ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

Development Of Novel Diagnostic Tool For Assessment Of Osteoarthritis.

M Subramoniam*, S Poornapushpakala, and S Barani.

Department of Electronics &Communication Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India.

ABSTRACT

Arthritis is a disease that occurs among all age group of people in the society. The disease emerges with mild joint stiffness and later leads to joint immobility. This disease cannot be cured but can be controlled at any stage of diagnosis. But the medical challenge is that, the occurrence of this disease at the early stage cannot be diagnosed with any of the diagnostic tools or methods followed in current medical society. This paper discusses a novel diagnostic tool developed for the assessment of arthritis from digital X-ray images. The tool was developed using Haralick Features extraction method and Bayesian classification algorithm for diagnosing abnormality in the bone joints under analysis. The performance of this developed tool is assessed using various statistical parameters such as accuracy, precision, specificity and sensitivity. The performance report produced by this diagnostic tool is satisfactory.

Keywords: Osteoarthritis, Bayes Classifier, Haralick Features, Radiographs

Corresponding author

INTRODUCTION

The application of engineering in the medical field has ease up the diagnostic procedures and non-invasive. They also act us life support tools in many complex scenarios. This paper also discusses one such tool developed for the diagnosis of osteoarthritis. This tool provides additional support to the physicians for the analysis of radiographic images in the diagnosis of osteoarthritis by classifying the bone joints under study into normal or abnormal subjects. Researchers throughout the world have proposed many such tools and algorithms for the diagnosis of Osteoarthritis. A few of them have been discussed below.

Ali Nouri et al. (2016) proposed a non-parametric window method for analyzing thermo graphic images. The temperature distribution over the joints is the parameter taken for the study. On evaluation this algorithm produced an error rate of 5 %.This method aids in obtaining the information related to the joints under study from the image.

Andreas Koch et al. (2017) developed a 3-D probe for imaging the bone joints affected by RA. The image formed with this probe using echo tomography provided more information than the image formed using 2-D probe. The signal acquisition using 3-D probe can be extended for the analysis of breast cancer.

Andrew D Weins (2016) developed an algorithm for analyzing the bone joints affected by arthritis. This computer vision based system used the signal from the accelerometers, electromyograms, LED's and eye cameras for analysis of the joints under study. The signal obtained with this method can be stored for future analysis, which forms the major advantage of this algorithm.

Fbiola et al.(2015) made a comparative study on various feature selection algorithms for the quantification of bone joints affected by RA. Algorithms such as Binomial distribution, Chi-square information gain, GINI and DKM were the methods taken for the study. The FS algorithm was applied on the data base obtained from Abertawae Bro Morgannwg University Health board. The final results concluded that the Chi-Square feature selection process is much superior to the other feature selection algorithms.

Gopi Krishnan .M et al. (2016) developed an algorithm for segmentation of RA affected portions from thermo graphic images. FCM technique was used for this study. The segmented portions were then subjected to statistical analysis to differentiate the normal and the abnormal portions in the segmented image. This algorithm plays as an automated tool for segmenting the radiographic images.

Kunlincao et al.(2016) developed an automated image analysis tool for analyzing the bone joints affected by RA. The developed tool was applied on the images acquired using ultrasound technique to locate the joint capsule region. 8 subjects were used for the evaluating the developed algorithm. The final results concluded that the proposed algorithm is much robust over the earlier algorithms.

Mathew chin Heng Chua et al.(2016) designed a device for providing support to the finger joints affected by RA. The device contains elastomeric actuators which can be self operated by the subjects affected by arthritis. This device aids in avoiding further deformation in the bone joints. This device acts as the best rehabilitation device for the subject's affected by RA.

Patrick Leinel et al.(2016) developed an application for sharing the Rheumatoid Arthritis (RA) data among the physicians. The aim of the work developed by the author is to create awareness related to health care among the subjects affected by RA. This tool also helped the researchers to gather data and make analysis for their research.

Seo Hyun Kim (2016) made a study on rehabilitation study by treating the subjects using electrical stimulation on the bone joints affected by arthritis. Currents in the range of micro amperes were applied on the affected bone joints.55 subjects of various classes of RA were treated for the period of three weeks. Noticeable improvement was observed on the subjects treated by this method.

