Intelligent Assistive Methods for Diagnosis of Rheumatoid Arthritis Using Histogram Smoothing and Feature Extraction of Bone Images

SP. Chokkalingam, K. Komathy

Abstract—Advances in the field of image processing envision a new era of evaluation techniques and application of procedures in various different fields. One such field being considered is the biomedical field for prognosis as well as diagnosis of diseases. This plethora of methods though provides a wide range of options to select from, it also proves confusion in selecting the apt process and also in finding which one is more suitable. Our objective is to use a series of techniques on bone scans, so as to detect the occurrence of rheumatoid arthritis (RA) as accurately as possible. Amongst other techniques existing in the field our proposed system tends to be more effective as it depends on new methodologies that have been proved to be better and more consistent than others. Computer aided diagnosis will provide more accurate and infallible rate of consistency that will help to improve the efficiency of the system. The image first undergoes histogram smoothing and specification, morphing operation, boundary detection by edge following algorithm and finally image subtraction to determine the presence of rheumatoid arthritis in a more efficient and effective way. Using pre-processing noises are removed from images and using segmentation, region of interest is found and Histogram smoothing is applied for a specific portion of the images. Gray level co-occurrence matrix (GLCM) features like Mean, Median, Energy, Correlation, Bone Mineral Density (BMD) and etc. After finding all the features it stores in the database. This dataset is trained with inflamed and non-inflamed values and with the help of neural network all the new images are checked properly for their status and Rough set is implemented for further reduction.

Keywords—Computer Aided Diagnosis, Edge Detection, Histogram Smoothing, Rheumatoid Arthritis.

I. INTRODUCTION

RHEUMATOID arthritis is one among the most painful diseases afflicted to the bones and generally comes with the onset of age. It causes deformation in joints as well as leads to inflammation also causing incapacitation in motion. Once it goes beyond a certain stage it cannot be reverted back and causes impairment in motion. Its diagnosis in an early stage plays an important part in the treatment of the disease and so accuracy in the prognosis is very important [1].

SP. Chokkalingam is Associate Professor with Saveetha School of Engineering, Saveetha University, Anna University Research Scholar, Chennai, (e-mail: cho_mas@yahoo.com).

K. Komathy is Professor & Head with Department of Information Technology, Hindustan University, Chennai, (e-mail: gomes1960@gmail.com).

The symptoms of the RA or joint pain, a flare, swelling of the joint etc. and functional disorder happens when a symptom is in progress.

If the bone is transformed once, it cannot be restored. Therefore, early detection and early treatment are the most important problems [2]. The main diagnosis of the RA is generally done by performing a blood test, diagnostic imaging, and a self-report of the patients. The RA cannot be diagnosed only with a numerical value of blood test, [3], [4] because the numerical value of the blood test is not affected to change for RA alone. Therefore, diagnostic imaging that diagnoses the RA is requisite. The treatment of the RA prevents the immobility of a joint and the bone by inflammation. More correct diagnosis, imaging is important to prevent an anti-rheumatic drug from administering it to a patient without the need.

The first step in scan analysis can be histogram analysis and normalization because it will help in proper contrasting of the various eccentricities present in the image and highlight the different areas termed as region of interest [5]. This will also help to differentiate the various objects present in the image. A series of morphing operation is to be applied so that the region of interest can be specified and focused on. This includes the process of dilation and erosion so that the focus of processing on the bones in the image can be made error free and more directed towards the joint areas. Finding the boundary of edges in the image is still a difficult task. Though multiple techniques like Sobel filter, canny edge detector and various snake algorithms exist; they still are not fully accurate and have their own drawbacks. Boundary detection is achieved with various techniques [6] like active contour model (ACM), geodesic active contour (GAC) model, active contours without edges (ACWE), gradient vector flow (GVF) snake model, vector field convolution (VFC) snake model, etc. For boundary detection snake model is taking vital role.

A bone disease always depends upon the tissue properties in the bone. Though HR-pQCT scans are available it is difficult to separate cortical and trabecular bone. We use an automatic thresholding method for segmentation in HR-pQCT scans. Threshold independent segmentation tool (TIST) [7] is used to classify bones and to find the thickness and tissue volume of the bones. It shows positive results with high accuracy in identifying the trabecular and cortical bones. Threshold values with slight change in it will affect segmentation. So we apply fully automatic threshold segmentation mechanism. TIST tool

used to capture 3-D images of gray scale images. For calculation 3 methods were used: 3D- GLCM (gray level Co occurrence matrix) - To extract statistical image information. Local structure tensor & Hessian Matrix were used to determine properties like tubularity and fatness.

Random Forest classifier is used to classify the features from the training sets .Segmentation was done by TIST and is compared with Standard Algorithm and CRTX which are threshold based segmentation methods [8]. It included 14 samples with an age for male as 72.9 yrs and female as 80.0 yrs and in the ratio of 1:1.CRTX and TIST performed more accurate than SA where CRTX accurate for female samples and TIST more precise for male samples. The important difference is that there is no need for operator interaction like SA or CRTX. TIST can be developed with few more different samples to increase its performance and its accuracy. This a 3D-texture can be easily analyzed based on automatic threshold based independent cortex segmentation which uses HR-pQCT images [9], [10].

