filter response for Normal Image (a) Input and Histogram (b) Output and Histogram.

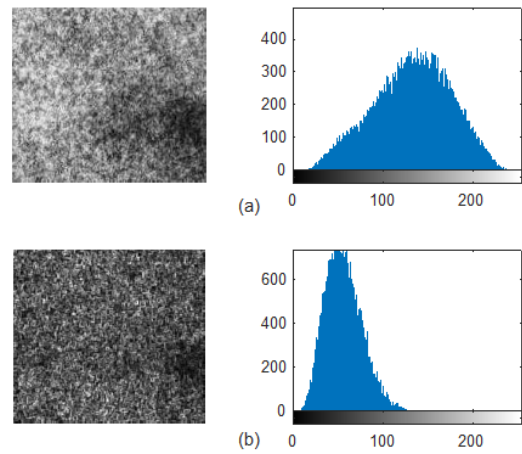


Fig. 6. Range Filter response for Abnormal Image (a) Input and Histogram (b) Output and Histogram.

### B. PHOG

By extracting pyramid histograms of oriented gradients (PHOG) from the image are useful to discriminate between normal and abnormal texture features. It was first presented in [25]. Using HOG, a range filter response image can be divided into small cells, HOGs for each cell are computed, normalized with block patterns, and each cell is given as a descriptor [26]. HOG features describe the outline and local structure appearance in the image by describing their distribution of intensity gradients. Using this method, intensity invariance is maintained by counting the appearance of gradient orientations. A descriptor is generated in four main steps:

- Computation of gradients and orientation.
- Bin Orientation.
- Block descriptors.
- Normalization of blocks.

1) *Computation of gradients and orientation*: In the first step, when an image sample is provided,  $f(i, j)$ , then the magnitude of the gradient,  $M(i, j)$ , and angle,  $\theta(i, j)$ , can be obtained using (2) and (3) by considering the pixel difference.

$$M(i, j) = \sqrt{(f(i+1, j) - f(i-1, j))^2 + (f(i, j+1) - f(i, j-1))^2} \quad (2)$$

$$\theta(i, j) = \tan^{-1} \left( \frac{f(i, j+1) - f(i, j-1)}{f(i+1, j) - f(i-1, j)} \right) \quad (3)$$

Histograms of orientations generated by gradients of samples within a region around the centered point are called orientation histograms. Histograms of orientations are divided into eight bins, each covering 360 degrees. In the histogram, samples are weighted according to their gradient magnitude.

2) *Bin orientation*: In the second step of the calculation, the cell histogram is created. The algorithm assigns a weighted vote to each pixel within each cell depending on the values found in the gradient calculation and its orientation. Depending on whether the magnitude gradient and its orientations from 0° to 360°, the histogram evenly distributed over 0° to 180° or 0° to 360°. The optimum number of bins in the histogram are considered as 8 bins as mentioned in [27]. Various bin number will be tested in this paper experimentally.

3) *Block descriptors*: For gradient strength to be normalized locally under different illumination conditions, cells must be organized into larger with the spatially connected blocks. HOG descriptors are then constructed as concatenated vectors of corresponding normalized cell histogram components compiled from all block regions. Each cell contributes to the final descriptor more than once if these blocks overlap. Block geometries consist of two main types: the rectangular R-HOG block and the circular C-HOG block. There are three parameters that characterize R-HOG blocks: the number of cells per block, the number of pixels per cell, and channel per cell histogram. Each R-HOG block is computed at one scale, but with an orientation not aligned. As well as encoding spatial form of information, R-HOG blocks are used. There are two types of circular HOG blocks (C-HOG): those with a single central cell and those with angularly divided central cells. The C-HOG blocks can be defined by four other factors: center bin radius, the additional bins radius expansion factor, and the number of radial bins. Shape context descriptors are similar to C-HOG blocks in appearance, but they differ greatly because they contain multiple orientation, whereas single edge presence count are used to formulate the shape contexts.