SudhirRathore and S.V Bhalerao (2015) proposed a method using Fuzzy-C Means (FCM) algorithm for identifying the disorders from thermo graphic images. Generally the infra-red radiations emitted from the body will show a difference in wavelength between the normal and disorder portions of the human body. This difference in the IR radiations will in turn produce a change in color while forming the image using thermal imaging method. The author used neuro-fuzzy methods to classify the normal and the abnormal joints. This method of analysis is simple, but they do not provide any information related to the intensity of disorders.

Suma.B eta l. (2016) made a comparative study on various segmentation techniques that can be applied on thermal images for identifying the joints affected by RA. K-means algorithm, color and manual segmentation methods were some methods taken for the study. The author concluded that K-means algorithm proved best segmentation results over the other algorithm. This study also paved the way for the development of automated segmentation algorithm.

YingheHuo et al. (2017) developed an algorithm for evaluating the wrist joint space from radiographic images. The author studied that the joints located around the scaphoid bone are frequently affected by rheumatoid arthritis in most of the subjects. The automated algorithm developed by the author used delineation method and back trace method for quantifying the joint space width. The developed algorithm was evaluated using 50 radiographic images. The accuracy rate reported with this method of analysis is 90 %.

From the survey made with the recent work carried out in the quantification of arthritis the following are the major drawbacks identified.

- Very few research works are carried out for diagnosing arthritis from the conventional X-ray imaging technique.
- The accuracy rate can be improved to have improved diagnostic rate.

Considering the above drawbacks, a simplified system for the diagnosis of arthritis was proposed.

PROPOSED SYSTEM

Radiography is the conventional method followed for diagnosing arthritis. But the early erosions caused of arthritis are not visible in most of the radiographs. Physicians will suggest for other imaging techniques if relevant information is not furnished by the image done using radiographs. However certainty of information is not an assurance with those other imaging techniques also. Moreover, the other imaging techniques are much expensive as compared to radiographic imaging techniques. Hence the work is focussed to diagnose arthritis from radiographic images. The proposed system is framed to define the normal and abnormal conditions from radiographic images for the subjects assumed to have arthritis. The block diagram of the proposed system is given in figure 1.

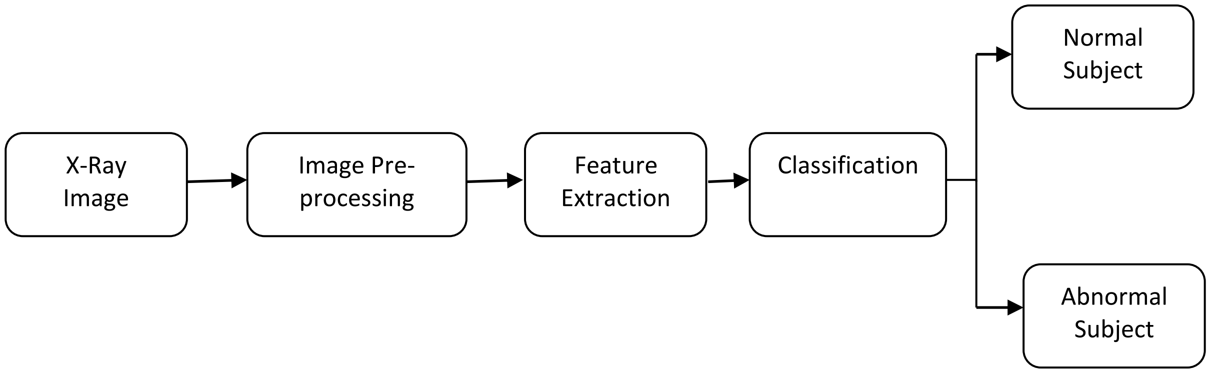

Figure 1: Block Diagram of the Proposed System for Diagnosis of Arthritis

IMAGE PRE-PROCESSING

In this work digital x-ray images of the subjects who are reported to have early morning stiffness, joint swelling and mild pain are taken. The knee bone joint images of these subjects are considered as abnormal subjects for training the images to the algorithms. The knee images of the subjects who are not reported to have any symptoms of Osteoarthritis are trained under normal subjects in the algorithm. These images are collected from Bharat Scans, Chennai. The actual size of the radiographic image is 1000 X 1000 pixels. As arthritis results in the erosion and reduction of joint space, those portions are cropped manually to a size of 200 X 200 pixels. The actual image and the Region of Interest (ROI) image are shown in figure 2.

