*Stochastic Modelling and Computational Sciences*

# CLASSIFICATION OF OSTEOARTHRITIS AND RHEUMATOID ARTHRITIS BASED ON ENSEMBLE MACHINE LEARNING

**R. Ranjani [1] and Dr. L. Thara [2]**
[1]Research Scholar, Department of Computer Science, P.S.G. College of Arts & Science and Assistant Professor, Department of Computer Science with Data Analytics, Dr.N.G.P. Arts and Science College, Coimbatore.
[2]Associate Professor and HOD, Department of MCA, P.S.G. College of Arts & Science
[1]mails2rranjani@gmail.com and [2]thara@psgcas.ac.in

## ABSTRACT
*Osteoarthritis is a non-inflammatory condition that results in the degeneration of cartilage, the flexible substance that softens the gap between bones. Degenerative joint disease and osteoarthritis are other names for the same condition. It is the most prevailing type of arthritis and commonly brought on by the deterioration of a joint over a lifetime. Rheumatoid arthritis is an autoimmune disease, RA's hallmarks include cartilaginous degradation, bone erosions, and joint deformity, the disease's progression can be highly unpredictable. Both hereditary and conservational variables may participate to the onset of the disease; a connection with HLA-DR4 has been found. This paper is focused on classifying the conditions of RA and OA, because both forms of arthritis have common Symptoms. It is very important to classify RA from OA at the initial stage itself, then only doctors can provide better treatment for patients. Early detection in OA can avoid the joint disability and RA can slow down the disease progression.*

***Keywords:*** *Osteoarthritis, Rheumatoid Arthritis, Autoimmune disease, joint deterioration.*

## I. INTRODUCTION
Knee Osteoarthritis disorder causes all structural components of the joint has undergone change in a pathologic manner, eventually resulting in joint failure Loss of articular cartilage is the pathophysiologic precursor form for OA, which is accompanied by the subchondral bone plate thickening and sclerosis, articular capsule stretching, growth of osteophytes at the joint's edge and the deterioration of the muscles that bridge the joint. Joint injury is frequently the first to happen when protective systems have failed, even though there are numerous other possible paths that could result in OA. It most frequently manifests in the hips, knees, hands and spine. The tiny joints in fingers and base thumb are the ones, osteoarthritis most frequently affects in the hand. The development of RA is influenced by synovial hyperplasia, synovial lymphocytic infiltration, locally localized cytokines and chemokines produced by activated macrophages, fibroblasts, and lymphocytes as well as an immunologically mediated process. The thin membrane lining the joints is attacked by the immune system due to the inflammatory and autoimmune nature of rheumatoid arthritis.

## II. LITERATURE SURVEY
Numerous antibody systems have been discovered in RA. These include antibodies against carbamylated(anti-CarP) and citrullinated (ACPA) protein, both of which have been used as diagnostic indicators (Fig. 1). Synovial fluid (SF) and serum are the main locations of these autoantibodies in RA patients [1]. (The classification based on 2010 ACR/EULAR criteria [2]. Joint involvement, provocative indicators such as ESR and CRP, illness duration, and serology (RF and ACPA autoantibodies) scores are all included in this classification scheme. In those with undifferentiated arthritis, the development of RA [3] and a poorer clinical course with greater joint erosions are both associated with the existence of autoantibodies. By means of an enzyme reaction with PAD(A), citrullination is the process by which an arginine is changed into a citrulline. PAD might be produced by bacteria or secreted by neutrophils. Carboxylation is a chemical reaction that converts lysine into homocitrulline through the interaction with cyanate (B).Numerous factors might lead to an elevated cyanate level, such as smoking, inflammation and renal illness. Image modified from references [4,5].

 

Autoantibodies can recognize these PTMs. The anti-Carp, anti-carbamylated protein and RF were some of the most prevalent antibodies during the rheumatoid arthritis diagnosing process. Ultrasound imaging method assesses synovial and emission changes. A minor flare-up of deterioration and bone rarefaction in the joint, which is incredibly painful and only becomes worse over time, might be acted out using an aggravation. Additionally, it could result in numerous flaws and practical problems in the bones. In this way, it is typically crucial to detect such issues in the early stages. All medications, therapies, and interventions used up to this point have only served to tame the infection's symptoms and slow its spread. We can comprehend that the causes of RA are unknown, and there are now no known treatments for these excruciating cases.