4) *Normalization of blocks*: Gradient magnitude varies greatly over a wide range due to variations in pixel intensity on a local level [28]. Therefore, effective local contrast normalization is essential for good performance. They typically involve separately normalizing each block of cells after grouping them into blocks and normalizing of each block is obtained by using L1-norm and L2 -norm illustrated in (4) and (6). L1- norm is defined as the sum of the vector's absolute values and L2 -norm as a square root of the sum of all squared vector values. Consider a non-normalized vector  $v_b$  which contains all histograms in a given block,  $\|v_b\|_k$  be its k-norm of the vector for  $k=1$  (L1-norm) and  $k=2$  (L2-norm) and  $e$  be the normalization constants should be small to prevent division by zero.

$$v_{b_{L1}} = \frac{v_b}{\|v_b\|_1 + e} \quad (4)$$

Square root of  $v_{b_{L1}}$  is described in (5), by using the Bhattacharya distance between pixels, descriptor vectors can be represented as a probability distributions.

$$\text{Square root of } v_{b_{L1}} = \sqrt{\frac{v_b}{\|v_b\|_1 + e}} \quad (5)$$

$$v_{b_{L2}} = \frac{v_b}{\sqrt{\|v_b\|_2^2 + e^2}} \quad (6)$$

L2- hys: By attenuating the maximum pixel values of  $v_b = 0.2$ . Renormalizing, hysteresis thresholding can be achieved for L2 norms [29]. Renormalize the univariate feature vector so the values in each are no larger than 0.2. This will limit the effect of large gradient magnitudes. As a result, it is less important to match the magnitudes for large gradients, while the distribution of orientations is more important.

Determine the HOG for Level-0 that indicates for entire image as shown in Fig. 7(a). Throughout the entire HOG operation,  $N=8$  bins are fixed. Each bin in the histogram represents how many pixels fall within a certain range of angles. In the next step, divide the image into four blocks, and compute HOG for each block as shown in Fig. 7(b). Repeat the process recursively until the depth  $L=3$  as shown in Fig. 7(c) and (d). Several levels of pyramids are used to divide the input image, histogram composed of  $N$  bins, and Level-0 is characterized by a  $N$ - vector. An image with Level-1 is described as a  $4 \times N$ -vector. The PHOG descriptor represented using (7).

$$D_{PHOG} = N \sum_{l=0}^3 4^l \quad (7)$$

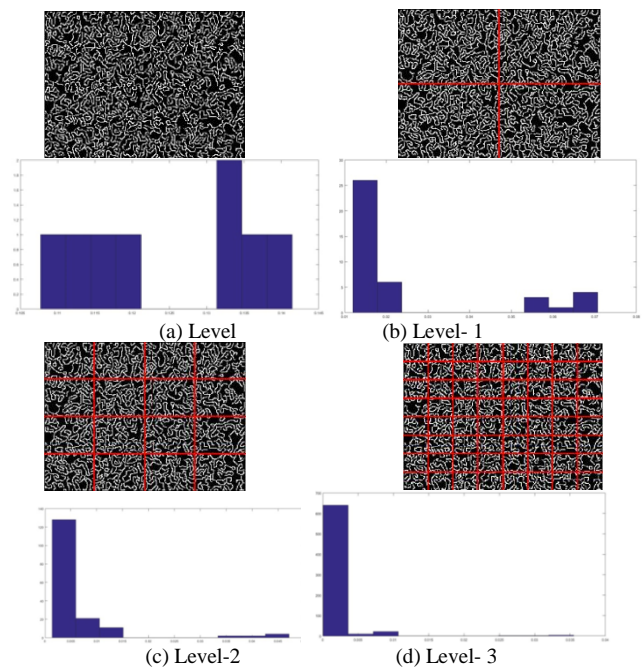


Fig. 7. PHOG Schematics at different Resolution Levels.

Create one vector by concatenating the histograms with the same level. The final PHOG descriptor of an image is obtained by concatenating all the vectors in each pyramid level that pertain to the spatial properties of an image. By setting  $N = 8$  and  $L = 3$ , the PHOG descriptor ( $D_{PHOG}$ ) has a length of  $8 * (1 + 4 + 16 + 64) = 680$ .

## VI. CLASSIFIER

Data can also be classified into identifiable groups or features using neural networks [30]. A neural network based classification becomes a very powerful prediction network. It achieves two goals: Lowering the prediction error by training and testing similar patterns of behavior; and minimizing the number of training samples within a training data set, resulting in more efficient training.

### A. Neural Networks with L2 Regularization

A technique known as L2 regularization of neural networks (NNs) reduces the likelihood of the model overfitting, thus improving its performance. NNs are mathematical functions used to train the network and make predictions. The training refers to finding values for weights and biases needed to define the NN. Excessive training of a neural network may lead to model overfitting. Thus, when a neural network model is trained to predict the values from used training data, it does so with little error and high accuracy, but when the model is applied to new, previously unknown data, it does not work well. Parameter optimization is used to determine weights and biases in a way that minimizes error between computed and true results.