a) Actual Imageb) ROI Image

Figure 2: Digital X-Ray images of Knee Joints

FEATURE EXTRACTION

In order to analyse and define the normal and abnormal condition of the bone joint under analysis, it is necessary to have the features from the ROI image. Extraction of more features from the image under analysis can provide improved classification which interns leads to better diagnosis. Haralick et al. studied that 13 parameters can be extracted from a grey scale image. These 13 parameters are known as Haralick features. The 13 parameters defined by Haralick and et al. are listed in table.1. Those parameters are estimated using the Grey Level Co-occurrence Matrix (GLCM). With the GLCM, these parameters can be derived for various angles such as 0°, 45°, 90°and 135°on an image. So for each image under analysis, possibly of (13 X 4)=52 features can be derived for the study. In this work, the Haralick features are derived for all the above mentioned angles and those features are trained for classification using Bayes classifier.

Table 1: List of Haralick Features

Energy	Sum Entropy
Contrast	Difference Entropy
Corelation	Information Measures of Corelation-1
Inverse Difference Moment	Information Measures of Corelation-2
Sum Variance	Sum of Variance
Difference Variance	Sum Average
Entropy	

BAYES CLASSIFIER

Bayes classifier is a simple probabilistic classifier based on applying Bayes theorem from Bayesian statistics with independence assumptions. It assumes that the presence or absence of a particular feature of a class is unrelated i.e. independent to the presence or absence of any other feature. Depending on the nature of the probability model, this classifier can be trained very efficiently in a supervised learning environment. An advantage of the Bayes classifier is that it only requires a small amount of training data to estimate the parameters e.g. means and variances of the variables necessary for classification.

Because independent variables are assumed, only the variances of the variables for each class need to be determined and not the entire covariance matrix. The probability model of a classifier is a conditional model, p(C/F1… … ., Fn) over a dependent class variable C with a small number of outcomes or classes, conditional on several feature variable F1 through Fn .

If the number of features n is large or if a feature takes on a large number of values, then basing such a model on probability tables is infeasible. For such cases the model can be reformulated as given in equation 1, and 2.

$$p(C/F1.......,Fn) = \frac{p(C)\,p(F_1......F_n/C)}{p(C)(F_1......F_n)} \quad (1)$$

It can be simplified as,

$$posterior = \frac{prior \times likelihood}{evidence} \quad (2)$$

The above expression is known as Bayes theorem and the classification is done on accordance with the above expression. Kernel-based methods are most popular non-parametric estimators. These Kernel Smoothers are used to estimate the probability density values to calculate the posterior of the posterior than the actual probability values. The Kernel Smoother produces a curve formed by repeatedly finding a locally weighted fit of a simple curve at sampled points in the domain.

Let $X_1, X_2 .... X_n$ be a random samples with a continuous, univariate density f. The Kernel smoother for the function is given by equation 3

Where

$$\hat{f}(x,h) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h}\right) \quad (3)$$

k	-	Kernel function
h	-	Bandwidth
n	-	Number of Samples.

As the n increases, h decreases and under this condition the kernel estimate converges in probability to true density. The kernel k can be unimodel or symmetric about zero. With this k value the influence of each point is spread about its neighbourhood. Finally the contribution from each point is summed for overall estimation of the probability density function. The bandwidth h is the scaling factor. The width of probability mass around a point depends on this value of h. It controls the smoothness and the roughness of the density estimate .So proper choice of bandwidth value is needed as the value can under or over smooth the function.

In this work three kernel functions such as normal or Gaussian, triangle and box kernel functions are used to estimate the kernel smoother are used. The default kernel function in Matlab is normal kernel function which has its bandwidth values h at 0.1. The ranges of h values used in this work for triangle and box kernel function are 0 to 1 and - 0.5 to 0.5 respectively. As better smoothening is obtained with these values in the work these values are fixed for the kernel functions. Bayes classifier has been used in many image classification applications.

VALIDATION

The validation of the work is carried out using K –fold cross validation method. In this the value of K chosen is 5. A total of 125 samples which includes 25 normal and 100 abnormal subjects are taken for the study. For each fold, one third of the samples are used as training set and the remaining will forms the testing set. The statistical parameters such as accuracy, precision, sensitivity, specificity F-Score and False Prediction Ratio (FPR) are the parameters taken for evaluation. The results obtained for each degree of angle in Haralick features is given from table 2-5.