This method is used to detect the size and the amount of minerals automatically in core samples. Normalized color based statistical segmentation method is used to determine the Avg value, S.D of RGB color components in the sample images. Segmentation is 90% accurate by this method. Some of the methods that are used to identify ores are X-ray diffraction, optical light microscopy, Scanning Electron microscope (SEM). SEM uses Charge-Couple Device (CCD) cameras with an electron beam as an input and gets backscattered electron detector (BSE) images as output which is then digitized. It can also do using an X-ray detector (EDX imaging). Ore microscopy is better than semi microscopy. It is fast and less expensive compared to semi microscopy. Segmentation is done using gray scale images. Ore microscopy is used to determine statistical properties and RGB color components in the images. A normalized colour method is used to size and amount of minerals. Image minerals are found by the Naïve Bayesian classifier.

This method uses a 5mp Nikon digital camera .It uses 3.2GHz CPU and 1024 MB RAM to perform the segmentation and measurements. Some of the segmentation methods are color based segmentation, normalized color based segmentation, Bayesian approach, Histogram based approach, Statistical segmentation technique, Region growing technique. The main aim is to identify simple and best method for segmentation.

A Histogram based approach [11] is used to compare the results of the segmentation process. Color based technique is used to find out the amount of RGB colors present in the minerals. It leads to find out the avg and standard deviation of segmentation. Normalised color based removes the intensity but preserves the color values in the image [11], [12]. Different Intervals for each colour is determined and compared with average colour value of Characteristic Texture for each mineral. Naïve bayes theorem is used to identify colour and intensity of each mineral. Histogram based approach is used to compare each methods with its intensity or color of each minerals. The region based approach is used to segment the images in to regions and from the regions it groups RGB colours of same type to identify the intensity. The percentage of each mineral in the sample is identified and compared to identify the best method. The colour based approach method segments all the minerals with good percentage. Normalized colour based with enhancements produce the results which are average. NCM method performs segmentation smoother than Bayesian and Histogram methods. To identify mineral textures Naïve Bayesian method was used. Around 86% of minerals were segmented and calculated accurately based on 60 micrographic images. It can be improved by high definition coloured images.

This paper is used to identify various tissue properties in a bone. It is used to determine the fracture risk of a bone based on its properties. We use quantitative computed tomography (QCT) to determine the volumetric BMD. It is used to find out whether BMD is positive or negative with fatigue life .It is found that the standard deviation of BMD is negatively associated with fatigue life. Fatigue life is defined as the no of applications of a given stress to which a sample of metal can be subjected before failing [13]. The QCT scans were obtained with the specimens and simulated and are in .8mm isotropic voxel size with 1mm thickness of images. All the specimens were fatigue loaded using load levels. Fatigue loading is stopped when 1000 cycles is reached or when vertebrae is fractured. Hazardal situations are measured during the calculation of BMD [14]. The analysis identifies the average and S.D of BMD for 18 values along with fatigue load. It also finds out the min BMD and max BMD to identify the nature of the information. It is then calculated by Fishers r to z transformation. The S.D of BMD identifies the fatigue résistance and AVG of BMD strength of vertebrae. This paper identifies the values with low cycle and high magnitude of fatigue, but values for high cycle and low amplitude is yet to be determined [15]. Finally, using trained set of images and comparison through subtraction of progressive images we can detect changes in the phalange joints. This will help in early and fast detection of any variations in the joints. The method is termed safe to improve on the no of true positives as it is independent of error dependent factors.

II. RELATED WORKS AND ISSUES

At the present a large amount of research has been started in image processing for medical field as per the specifications of diagnosis, but the number of true positives generated is very low and limited. This leads to inaccuracy and the methods are not reliable for continuing. Various segmentations, and clustering techniques based on discrete cosine transform (DCT), NCuts Algorithm, SOFES Algorithm and even manual segmentation have been employed to perform and help in computer aided diagnosis from scans with accuracy [16]. The various disadvantages result from the methods inefficiency for the procedure as they cannot sufficiently deal with the respective image scans properly or give a proper diagnosis. Also a certain amount of negative results ensure that the methods be refined more to give precise results. Methods

employing canny edge detector are less accurate and the sobel filter smoothens the image allowing edge analysis more accurately by producing considerably higher values for similar edges and hence giving a better detection [17].

The paper highlights the use of better methods for image analysis for the image scan provided so that the procedure followed is less error prone and can be used more efficiently [18]. To improve the image quality histogram equalization has been practiced for a long time, but it has been experienced that it is not effective for complex noisy data which can hence lead to inappropriate handling of the image. So we follow another process called histogram smoothing and specification is followed which allows normalization of image quality and brightness and hence enhances the quality of the image. This hence provides immunity to noise in the image making the procedure work sensitive rather than noise sensitive [18], [19]. The choice of various methods used here, provide a complete knowhow about the image dependent on the method of processing and the steps to be followed for the same. Though simple the steps provide an effective diagnosis compared to all previous methods [20].

III. SYSTEM DESIGN

The design of the system is based on various technologies from image processing likewise for providing complete diagnosis automatically. Our proposed method consists of a stepwise processing where first the image is normalized through histogram smoothing and specification, and then it undergoes morphological operation of dilation and erosion to enable extraction of the bones area termed as the region of interest (ROI). Next using boundary detection we find out the edges of the bones sufficiently underlining the joint area for analysis which is the subtracted from previous image and the results are generated as a diagnosis where if the value is above a specific standard it can be positive for RA.