1) *L2-Regularization*: During L2 regularization, a new regularization term was added to the loss function of the neural network [31]. As part of regularization, the proposed work utilizes Euclidean-Norm or L2-Norm for the weight matrices, which is the sum of all the squared weight values of any given matrix. By dividing the scalar regularization parameter  $\lambda$  by two, add the regularization term to the regular loss function. The following steps are used to implement the neural network with L2 regularization.

2) *Activation function*: The logistic regression [32] hypothesis is defined as in (8)

$$h_{\theta}(x) = g(z); z = \theta^T \times x \quad (8)$$

$\theta$  – weights and  $x$  -input

The sigmoid function  $g$  is defined as in (9).

$$g(z) = \frac{1}{1 + e^{-z}} \quad (9)$$

3) *Cost function*: Probabilities of an observation can be predicted with the cost for the observation, with a minimum amount of error. A given observation  $x$  has a cost (error) associated with it if the class label is  $z$  can be defined in (10).

$$Cost(h_{\theta}(x), z) = \begin{cases} -\log(h_{\theta}(x)) & ; \text{if } z = 1 \\ -\log(1 - h_{\theta}(x)) & ; \text{if } z = 0 \end{cases} \quad (10)$$

As a result, the total cost for all the  $m$  observations in a dataset is given in (11).

$$J(\theta) = \frac{1}{m} \sum_{i=1}^m Cost(h_{\theta}(x^{(i)}), z^{(i)}) \quad (11)$$

$(x^{(1)}, z^{(1)}), (x^{(2)}, z^{(2)}), \dots, (x^{(m)}, z^{(m)})$  are the  $m$  number of training datasets.

4) *Cost function and gradient*: The cost function in logistic regression is defined in (12).

$$J(\theta) = -\frac{1}{m} \left[ \sum_{i=1}^m z^{(i)} \log(h_{\theta}(x^{(i)})) + (1 - z^{(i)}) \log(1 - h_{\theta}(x^{(i)})) \right] \quad (12)$$

Regularization L2 helps to avoid over fitting in the model by adding a penalizing component for high weights. Equation (13) gives the regularized cost function for logistic regression.

$$J(\theta)_{new} = \frac{1}{m} \sum_{i=1}^m Cost(h_{\theta}(x^{(i)}), z^{(i)}) + \left( \sum_{j=1}^n \|\theta_j\|^2 \right) \frac{\lambda}{2m} \quad (13)$$

- $\left( \sum_{j=1}^n \|\theta_j\|^2 \right) \frac{\lambda}{2m}$  is the regularization term.
- $\lambda$  is the regularization factor.
- $\frac{1}{m} \sum_{i=1}^m Cost(h_{\theta}(x^{(i)}), z^{(i)})$  loss function.

The gradient of the cost function is a vector where the  $i^{\text{th}}$  element is defined as follows:

$$\frac{\partial J(\theta)}{\partial \theta_{(i)}} = \frac{1}{m} \left[ \sum_{j=1}^m (h_{\theta}(x^{(j)}) - z^{(j)}) x_i^{(j)} \right]; \text{ for } i = 0$$

$$\frac{\partial J(\theta)}{\partial \theta_{(i)}} = \frac{1}{m} \left[ \sum_{j=1}^m (h_{\theta}(x^{(j)}) - z^{(j)}) x_i^{(j)} + \lambda \theta_{(i)} \right]; \text{ for } i \geq 1$$

5) *Learning parameter optimization*: In order to efficiently train NN, gradient descent can be implemented to determine the optimal parameters for a regression model [33]. During training and testing, the accuracy of the algorithm was maximized, while the computation time was reduced through parametric  $\lambda$ ,  $\theta$  optimization and number of iterations. Calculate the optimal cost function  $J(\theta)$  according to parameters  $\theta$ . Determine the optimal parameters  $\theta$  for the logistic regression cost function based on a fixed dataset of  $x$  and  $z$  values. For a given dataset  $(x, z)$ , computes the logistic regression cost and gradient based on a training data and a specific value of  $\theta$ . Lastly, it gives the cost and  $\theta$  values. After the trained data is plotted, the decision boundary will be based on the final  $\theta$  value as shown in Fig. 8.

6) *Prediction*: A NN's hidden layers make it better suited for predictive analytics. Predictions are based on the output and input nodes of linear regression models. Consequently, they are cost-prohibitive because they require enormous computing power [34]. Furthermore, neural networks perform best when trained with extremely large data sets. By using supervised learning techniques such as data clustering, each neuron's weight can be determined. In the supervision training process, by feeding sample inputs and outputs to the algorithm, weights are derived until the inputs and outputs are closely matched, that is, weights are optimized. Using a sigmoid activation function on the output layer, it is easy to make predictions for the input test, so predicted label will be either 0 or 1.