Table 2: Results with 0˚Haralick features and Bayes Classifier

Kernel function	K-Fold	Performance Metrics (X 100 %)					
		Accuracy	Precision	Sensitivity	F-score	FPR	Specificity
Normal	1	0.788	0.750	0.692	0.720	0.150	0.850
	2	0.818	0.818	0.90	0.857	0.308	0.692
	3	0.909	0.824	1.000	0.903	0.158	0.842
	4	0.758	0.818	0.600	0.692	0.111	0.889
	5	0.818	0.813	0.813	0.813	0.176	0.824
	Avg.	**0.818**	**0.804**	**0.801**	**0.797**	**0.181**	**0.819**
Triangle	1	0.788	0.750	0.692	0.720	0.150	0.850
	2	0.818	0.818	0.90	0.857	0.308	0.692
	3	0.848	0.737	1.000	0.848	0.263	0.737
	4	0.818	1.000	0.600	0.750	0.000	1.000
	5	0.818	0.813	0.813	0.813	0.176	0.824
	Avg.	**0.818**	**0.824**	**0.801**	**0.798**	**0.179**	**0.821**
Box	1	0.788	0.800	0.615	0.696	0.100	0.90
	2	0.788	0.810	0.850	0.829	0.308	0.692
	3	0.818	0.786	0.786	0.786	0.158	0.842
	4	0.788	0.90º	0.600	0.720	0.056	0.944
	5	0.848	0.867	0.813	0.839	0.118	0.882
	Avg.	**0.806**	**0.832**	**0.733**	**0.774**	**0.148**	**0.852**

Table 3: Results with 45˚Haralick features and Bayes Classifier

Kernel function	K-Fold	Performance Metrics (X 100%)					
		Accuracy	Precision	Sensitivity	F-score	FPR	Specificity
Normal	1	0.788	0.714	0.769	0.741	0.200	0.800
	2	0.758	0.750	0.90	0.818	0.462	0.538
	3	0.879	0.778	1.000	0.875	0.211	0.789
	4	0.788	0.90	0.600	0.720	0.056	0.944
	5	0.818	0.778	0.875	0.824	0.235	0.765
	Avg.	**0.806**	**0.784**	**0.829**	**0.795**	**0.233**	**0.767**
Triangle	1	0.788	0.714	0.769	0.741	0.200	0.800
	2	0.758	0.750	0.90	0.818	0.462	0.538
	3	0.818	0.700	1.000	0.824	0.316	0.684
	4	0.788	0.90	0.600	0.720	0.056	0.944
	5	0.788	0.737	0.875	0.800	0.294	0.706
	Avg.	**0.788**	**0.760**	**0.829**	**0.780**	**0.265**	**0.735**
Box	1	0.788	0.714	0.769	0.741	0.200	0.800
	2	0.727	0.739	0.850	0.791	0.462	0.538
	3	0.788	0.684	0.929	0.788	0.316	0.684
	4	0.788	0.90	0.600	0.720	0.056	0.944
	5	0.788	0.737	0.875	0.800	0.294	0.706
	Avg.	**0.776**	**0.755**	**0.805**	**0.768**	**0.265**	**0.735**

ISSN: 0975-8585

Table 4: Result for 90ºHaralick features and Bayes Classifier

Kernel function	K-Fold	Performance Metrics (X100%)					
		Accuracy	Precision	Sensitivity	F-score	FPR	Specificity
Normal	1	0.758	0.778	0.538	0.636	0.100	0.90
	2	0.667	0.765	0.650	0.703	0.308	0.692
	3	0.848	0.800	0.857	0.828	0.158	0.842
	4	0.545	0.500	0.333	0.400	0.278	0.722
	5	0.727	0.706	0.750	0.727	0.294	0.706
	Avg.	**0.709**	**0.710**	**0.626**	**0.659**	**0.227**	**0.773**
Triangle	1	0.818	0.818	0.692	0.750	0.100	0.90
	2	0.667	0.765	0.650	0.703	0.308	0.692
	3	0.788	0.706	0.857	0.774	0.263	0.737
	4	0.576	0.556	0.333	0.417	0.222	0.778
	5	0.818	0.778	0.875	0.824	0.235	0.765
	Avg.	**0.733**	**0.724**	**0.682**	**0.693**	**0.226**	**0.774**
Box	1	0.788	0.750	0.692	0.720	0.150	0.850
	2	0.667	0.765	0.650	0.703	0.308	0.692
	3	0.788	0.706	0.857	0.774	0.263	0.737
	4	0.576	0.556	0.333	0.417	0.222	0.778
	5	0.818	0.778	0.875	0.824	0.235	0.765
	Avg.	**0.727**	**0.711**	**0.682**	**0.687**	**0.236**	**0.764**