A. Histogram Smoothing and Specification

Histogram of an image shows a strong peak because the image is dominated by a large area of a single gray-level value, this can cause problems when using Histogram Equalization. One method, Piecewise Therefore preference to this method in which segments yield the original means of the modes, maximizing their entropy as suggested in the method of Brightness Preserving Histogram Equalization with Maximum Entropy, is natural. Further contrast improvements can be achieved if the valleys between the modes that have areas much smaller than the rest can be eliminated, choosing a single valley in between. Additionally, the number of valleys between two consecutive peaks that are close together should be reduced by eliminating the valley in the middle.

B. Morphological Operation

Morphological operation is a transformation technique in the object image that consists of the set of theory operation to extract the specific feature from given binary image or gray image. There are two types of morphological operations; erosion and dilation. The opening processing is processing that performs dilation after erosion and the closing processing is processing that performs erosion after dilation. When the opening processing is performed on the image, the part, where pixel value is large, becoming not dim and disappears. Finally we obtain the differences of original image and the opening processing image is performed, the outline part of which becomes dim will be extracted.

The general expressions for the two operations are:

Dilation: $[f \oplus g](x) = \max_{\substack{x+ueF \\ ueG}} \{f(x+u) + g(u)\}$

Erosion: $[f \oplus g](x) = \min_{ueG} \{f(x+u) - g(u)\}$

These two processes hence help in extraction of our ROI from the image scan by gradual processing.

In Fig. 1 Under preprocessing noises are removed from images and using segmentation, region of interest are found and Histogram smoothing is applied for a specific portion of the images. Gray level co-occurrence matrix (GLCM) is applied and found many features like Mean, Median, Energy, Correlation, Bone Mineral Density (BMD) and etc. After finding all the features it stores in the database. This dataset is trained with inflamed and non-inflamed values and with the help of neural network all the new images are checked properly for their status and Rough set is implemented for further reduction.
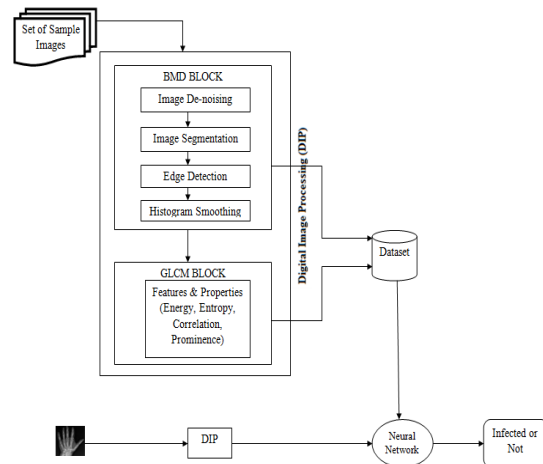

Fig. 1 Block Diagram

C. Boundary Detection

A wide range of edge detection techniques has been available in the image processing. Sobel, Prewitt, Costella, Laplacian and Canny operators which are all based on the difference in intensity of gray levels but still different in the mathematical aspects. Therefore, we extend the capability of the previous edge vector field by applying a local averaging operation where the value of each vector is replaced by the average of all the values in the local neighborhood, i.e.

$$M(i,j) = \frac{1}{Mr} \sum_{i,j \in n} \sqrt{Mx(i,j)^2 + My(i,j)^2}$$

$$D(i,j) = \frac{1}{Mr} \sum_{i,j \in n} \tan^{-1} \frac{My(i,j)}{Mx(i,j)}$$

The texture feature images of Law's texture are computed by convolving an input image with each of the masks. Given a column vector $L=(1, 4, 6, 4, 1)^T$, the 2-D mask $l_{(i, j)}$ used for texture discrimination in this research is generated by $L \times L^T$. The output image is obtained by convolving the input image with the texture mask.

The Sobel operator performs a 2-D spatial gradient measurement on an image and so emphasizes regions of high spatial frequency that correspond to edges. Typically it is used to find the approximate absolute gradient magnitude at each point in an input grayscale image. The convolution kernels are given as:

$$G_x = \begin{matrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{matrix} \qquad G_y = \begin{matrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{matrix}$$

The gradient magnitude is given by

$$|G| = \sqrt{Gx^2 + Gy^2}$$

Result of Sobel operator is in the form of 2-Dimensional map of the gradient at each point. This method is effective for proper evaluation of image.

D. Image Subtraction

This step requires the subtraction of images, one of a previous time and one processed image to find out the changes affected into the bones at the ROI and hence helps in diagnosing the presence of RA in the patient more effectively.

E. Working Principle

Our proposed method consists of a stepwise processing where first the image is normalized through histogram smoothing and specification, and then it undergoes morphological operation of dilation and erosion to enable extraction of the bones area termed as the region of interest (ROI). Next using boundary detection we find out the edges of the bones sufficiently underlining the joint area for analysis which is the subtracted from previous image and the results are generated as a diagnosis where if the value is above a specific standard it can be a positive for RA. The complete flow of proposed system can be seen in Fig. 1.