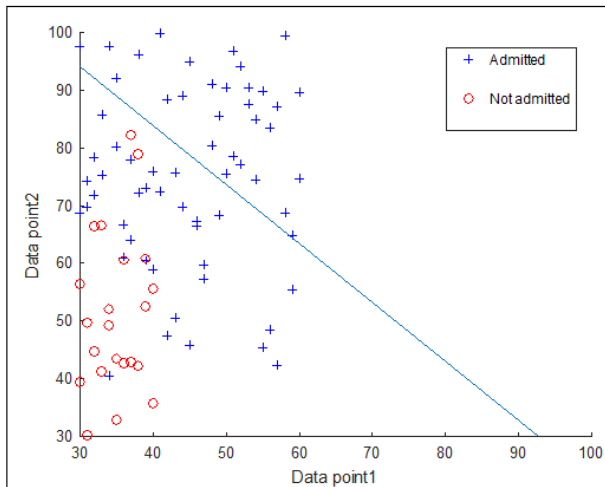


Fig. 8. Trained Data with Decision Boundary.

### VII. RESULT AND DISCUSSION

An experimental result is carried out in three stages: A) preprocessing B) feature extraction C) classification.

#### A. Preprocessing Stage

There are 80 X-ray lumbar spine images collected; 80 x 4 =320, vertebrae of L1-L4 are the actual ROI images, but in reality only 76 ROI images have been correctly segmented (76 x 4 =304), where all four segmentations (L1-L4) are correct in 76 images. Table II shows the true and false segments of L1-L4. Out of 80 images, 77 reflect true segmentation of L1, while 3 reflect false segmentation. The results for L2 were 78 true and 2 false, L3 was 77 true and 3 false, and L4 was 76 true and 4 false.

As shown in Table III, the data base can be used to detect osteoporosis in the lumbar vertebrae L1 - L4. A total of 76 correctly segmented images are considered for the processing (76 x 4 = 304). For testing the system, a 2-fold cross validation is applied. This is accomplished by using 152 samples for supervised samples and 152 samples for testing.

#### B. Feature Extraction Stage

The texture features were extracted from 304(L1-L4) samples of ROI images using PHOG. For N=8 bins, Table IV shows the levels and the corresponding descriptors.

Concatenated level descriptors are used to classify samples based on their attributes.

#### C. Classification Stage

Table V shows the osteoporosis detection classification accuracy based on the concatenation of all the levels descriptor are 696 with parameter optimization lambda varying from 0 to 1. If lambda is greater than 1, the system takes more iteration and becomes more complicated. Consequently, it was determined that for classification accuracy, when lambda is higher, the likelihood of over fitting during training is reduced, and gives maximum accuracy. During training, the model was iterated 27 times, with  $\lambda = 0.9$ . In large-scale implementations, a small change in the parameter can produce more effects. By taking the 184 and 40 descriptor with parameter optimization lambda varying from 0 to 1, the classification accuracy is reduced by reducing the size of the descriptor. Fig. 9 displays the graphical presentation of the system's osteoporosis detection classification accuracy with lambda. In order to improve the accuracy, lambda plays a vital role in terms of number of iterations and computation time.

TABLE II. SEGMENTATION OF L1-L4

ROI Segmentation	L1	L2	L3	L4
True	77	78	77	76
False	3	2	3	4

TABLE III. DETAILS OF THE EXPERIMENTAL DATA BASE

Subjects	No. of Subjects (76)	No. of ROI samples (L1-L4=304)	DXA reports	No. of supervised samples (152)	No. of testing samples (2-fold =152)
Controlled	18	72 (18x4)	18	36	36
Osteoporotic	58	232 (58x4)	58	116	116

TABLE IV. DETAILS OF THE LEVELS AND DESCRIPTORS

No. of levels	No. of descriptor	Concatenated levels	No. of descriptor
L0	8	-	-
L1	32	L0+L1	40
L2	144	L0+L1+L2	184
L3	512	L0+L1+L2+L3	696

TABLE V. SYSTEM ACCURACY WITH PARAMETER OPTIMIZATION

Lambda values	Instances of iteration	Accuracy in Training set (%)	Accuracy of Test sets (%)	Computed time (in seconds)
0.1	25	96.89	99.01	19
0.3	25	97.45	99.05	21
0.6	26	98.56	99.12	22
0.7	27	100	99.34	25
0.9	27	100	99.34	25