Table 5: Result with 135˚Haralick Features and Bayes Classifier

Kernel function	K-Fold	Performance Metrics (X 100 %)					
		Accuracy	Precision	Sensitivity	F-score	FPR	Specificity
Normal	1	0.788	0.714	0.769	0.741	0.200	0.800
	2	0.848	0.857	0.90	0.878	0.231	0.769
	3	0.848	0.737	1.000	0.848	0.263	0.737
	4	0.758	0.818	0.600	0.692	0.111	0.889
	5	0.788	0.737	0.875	0.800	0.294	0.706
	Avg.	**0.806**	**0.773**	**0.829**	**0.792**	**0.220**	**0.780**
Triangle	1	0.788	0.714	0.769	0.741	0.200	0.800
	2	0.848	0.857	0.90	0.878	0.231	0.769
	3	0.788	0.667	1.000	0.800	0.368	0.632
	4	0.758	0.818	0.600	0.692	0.111	0.889
	5	0.788	0.737	0.875	0.800	0.294	0.706
	Avg.	**0.794**	**0.759**	**0.829**	**0.782**	**0.241**	**0.759**
Box	1	0.818	0.769	0.769	0.769	0.150	0.850
	2	0.818	0.850	0.850	0.850	0.231	0.769
	3	0.758	0.650	0.929	0.765	0.368	0.632
	4	0.788	0.90	0.600	0.720	0.056	0.944
	5	0.788	0.737	0.875	0.800	0.294	0.706
	Avg.	**0.794**	**0.781**	**0.805**	**0.781**	**0.220**	**0.780**

In order to analyze the performance, a study is also made by aggregating the Haralick features derived over the individual angles. The results obtained with aggregated Haralick features are shown in table 6.

Table 6: Result with aggregated Haralick Features and Bayes Classifier

Kernel Function	K-Fold	Performance Metrics (X 100 %)					
		Accuracy	Precision	Sensitivity	F-score	FPR	Specificity
Normal	1	0.818	0.889	0.615	0.727	0.050	0.950
	2	0.848	0.857	0.90	0.878	0.231	0.769
	3	0.879	0.778	1.000	0.875	0.211	0.789
	4	0.818	1.000	0.600	0.750	0.000	1.000
	5	0.818	0.813	0.813	0.813	0.176	0.824
	Avg.	**0.836**	**0.867**	**0.786**	**0.809**	**0.134**	**0.866**
Triangle	1	0.879	0.909	0.769	0.833	0.050	0.950
	2	0.848	0.857	0.90	0.878	0.231	0.769
	3	0.818	0.700	1.000	0.824	0.316	0.684
	4	0.818	1.000	0.600	0.750	0.000	1.000
	5	0.909	0.842	1.000	0.914	0.176	0.824
	Avg.	**0.855**	**0.862**	**0.854**	**0.840**	**0.155**	**0.845**
Box	1	0.909	1.000	0.769	0.870	0.000	1.000
	2	0.879	0.90	0.90	0.90	0.154	0.846
	3	0.818	0.700	1.000	0.824	0.316	0.684
	4	0.818	1.000	0.600	0.750	0.000	1.000
	5	0.939	0.889	1.000	0.941	0.118	0.882
	Avg.	**0.873**	**0.898**	**0.854**	**0.857**	**0.117**	**0.883**

An improvement is classification rate is observed with aggregated Haralcik features. A further evaluation is also made using aggregated Haralick Features and by using hybrid kernel smoothers. The result obtained with this combination is given in table 7.