In our proposed system, Neural Network is used for classification. Whenever a new bone image is given as input to the proposed system, it is pushed into the BMD block. Then various steps are performed on this input bone images. Finally BMD and GLCM features are extracted from input bone image. These values are compared against the dataset using neural network and then it is classified as whether infected or not infected.
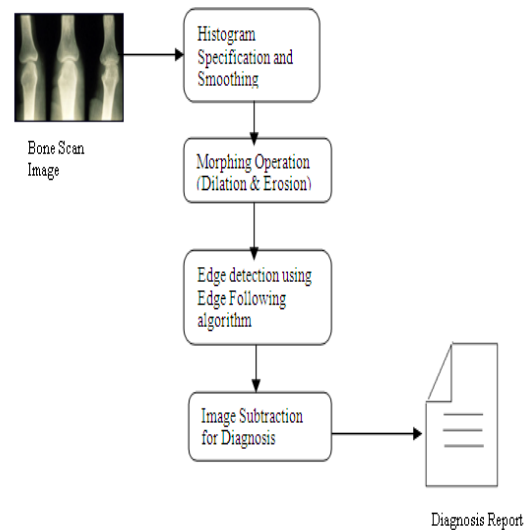

Fig. 2 System Design and Flow

Fig. 3 Original Image
Bone images are collected from various laboratories and online image database

Fig. 4 Image after Denoising using Median Filter

Noise may be accumulated in the image during image acquisition. So, the given image is subjected to Median filtering techniques in order to enhance the image because median filter shows average performance on all types of noise. New procedure is created for measuring these performance parameters. From the result of these performance parameters we can identify that median filter gives average performance on all type of noise.
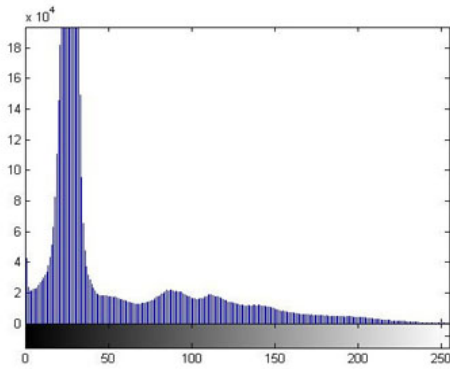
Fig. 5 Histogram of Noise free Images

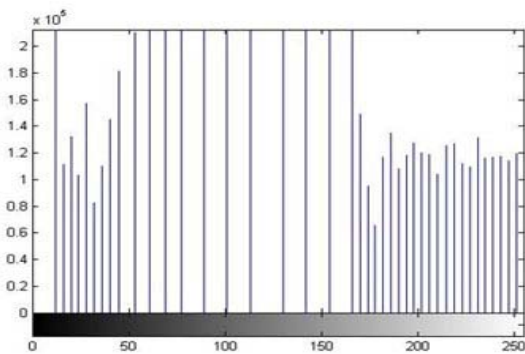Fig. 6 Smoothing using Probability and Cumulative Function

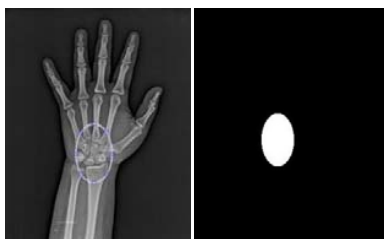Fig. 7 Histogram of Smoothed Image

Fig. 8 Segmentation using threshold approach

Probability and Cumulative distribution function method based smoothing is used to identify those different types of objects appearing in the image and makes segmentation process easier.

```
Algorithm 1: The probability of each occurrence
Input: Noise free Image
Output: Smoothed Image
Procedure:
numofpixels=size(GIm,1)*size(GIm,2);

%freq counts the occurrence of each pixel value.
%The probability of each occurrence is calculated by
probf.

for i=1:size(GIm,1)
   for j=1:size(GIm,2)
      value=GIm(i,j);
freq(value+1)=freq(value+1)+1;
probf(value+1)=freq(value+1)/numofpixels;
   end
end

sum=0;
no_bins=255;
```

```
Algorithm 2: The cumulative distribution  probability

Input: Noise free Image
Output:  Smoothed Image

Procedure:

for i=1:size(probf)
  sum=sum+freq(i);
  cum(i)=sum;
probc(i)=cum(i)/numofpixels;
  output(i)=round(probc(i)*no_bins);
end

for i=1:size(GIm,1)
   for j=1:size(GIm,2)
HIm(i,j)=output(GIm(i,j)+1);
   end
end
```

This process involves splitting of given bone image into multiple segments in order to separate the Region of Interest. Threshold based segmentation makes the Region of Interest more accurate. In segmentation procedure, new method is developed in order to select the region of interest by drawing ellipse. Automatic selection of region of interest has a disadvantage that phalanges position in digital x-ray image are differ from one image to another image and joints of phalanges should be take for analysis. In order to produce good accuracy and better performance of proposed system region of interest should be taken manually. In our proposed system other steps are created in automatic manner. Initially region of interest is selected by drawing ellipse and then selected portion are segmented. Then the selected portion is mapped with white color and again selected portions are mapped with black color and stored as new image. Finally these white and black images are subtracted in order to identify region of interest. From this, inner portion of region of interest are identified clearly and accurately.