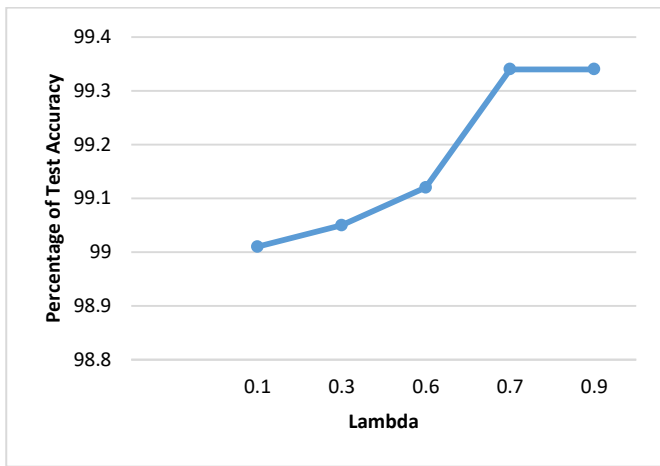


Fig. 9. Graphical Representation of Accuracy with Lambda.

A confusion matrix with multiple classifiers can be used to evaluate system performance as shown in Table VI. Testing the proposed system is based on a detailed analysis of the database, as shown in Table III. With L2 regularization, NN has fewer False Positives (FP) and False Negatives (FN), which leads to less overfitting and more accuracy, flexibility, and works well in classifying events. When compared with other classification techniques, the overall performance of the system leads to better True Positives (TP) and True Negatives (TN). Based on NN with L2 regularization, a 2-fold cross-validation gives good results for L1 through L4. Fig. 10 illustrates 2-fold cross validation using different classifiers for system assessment. Thus, NN with L2 Regularization can produce the smallest FP and FN, resulting in a more accurate system. Using 2-fold cross-validation for the lumbar spines L1-L4, Table VII describes the systems average performance using the various classification methods. The system is tested for 2-fold cross validation for descriptor 696 with lambda 0.9. Sensitivity (True positive Rate-TPR) is the percentage of actual positive results that were correctly predicted. The specificity indicates how many negatives were correctly predicted out of actual negatives. Probability of successful positive predictions is measured by precision [35,36]. The recall measures the proportion of true positives for which the prediction was correct. Precision and recall are expressed as the harmonic mean, this method accounts for both false positives as well as false negatives which is termed as F1-score [37]. NPV (Negative Predictive Value) is the probability that someone will not have the specific disease, even after testing negative for it. According to all the above mentioned measures the NN with L2 regularization techniques attain the best classifications and overall the system measures the trueness of the test samples accurately.

Fig. 11 shows how system performance can be represented from a graphical point of view using the Receiver Operating Characteristic (ROC) curve. The red line represents the area under (AU) ROC. Essentially all of the images of test samples are classified into abnormal and normal images based on the curves occupying 99 percent of the region. This diagonal line

shows an average AUROC value of 0.5 is found in ROC curves of random predictors. In order to determine if the model is useful, the random predictor is often used. ROC curves plotted in terms of sensitivity (TPR) versus 1-specificity (false positive rate). This model gives better classification against false positive and true positive rates since the AUC approaches 1, that is it is more accurate. Based on the comparison of different methods in Table VIII, the proposed approach is found to be better than those adopted by the other methods in terms of accuracy, sensitivity, specificity, and F1-score. Because, the overall design will enable more accurate classification methods of osteoporosis detection with the help of preprocessing and significant image texture feature extraction using PHOG.

TABLE VI. EVALUATION OF DIFFERENT CLASSIFIERS

Classifiers	TP	FP	TN	FN
Decision Tree (DT)	113	3	34	2
Naïve Bayes (NB)	108	5	32	7
Support Vector Machine (SVM)	102	9	30	11
K-Nearest Neighbor (KNN)	110	6	32	4
Neural Network with L2 Regularization	115	1	36	0

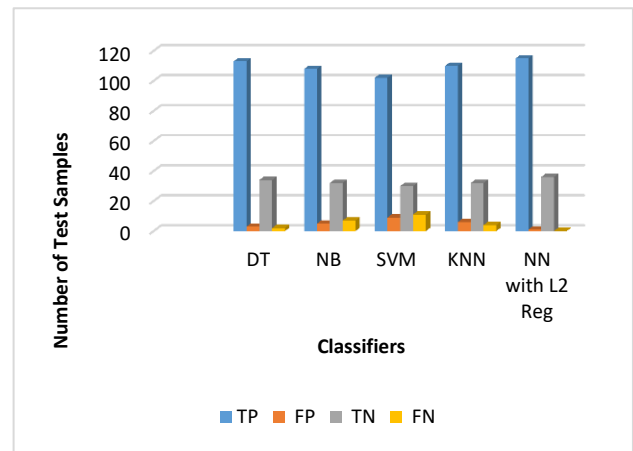


Fig. 10. An Evaluation System represented Graphically.