Table 7: Result with Aggregated Haralick Features and Hybrid Kernels

	Performance Metrics (X 100 %)					
Angle	**Accuracy**	**Precision**	**Sensitivity**	**F-score**	**FPR**	**Specificity**
0	0.879	0.90	0.90	0.90	0.154	0.846
45	0.848	0.857	0.90	0.878	0.231	0.769
90	0.831	0.788	0.867	0.825	0.200	0.800
135	0.879	0.90	0.90	0.90	0.154	0.846
Hybrid	**0.939**	**0.889**	**1.000**	**0.941**	**0.118**	**0.882**

ISSN: 0975-8585

RESULTS AND DISCUSSION

From the results it is observed that improved classification rate is obtained with aggregated kernel smoothers and hybrid kernel functions. The accuracy achieved with this study is 93.9 percentages. This accuracy rate higher and satisfactory as compared as compared to the techniques discussed in the literature survey. This shows the performance rate of Bayesian classifier higher than the other classification algorithm in the diagnosis of Ostoarthritis. The study carried out in this method, purely focused the subjects who are reported with the early symptoms of Osteoarthritis. From the classification rate reveals that this algorithm is diagnose the early erosions in bone joints from the conventional radiographic images. Therefore this tool can be used for diagnosing the early stage of Osteoarthritis.

REFERENCES

[1] Ali Nouri; RasoulAmirfattahi, "Mutual information based detection of thermal profile in hand joints of rheumatoid arthritis patients using Non – Parametric windows", Proceedings of IEEE Canadian Conference on Electrical and Computer Engineering ,PP-1-4,2016.

[2] Andreas Koch ; Florian Stiller, "3D-pulse-echo-tomography for breast cancer and rheumatoid arthritis diagnosis – add-on-system and latest in vivo results", Proceedings of IEEE International Symposium on Unltrasonics , PP-1-4,2016.

[3] Andrew D Wiens ;SampathPrahalad, "Vibro CV: A Computer vision-based vibroarthography platform with possible application to Juvenile idiopathic arthritis", Proceedings of IEEE 8th International Conference on Engineering in Medicine and Biology Society, PP-4431-4434,2016.

[4] B Suma ; U Snekkalatha, "Automated thermal image segmentation of knee rheumatoid arthritis", Proceedings of International Conference on Communication and Signal Processing, PP-535-539,2016.

[5] Fabiola Fernandez-Gutierrez; "Comparing feature selection methods high dimensional imbalanced data; identifying Rheumatoid arthritis cohorts from routine data", Proceedings of International Conference on Industrial Engineering and Systems Management, PP-236 - 241,2015.

[6] KunlinCao;David M Mills; "Toward Quantative Assessment of Rheumatoid Arthritis Using Volumetric Ultrasound", IEEE Transactions on Biomedical Engineering ,Vol.63,issue-2,PP-449-458,2015.

[7] M Gopikrishnan , T Rajalajshmi , "Diagnosis of Rheumatoid arthritis in knee using fuzzy C means segmentation technique", Proceedings of international Conference on Communication and Signal Processing, PP-430-433,2016

[8] Matthew Chin HengChua;LimJeongHoon, "Design and evaluation of Rhematoid Arthritis rehabilitative Device (RARD) for laterally bent fingers", Proceedings of 6 thIEEE International Conference on Biomedical Robotics and Biomechatronics, PP-839 - 843,2016.

[9] Patrick Leiniel H Domingo; Jaime D L Caro , "Philippine app for Rheumatoid Arthritis Data Exchange (PARADE)", Proceedings of 7th International Conference on Information,Intelligence,Systems& Applications, PP-,2016

[10] Robert M.Haralick and K.Shanmugam. "TexturalFeatures for Image Classification", IEEETransactions on Systems,Man andCybernetics,Vol.3,Issue 6,pp 610-621.

[11] Seo Hyun Kim, Hana Lee, "Effectiveness of micro-current electrical stimulation for treating Rheumatoid Arthritis", Proceedings of IEEE EMBS International Student Conference, PP-1-4,2016.

[12] SudhirRathore; S V Bhalerao, "Implementation of neuro-fuzzy based portable thermographic system for detection of Rhematoid Arthritis", Proceedings of Global Conference on Communication Technologies, PP-,2015.

[13] YingheHuo,Koen.I.Vincken, "Automatic Quantification of radiographic wrist joint space width of patients with rheumatoid arthritis", IEEE Transactions on Biomedical Engineering ,Vol.63,Issue-10,PP-2177-2186,2016.

[14] YingheHuo,Koen.I.Vincken, "Automatic Quantification of radiographic Finger Joint Space width of patients with early Rhematoid Arthritis", IEEE Transactions on Biomedical Engineering,Vol.63,Issue-10,PP-1-10,2016.