Algorithm 3: To Segmentation hand portion from digital x-ray image

Input: Digital X-ray image contains hand portion and some other extra waste space
Output: Image containing hand portion

Procedure: Threshold (Image, Slider_value)
Image is the input and segmented from the whole image. Slider values are taken as a threshold value and applied to the Image for Segmentation.
Begin
    if Image is in RGB
    then image = RGB-to-Grey(Image)
End

if slider_value gets changed
    then size(row , col) = size(Image)
foreachrow of Image do
foreachcol of Image do
    if pixel_val(row , col) >slider_value
        then pixel_val(row , col) = 255
        else pixel_val(row , col) = 0
End
End
End
End
    return Image
End
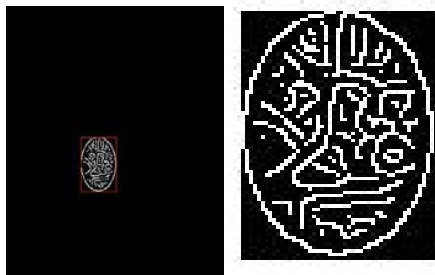

Fig. 9 Edge Detection using Canny Algorithm

Fig. 10 Applying Bounding Box and Cropped Image

Edge detection using canny algorithm identifies the edges of region of interest accurately.

Images are cropped for extracting different features. This process will lead us to more accurate results. The RGB images are converted into gray scale image. Sizes of images are in the matrix forms and it can be assigned into row and column. It is required to process up to the n number of rows and columns and each time it will check whether row and column values are greater than slider values and if yes, it can be converted into a white pixel otherwise it can be taken as black. In Image cropped Lymphocyte image centroid of X- Coordinates and Y- Coordinates are assigned to meanx and meany variables. Initially calculate the minimum and maximum distance from centroid to edges and divide it ten intervals. Images are scanned from left to right.

$$BMD = \frac{BMC}{Volume}$$

$$BMC = IntegratedDensity \times \frac{1.073\,gram}{perimeter}$$

$$Volume = \prod (diameter\,/\,2)^2\,Fixedlength$$

BMD –Bone Mineral Density
BMC – Bone mineral Content
Volume – Volume of Region of Interest

Algorithm 4: Features Extraction

Input: Cropped and Reshaped image after Edge detection using canny algorithm
Output: Features Mean, Median, Variance, Standard Deviation

Procedure: Image is the input and features are taken from whole image.
[r,c]=size(n);

%To calculate the mean of the matrix
totmean=sum(n(:))/(r*c);

%To calculate the variance and standard deviation
totdiff=(n-totmean).^2;
totsum=sum(totdiff(:));
nele=(r*c)-1;
totvar=totsum/nele;
totstd=sqrt(totvar);

%To find the total median of the image
C=reshape(n,[],1);
min1=min(C);
max1=max(C);
E=sort(n(:));
num=round((r*c)/2);
if((mod(r,2)==0) || (mod(c,2)==0))
totmedian=(E(num)+E(num+1))/2;
else
totmedian=E(num);
end

Edge detected image taken as the input and region of interest should be separated from the edge detected image. Bounding box is one of the image processing novel algorithms which is used to crop the required portion of connected components using maximum and minimum coordinates of edge detected images. Then colors are spread across cropped

image using spread spectrum in order calculate density of region of interest. Bone Mineral Density refers to the amount of mineral present in an area of bone. This work focuses on finding density of bone and do not deal with minerals of bone. We have proposed a new mathematical relationship and procedure to calculate bone mineral density from the region of interest.

It is important to understand what is meant by "density" in a microCT image. The reconstructed grey-scale intensity of each image voxel does not relate directly to mass density alone, thus do not expect it to correspond with measurements of weight per volume, e.g. g.cm-3. The primary entity that is measured is x-ray absorption, defined as the attenuation coefficient in units of 1/distance, specifically, mm-1. This is determined both by mass density and elemental composition of the material. However, when we know that the x-ray absorption of a material is dominated by one specific material, then we can relate the measured x-ray attenuation coefficient (AC) to the mass density of that material.

Bone mineral density (BMD) is defined as the volumetric density of calcium hydroxyapatite (CaHA) in a biological tissue in terms of g.cm-3.Itis calibrated by means of phantoms with known density of CaHA. Bone mineral density can refer to two different measurements: HU and BMD calibration in Bruker-MicroCT CT-analyser

(a) the combined density of a well-defined volume which contains a mixture of both bone and soft tissue, such as a selected volume of medullary trabecular bone in a femur or tibia, is measured as "bone mineral density", or BMD. This parameter BMD relates to the amount of bone within a mixed bone-soft tissue region, but does not give information about the material density of the bone itself.

(b) the density measurement restricted to within the volume of calcified bone tissue, such as cortical bone, excluding surrounding soft tissue, is called "tissue mineral density" or TMD [1]. By contrast to BMD, this gives us information about the material density of the bone itself, and ignores surrounding soft tissue.

Bone Mineral Density is derived from Bone Mineral Content (BMC) of the region of interest and volume of region of interest. BMD is measured in grams/cm^3, BMC is measured in grams and volume of ROI is measured in cm³. Sum of volume of pixels in ROI refers to the integrated density of ROI. 1.073 grams is the default value for all the images that refers to the amount of air and water present in bone. Calibration of ROI refers to some measurements of region of interest, in this work perimeter is considered for calculation. Volume of ROI is measured using above mentioned formula. In that formula, diameter is calculated from area of region of interest and k refers to that factor used to identify pixels relationship in an image or length of ROI. In this work, k is calculated using entropy of region of interest/total mean of region of interest. New procedure is created for calculating area, perimeter and integrated density.