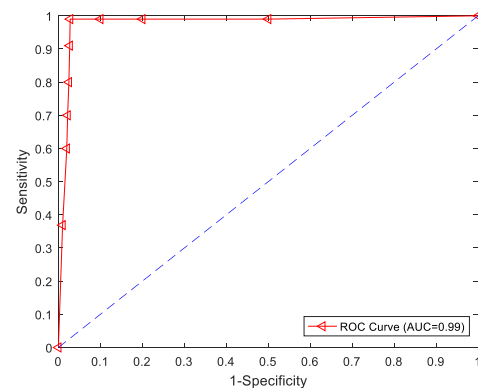


Fig. 11. Analyzing Classifications with AUROC.

TABLE VII. PERFORMANCE OF DIFFERENT CLASSIFIERS

Classifiers	Sensitivity (%)	Specificity (%)	F1-Score (%)	PPV (%)	NPV (%)	Accuracy (%)
Decision Tree (DT)	98.26	91.89	97.83	97.81	94.44	96.71
Naïve Bayes (NB)	93.81	86.48	94.73	95.57	82.05	92.10
Support Vector Machine (SVM)	90.26	76.92	91.06	91.89	73.17	86.84
K-Nearest Neighbor (KNN)	96.49	84.21	95.61	94.82	88.88	93.42
<b>Neural Network with L2 Regularization</b>	<b>100</b>	<b>97.29</b>	<b>99.56</b>	<b>99.13</b>	<b>100</b>	<b>99.34</b>

TABLE VIII. PERFORMANCE COMPARISON WITH DIFFERENT METHODS

Techniques and features	Image modalities	Classifiers	Cross -validation	Specificity (%)	Sensitivity (%)	Accuracy (%)
Fractal Analysis [4]	CT	k-NN	3-fold	90	78	81
Morphometric [7]	CT	k-NN	10-fold	83.3	92.5	88.3
3D feature with vBMD [8]	CT	Random forest	10-fold	78	77	
CNN[9]	CT	BCC- net MS-net	7- fold 2- fold	-	-	76.65
OSTA[10]	Person attribute values	ANN SVM RF KNN LoR	OSTA[10]	64.6 54.7 70 72.9 53.3	91.7 91.7 87.5 83.3 95.8	-
<b>Proposed</b>	<b>X-ray and DXA Report</b>	<b>NN with L2 regularization</b>	<b>2- fold</b>	<b>97.29</b>	<b>100</b>	<b>99.34</b>

### VIII.CONCLUSION

Millions of people worldwide suffer from osteoporosis, especially those who are aging. The purpose of this study is to reduce risks of bone fracture by exploiting the effectiveness of osteoporosis with the help of X-ray images. The aim of this study is to reduce the bone fracture risks by exploring the effectiveness of osteoporosis. In this research work, preprocessing step allows experimentally extracting the texture features of trabecular bone structure, allowing visualization of fine texture features of internal trabecular bone. An analysis of texture features is helpful for identifying the significant features that allow the images to be classified. By using these texture features, NN with L2 regularization achieves better system performance in regards to classifying the normal and osteoporotic L1-L4 images. The contribution of this paper is to understand the diagnose of osteoporosis in lumbar vertebrae L1-L4 of spine X-ray image more accurately so that risk factors can be avoided with less cost.

### IX. FUTURE WORK

To improve the diagnostic accuracy with the help of spine X-ray image of lumbar vertebrae, it is necessary to make a standard dataset that can be used universally by researchers so that one can set the standard analysis of trabecular bone micro architecture to reduce error classification. A reliable and cost-effective system is needed to overcome the drawbacks of DXA, such as lack of measurement of soft bone (trabecular) and set the standard benchmark to analyze BMD. For the bone density to be calculated, a machine learning algorithms should combine with larger number of clinical reports and X-ray imaging data in order to avoid bone risk factor.

### ACKNOWLEDGMENT

The authors would like to thank Dr. A Ramalingaiah M.S (ortho) Orthopedic Doctor, Bangalore. No.271 Abhilasha Orthopedic Hospital 5th Block 100ft Rd Banashankari 3rd stage Bangalore - 560085, India for providing X-ray image of spine (raw-data) and DXA report. We sincerely acknowledge and are grateful for their guidance and suggestions.