In this work, Gray Level Co-occurrence Matrix (GLCM) features (Energy, Entropy, Contrast, Homogeneity and Correlation) for bone images are calculated. GLCM is one of the very famous and popular image analysis techniques which are used to extract above mentioned texture features and classification. GLCM is a statistical way to identify the occurrence of gray level values in an image and to identify how each pixel is related with adjacent pixels. Correlation refers that positively or negatively a pixel is correlated with neighbor pixel. Contrast refers to intensity contrast or brightness of a pixel with its neighbor pixel over whole image. In other way, in a continuous set of pixels, difference between highest and lowest values is contrast. Uniformity of pixels pairs or similarity between pixels is called as energy. Homogeneity returns the values of GLCM to GLCM diagonal by applying closeness distribution of elements. Calculation of GLCM texture features can be performed using below listed formulae.

$$\text{Contrast} = \sum_i \sum_j (i-j)^2 C_{ij}$$

$$\text{Entropy} = \sum_i \sum_j C_{ij} \log C_{ij}$$

$$\text{Energy} = \sum_i \sum_j C_{ij}{}^2$$

$$\text{Homogeneity} = \sum_i \sum_j \frac{C_{ij}}{1 + |i-j|}$$

In this work, some more features like Mean, Median, Variance and Standard deviation are also derived for both normal and eroded bone images. These values are used for classification because these values will be different for normal and eroded bone images. New procedure is created for calculating Mean, Median, Variance and Standard deviation. Calculation of these features can be performed in this manner. For calculating mean value, intensity value of each pixel are added and divided by number of pixel in region of interest. For calculating median value, each pixel values are arranged in ascending order. If the count of pixel is odd number, middle value from that ascending order consider as median value. If the count of pixel is even number, then middle two values are taken and average of that two values consider as median value. Variance and standard deviation of region of interest can be calculated from below listed formulae.

$$\text{Mean} = \frac{\text{Sum of pixel values of region of interest}}{\text{Number of pixels in region of interest}}$$

$$\text{Variance } (\sigma^2) = \frac{\sum(X_i - \overline{X})^2}{N}$$

TABLE I
FEATURES OF IMAGE

Images	Mean	Feret	Skew	Kurt	Standard Deviation	Entropy	Area	Perimeter	Bone Mineral Density	Result
Image 1	2.5444	99.043	0.468	-1.4	6.6913	0.6413	5442	998	0.4823	Inflamated
Image 2	2.6782	99.043	0.444	-1.6	6.8943	0.6567	5413	994	0.3224	Inflamated
Image 3	2.9396	99.049	0.439	-1.4	7.2366	0.6523	5545	997	0.2678	Inflamated
Image 4	2.3840	99.045	0.498	-1.8	7.5681	0.6654	5586	993	0.4113	Inflamated
Image 5	2.8723	99.047	0.448	-1.6	7.8153	0.6678	5796	993	0.4544	Inflamated
Image 6	2.9589	99.041	0.468	-1.7	6.9900	0.6112	5167	998	0.3867	Inflamated
Image 7	3.1980	99.049	0.493	-1.6	8.4563	0.6693	5487	999	0.4118	Inflamated
Image 8	2.5707	99.045	0.539	-1.8	7.6969	0.6817	5938	1002	0.8823	Non Inflamated
Image 9	2.6899	99.048	0.540	-1.9	7.6852	0.6999	5857	1001	0.7890	Non Inflamated
Image 10	2.4678	99.047	0.499	-1.5	7.7489	0.6789	5894	997	0.4899	Inflamated
Image 11	2.5894	99.063	0.559	-1.8	7.4900	0.6893	5811	1001	0.8859	Non Inflamated
Image 12	2.3840	99.045	0.498	-1.8	7.5681	0.6554	5786	994	0.4179	Inflamated
Image 13	2.9605	99.0467	0.488	-1.3	7.9599	0.6678	5699	999	0.3890	Inflamated
Image 14	2.9398	99.047	0.489	-1.3	6.9833	0.6489	4769	967	0.2588	Inflamated
Image 15	2.3568	99.048	0.454	-1.4	7.9849	0.6749	6299	996	0.4588	Inflamated
Image 16	2.4678	99.047	0.499	-1.5	7.7489	0.6789	5894	995	0.4545	Inflamated
Image 17	2.9599	99.042	0.423	-1.8	6.8599	0.6649	5787	995	0.5859	Inflamated
Image 18	2.5003	99.040	0.534	-1.5	7.4889	0.6839	5978	1002	0.5859	Non Inflamated
Image 19	2.6782	99.043	0.444	-1.6	6.8943	0.6567	5713	1002	0.3287	Inflamated
Image 20	2.4573	99.067	0.555	-1.6	7.6536	0.6856	5965	1000	0.6899	Non Inflamated
Image 21	2.4589	99.094	0.504	-1.2	7.6950	0.6489	5700	998	0.6283	Inflamated
Image 22	2.4940	99.098	0.594	-1.3	7.6403	0.6940	5700	1002	0.6283	Non Inflamated
Image 23	2.4399	99.038	0.504	-1.5	7.6940	0.6590	5700	994	0.459	Inflamated

F. Rough Set Implementation

In this article the information about **Bone Image i**s called a decision table. This table represents the input data, gathered from the domain medicine. In the table (Table I) rows of a table, labeled Image1, Image2, Image3,… till Image23 are called examples (objects, entities). Properties of examples are perceived through assigning values to the variables. The main concept of rough set theory is an indiscernibility relation, normally associated with a set of attributes. The attributes are Image, Area, Perimeter, Mean, Stand deviation, Skew, Feret, kurt and Bone Mineral Density.