### REFERENCES

- [1] Riggs, B. Lawrence. "Overview of osteoporosis." Western journal of medicine 154, no. 1,63,1991.
- [2] On Osteoporosis, NIH Consensus Development Panel, and Diagnosis Prevention, "Osteoporosis prevention, diagnosis, and therapy," JAMA 285, no. 6, pp. 785-795, Feb. 2001.
- [3] H. P. Dimai, R. Ljuhar, D. Ljuhar, B. Norman, S. Nehrer, A. Kurth, and A. Fahrleitner-Pammer, "Assessing the effects of long-term osteoporosis treatment by using conventional spine radiographs: results from a pilot study in a sub-cohort of a large randomized controlled trial," Skeletal radiology 48, Vol. 7, pp.1023-1032, Dec.2019.
- [4] HY. Park, JY. Ha, KW. Kim, I. H. Baek, S. B. Park, and J. S. Lee, "Effect of lumbar spinal stenosis on bone mineral density in osteoporosis patients treated with ibandronate," BMC Musculoskeletal Disord 22, Vol.1, pp.1-7, May 2021.
- [5] Z. Omiotek, R. Dzierżak, and S. Uhlig, "Fractal analysis of the computed tomography images of vertebrae on the thoraco-lumbar region in diagnosing osteoporotic bone damage," In: Proc.of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine, Vol. 233, No. 12, pp. 1269-1281, Oct. 2019.
- [6] K. H. Nam, I. Seo, D. H. Kim, J. I. Lee, B. K. Choi, and I. H. Han, "Machine Learning Model to Predict Osteoporotic Spine with Hounsfield Units on Lumbar Computed Tomography," Journal of Korean Neurosurgical Society, Vol. 62, No. 4, pp. 442-449, Jul. 2019.
- [7] Patil, Kavita Avinash, and KV Mahendra Prashanth, "Texture Feature Extraction of CTB Radiograph Image Using Derivative Gaussian filter With NN Classification to Diagnose Osteoporosis," International Journal of Innovative Technology and Exploring Engineering, Vol.8, June 2019.

- [8] A. S. Shaker, "Detection and Segmentation of Osteoporosis in Human Body using Recurrent Neural Network," International Journal of Advanced Science and Technology, Vol. 29, No. 2, pp. 1055-1066, Jan. 2020.
- [9] H. Yousefi, E. Salehi, O. S. Sheyjani, and H. Ghanaati, "Lumbar Spine Vertebral Compression Fracture Case Diagnosis Using Machine Learning Methods on CT images", In: Proc. of 4th International Conference on Pattern Recognition and Image Analysis (IPRIA) IEEE, Tehran, Iran, pp. 179-184, Mar. 2019.
- [10] Patil, Kavita Avinash, KV Mahendra Prashanth, and A. Ramalingaiah. "Osteoporosis Detection in Lumbar spine L1-L4 based on Trabecular Bone Texture Features." International Journal of Intelligent Engineering and Systems 14, no. 6 (2021): 80-94.
- [11] A. Valentinitsch, S. Trebeschi, J. Kaesmacher, C. Lorenz, M. T. Löffler, C. Zimmer, T. Baum, and J. S. Kirschke, "Opportunistic osteoporosis screening in multi-detector CT images via local classification of textures," Osteoporosis international, Vol. 30, No. 6, pp. 1275-1285, Jun. 2019.
- [12] C. Tang, W. Zhang, H. Li, L. Li, Z. Li, A. Cai, L. Wang, D. Shi, and B. Yan, "CNN-based automatic detection of bone conditions via diagnostic CT images for osteoporosis screening," arXiv preprint arXiv:1910.06777, Oct. 2019.
- [13] WY Ou Yang, CC Lai, MT Tsou, and LC Hwang, "Development of Machine Learning Models for Prediction of Osteoporosis from Clinical Health Examination Data," International journal of environmental research and public health 18, Vol no.14, pp. 7635, Jan. 2021.
- [14] RC Gonzalez, "Richard E. woods", Digital image processing, Vol. 2, pp. 550-570, 2002.
- [15] DJ Bora, "Importance of image enhancement techniques in color image segmentation: A comprehensive and comparative study," arXiv preprint arXiv:1708.05081, Aug. 2017.
- [16] Mishra, Akshansh. "Contrast Limited Adaptive Histogram Equalization (CLAHE) Approach for Enhancement of the Microstructures of Friction Stir Welded Joints." arXiv preprint arXiv:2109.00886 ,2021.
- [17] L Tang, Y Li, T Ling, L He, T Luo, and D. Guo, "UAV image registration algorithm based on overlapping region detection," In 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vol.10, pp. 6706-6709, Jul. 2016.
- [18] A Plyer, E Colin-Koeniguer, and Weissgerber F, "A new coregistration algorithm for recent applications on urban SAR images," IEEE Geoscience and Remote Sensing Letters 12, Vol.11, pp. 2198-202, Aug. 2015.
- [19] G Brigot, E Colin-Koeniguer, A Plyer, and Janez F, "Adaptation and evaluation of an optical flow method applied to coregistration of forest remote sensing images," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 9, Vol. 7, pp. 2923-39, Jul. 2016.
- [20] F Champagnat, A Plyer, G Le Besnerais, B Leclair, and Y Le Sant, "How to calculate dense piv vector fields at video rate," In Proceedings of 8th International Symposium on Particle Image Velocimetry-PIV09, Vol. 11, pp. 15-20, Aug. 2009.
- [21] A Plyer, G Le Besnerais, F Champagnat, "Massively parallel Lucas Kanade optical flow for real-time video processing applications," Journal of Real-Time Image Processing 11, Vol. 4, pp. 713-30. Apr. 2016.
- [22] Frost BA, Camarero-Espinosa S, Foster EJ, "Materials for the Spine: Anatomy, Problems, and Solutions," Materials (Basel).12(2):253,PMID: 30646556,Jan 14 2019.
- [23] Patil, Kavita Avinash, and KV Mahendra Prashant. "Segmentation of Lumbar [L1-L4] AP Spine X-ray images using various Level Set methods to detect Osteoporosis." In 2021 IEEE Bombay Section Signature Conference (IBSSC), pp. 1-6. IEEE, 2021.
- [24] Bailey, Donald G., and Robert M. Hodgson, "Range filters: Local intensity subrange filters and their properties," Image and Vision Computing 3, no. 3 ,pp: 99-110, 1985.
- [25] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol. 1, pp. 886-893, CVPR.2005.177, June 2005.
- [26] OL Junior, D Delgado, V Gonçalves, and U Nunes, "Trainable classifier-fusion schemes: An application to pedestrian detection," In 2009 12th international IEEE conference on intelligent transportation systems 4, pp. 1-6, Oct. 2009.
- [27] A Bosch, A Zisserman, and X Munoz, "Representing shape with a spatial pyramid kernel," In Proceedings of the 6th ACM international conference on Image and video retrieval, pp. 401-408, Jul. 2007.
- [28] EB Sudderth, A Torralba, WT Freeman, and AS Willsky, "Learning hierarchical models of scenes, objects, and parts," In Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, Vol. 2, pp. 1331-1338, Oct. 2005.
- [29] DG Lowe, "Distinctive image features from scale-invariant key points," International journal of computer vision 60, Vol. 2, pp. 91-110, Nov. 2004.
- [30] DR Baughman, and YA Liu, "Classification: fault diagnosis and feature categorization," In Neural Networks in Bioprocessing and Chemical Engineering, pp. 110-171, Academic Press 1995.
- [31] AY Ng, "Feature selection, L 1 vs. L 2 regularization, and rotational invariance," In Proceedings of the twenty-first international conference on Machine learning, p. 78, Jul. 2004.
- [32] Demir-Kavuk, O., Kamada, M., Akutsu, T., & Knapp, E. W, " Prediction using step-wise L1, L2 regularization and feature selection for small data sets with large number of features ," BMC bioinformatics, 12, 412, 25 Oct 2011.
- [33] N Jain, "Optimization of Regularization and Early Stopping to Reduce Overfitting in Recognition of Handwritten Characters," In Congress on Intelligent Systems, pp. 305-323, Sep. 2020.
- [34] Iftikhar, Asim, Muhammad Alam, Rizwan Ahmed, Shahrulniza Musa, and Mazliham Mohd Su'ud, "Risk Prediction by Using Artificial Neural Network in Global Software Development," Computational Intelligence and Neuroscience 2021 (2021).
- [35] Patil, Kavita Avinash, KV Mahendra Prashanth, and A. Ramalingaiah. "Texture Feature Extraction of Lumbar Spine Trabecular Bone Radiograph Image using Laplacian of Gaussian Filter with KNN Classification to Diagnose Osteoporosis." In Journal of Physics: Conference Series, vol. 2070, no. 1, p. 012137. IOP Publishing, 2021.
- [36] Powers, David MW, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," arXiv preprint arXiv:2010.16061 ,2020.
- [37] Townsend, James T., "Theoretical analysis of an alphabetic confusion matrix," Perception & Psychophysics 9, no. 1 , pp. 40-50, 1971.