1. Reduced Information Table: (Table II)

This table is the reduced table which mainly depends on the value of Area and BMD value. If the area is larger and BMD value is more than 0.550 then the results are non-inflamed else the results are inflamed. For e.g. The Image 17 has larger area value and Large BMD value but the result is Inflamated.

TABLE II
REDUCED INFORMATION TABLE

Image	Area	BMD	Result
Image 5	5796	0.4544	Inflamated
Image 8	5938	0.8823	NonInflamated
Image 17	5687	0.5859	Inflamated
Image 18	5978	0.5859	Non Inflamated
Image 11	5811	0.8859	Non Inflamated
Image 16	5494	0.4545	Inflamated
Image 20	5965	0.6899	Non Inflamated

2. Inconsistent Reduced Table: (Table III)

This table consists of additional inputs like Image21, Image22, Image23. It consists of attributes area and BMD. But value of both Image21 and Image22 has the same value for both the images and thus it makes inconsistent information because the result of Image21 is inflamated and for Image22 is Non-inflamated. Those attributes also depends on other values like Perimeter and Entropy. In this place we need to concentrate all the attributes and values based on that it needs to be segregated.

TABLE III
INCONSISTENT REDUCED TABLE

Image	Area	Entropy	Perimeter	BMD	Result
Image 5	5796	0.6678	993	0.4544	Inflamated
Image 8	5938	0.6817	1002	0.8823	Non Inflamated
Image 17	5787	0.6649	995	0.5859	Inflamated
Image 18	5978	0.6839	1002	0.5859	Non Inflamated
Image 11	5811	0.6893	1001	0.8859	Non Inflamated
Image 16	5494	0.6789	995	0.4545	Inflamated
Image 20	5965	0.6856	1000	0.6899	Non Inflamated
Image 21	5700	0.6489	998	0.6283	Inflamated
Image 22	5700	0.6940	1002	0.6283	Non Inflamated
Image 23	5700	0.6590	994	0.459	Inflamated

Same attributes may have different values and same attributes may have different values are also possible in this place. The larger area depends on the value of perimeter and the attribute BMD depends on the Entropy value. The overall value will provide the result as inflamated or non inflamated.

3. Reduced Table: (Table IV)

This table consists of additional attributes like Perimeter and Entropy. The rules to follow to gain the consistent result. In this table the value of BMD is Similar for both images image17 and Image18 and the value is 0.5859. this attribute

depends on perimeter value. The perimeter value is less than 1000 for Image21 and the result is inflamated. The perimeter value is greater than 1000 for Image22 and the result is non Inflamated. Table IV follows the rule:

(BMD, <0.550) →(result, Inflamated)

(BMD, >=0.550) and (Entropy,<0.6800) → (result, Inflamated)

(BMD, >=0.550) and (Entropy,>=0.6800) → (result, Non-Inflamated)

(BMD, >=0.550) and (Entropy,>0.6800) and (area,>= 5800) → (result, Non Inflamated)

(BMD, >=0.550) and (Entropy,>=0.6800) and (area, < 5800) and (perimeter < 1000) → (result, Non Inflamated)

The subset generated by lower approximation is characterized by the objects that will definitely form a part of an interest subset, and the upper approximation is characterized by objects that will be possibly from a part of an interest subset. Every subset defined through upper and lower approximation is called the Rough Set. Decisions are made based on the parameters like area and Circularity since compare with other attributes these attributes values are very useful to take decision. From the results found the lower approximation attributes are area and skew and the upper approximation attributes are Circ and median.

TABLE IV
REDUCED TABLE

Image	Area	BMD	Result
Image 5	5796	0.4544	Inflamated
Image 8	5938	0.8823	Non Inflamated
Image 17	5787	0.5859	Inflamated
Image 18	5978	0.5859	Non Inflamated
Image 11	5811	0.8859	Non Inflamated
Image 16	5494	0.4545	Inflamated
Image 20	5965	0.6899	Non Inflamated
Image 21	5700	0.6283	Inflamated
Image 22	5700	0.6283	Non Inflamated
Image 23	5700	0.459	Inflamated

IV. SUMMARY AND CONCLUSION

In this image we made a sincere effort to implement a new scheme to diagnose the presence of rheumatoid arthritis by a series of image processing techniques which have been termed to be more effective than the other methods which perform the same task and hence provide a more effective approach in computer aided diagnosis. This is a huge step in the medical field and can be utilized further. The system may be further enhanced by the improvement of the edge detection as well as finding a better segmentation technique. Gray level co-occurrence matrix (GLCM) features like Mean, Median, Energy, Correlation, Bone Mineral Density (BMD) and etc. After finding all the features it stores in the database. This dataset is trained with inflamed and non-inflamed values and with the help of neural network all the new images are checked properly for their status and Rough set is implemented for further optimization of results.

REFERENCES

[1] Aman Kumar Sharma and SuruchiSahni, "A Comparative Study of Classification Algorithms for Spam Email Data Analysis", *International Journal on Computer Science and Engineering (IJCSE),* Vol. 3, No 5, 2011.
[2] Anache, N., "Medical image analysis a challenge for computer vision research," *Pattern Recognition, 1998. Proceedings. Fourteenth International Conference on* , vol.2, no., pp.1255,1256 vol.2, 16-20 Aug 1998.
[3] Anoraganinrum, "Cell Segmentation with Median Filter and Mathematical Morphology Operation", *International Conference on Image Analysis and Processing*, pp. 1043-1046, 1999.
[4] Bharanidharan, T., and D.K., Ghosh, 2012. "A Two Dimensional Image classification Neural Network for Medical Images", *European Journal of Scientific Research*,Vol.74 No.2 (2012), pp. 286-291.
[5] ChokkalingamS.P and K. Komathy, 2014. Classification and Segregation of Abnormal Lymphocytes through Image Mining for Diagnosing Rheumatoid Arthritis Using Min-max Algorithm. Research Journal of Applied Sciences, Engineering and Technology, 7(18): 3926-3934.
[6] ErmaiXie, T., M., McGinnity, QingXiang, Wu, "Automatic Extraction of Shape Features for Classification of Leukocytes", *International Conference on Artificial Intelligence and Computational Intelligence, 2010.*
[7] http://minitab.en.softonic.com.
[8] Krishnapuram, R., and J. M., Keller, "A possibilistic approach to clustering," *IEEE Transactions on Fuzzy Systems,* Vol. 1, No. 2, pp. 98-110, 1993.
[9] NiponTheera-Umpon and SompongDhompongsa, "Morphological Granulometric Features of Nucleus in Automatic Bone Marrow White Blood Cell Classification". *IEEE Trans.Infor.Tech.Bio.Medi,* Vol. 11, No. 2, 2007.
[10] Pavlova, P.E., K.P., Cyrrilov, and I. N., Moumdjiev, "Application of HSV colour system in the identification by colour of biological objects on the basis of microscopic images," *Computerized Medical Imaging and Graphics* 20(5), 1996, pp.357-64.
[11] SubrajeetMohapatra, DiptiPatra, and SanghamitraSatpathi, "Image Analysis of Blood Microscopic Images for Acute Leukemia Detection". *International Conference on Industrial Electronics Control and Robotics*, 2010.
[12] SubrajeetMohapatra, DiptiPatra, and Kundan Kumar, "Blood Microscopic Image Segmentation using Rough Sets", *International Conference on Image Information Processing* ,2010.
[13] Tabrizi P.R, S.H Rezatifighi and M.J. Yazdanpanah, "Using PCA and LVQ Neural Network for Automatic Recognition ofFive Types of White Blood Cells", *Inter. Conf. of the IEEE EMBS Buenos Aires, Argenti, 2010.*
[14] Tycko, D.H., S., Anbalagan, H.C., Liu, and L., Ornstein, "Automatic leukocyte classification using cytochemically stained smears," *J HistochemCytochem.* pp. 178-94, 1976.
[15] Wei Xiong, Sim-Heng Ong, joo-Hwee Lim, Kelvin Weng Chiong Foong, Jiang Liu, Daniel Racoceanu, Alvin G. L., Chong and Kevin S.W., and Tan, "Automatic area classification in peripheral blood smears", *IEEE Trans.Bio.Engg,* Vol.57, No. 8, 2008.
[16] www.cs.waikato.ac.nz/ml/weka
[17] www.cs.utexas.edu/users/ml/tutorials/Weka-tut
[18] www.ehow.com "rheumatoid arthritis can cause low lymphocytes"
[19] www.minitab.com/en-US/minitab-statistical-software-40-years.aspx
[20] Yampri, P., C.Pintavirooj, S., Daochai and S., Teartulakarn, "White Blood Cell Classification based on the Combination of Eigen Cell and Parametric Feature Detection", *IEEE Conference on Industrial Electronics and Applications, 2006.*

Chokkalingam. SP completed his B.SC(Maths) in 1995 and MCA in 1999 from Bharathidasan University, Tiruchirappalli, India and obtained his M.Tech (CSE) with Gold medal from MGR University, Chennai, India and Pursuing Doctorate degree in Computer Science and Engineering in Anna University, India. Heacquired his expertise from 3 years of Industrial experience, 15 years from academic and 2 years from research works. His recent

research publications of about 5 research papers are based on Medical Image processing and Data Mining.

Komathy K graduated from Madras University, India in 1982, obtained her Master and Doctorate degree in Computer Science and Engineering from Anna University, India. She acquired her expertise from 8 years of Industrial experience, 15 years from academic and 5 years from research works. Her recent research publications of about 50 research papers are based on wired and wireless networking, network security, social networking and artificial intelligence. Under her supervision, six students are pursuing their Doctorate degree. She has also received a research fund from AICTE India, for monitoring air pollution of metro roads using VANET.