The sum of the points from each of the domains A through D is taken as the final score. A patient must have a total score of 6 in order to be diagnosed with definite RA. patient radiographic data gathered today. The conventional disparities between ACPA-positive and negative have vanished as a result of improvements in disease progression and cartilage damage as well as in disease outcomes that tend to be most significant among patients (e.g., pain, fatigue, workability) [6]. In other words, both ACPA-positive and ACPA-negative cases are severe in terms of a number of outcomes, RA has similarized with current therapeutic approaches, with the exception that a higher percentage of patients with ACPA-negative RA appear to reach drug-free remission. The best treatment plans are those that are based on the pathogenesis that underlies the symptoms of the disease. It is likely that the disease progression is partly similar in seropositive and seronegative illness, which could end up resulting in similar treatment plans, and somewhat different, which would lead to various treatment strategies for diseases that are both seropositive and seronegative. Only disease progression can be monitored by measurements of bone and cartilage, such as by calculating temporal variations in bone lesions [7].

Recently, methods have been developed to assess synovitis directly. A CNN is trained to extract reality from the entire image using the visual score as a foundation [8]. To assess the surrounding synovium and estimate inflammation using contrast-enhanced magnetic resonance imaging, measurement of tenosynovitis also requires automated segmentation of tendons. BME could also be automatically determined using static postcontrast wrist MRI imaging, especially in early RA [9]. Numerous traditional AI techniques have been used to estimate cartilage thickness and volume in MRI scans [10]. Deep learning has been successfully applied to knee cartilage detection [11], and more recently, it has been further refined to classify cartilage abnormalities.[12] CNNs have also been used in a similar way to find cartilage in wrist joints [13].AI might be able to identify patterns that are hidden from the human sight and brain or reduce the number of dimensions. This is especially true when clinical records are joined with photographic images to create substantially large data that cannot be processed just by human effort. ML-based decision-making outperformed physician-only decision-making in its initial prospective clinical trials when it came to treating intensive care patients [14].

Random forests algorithm [15], in which samples are categorized by majority vote of all decision trees are an enhancement over decision trees. Less variance and bias result from the use of ensembles. SVMs are trained to adjust the weights of polynomial functions to catch the predominant feasible departure of several classifications. Typically, real data is utilized to train ML systems. Artificial data gathered through emulators, where engagements can be completed by experimentation in order to study various diverse outcomes, can also be used to train ML systems. Simulators can be used to expose Machine Learning approaches to an extensive variety of novel circumstances [16, 17] during training in the areas of automatic driving, gaming and robotics. RNN exists in a division of ANN that can evaluate the patterns of inputs like handwriting, voice, or numerical time series, thanks to their internal memory. In the field of medicine, one type of recurrent neural network LSTMs are used to forecast outcomes in intensive units [18], failure in heart functions [19], or trajectories in health care using medical records [20].

## III.METHODOLOGY

The major goals of this research are to examine ensemble learning methods to enhance the efficiency as well as accuracy of machine learning algorithms in differentiating between rheumatoid arthritis and other types of arthritis. To prevent the data from becoming over-fit, in tree-based techniques classification and regression

models temporal variations in bone lesions processing errors in algorithms, which are used in machine learning. SVM, Naive Bayes, KNN, Random Forest are well-known diversified classifiers with extreme gradient boosting (XGBoost) classifiers to enrich accuracy in classification. Each of these classifiers has a unique architecture and initial learning characteristics. Healthy tissues could be harmed by the immune system protein rheumatoid factor. Your doctor may use elevated rheumatoid factor levels to help make the diagnosis of RA.

**XGBoost:**

The paper discusses ensemble learning methods, notably bagging and boosting methods like XGBoost, Gradient Boost, AdaBoost, and Random Forest. Bagging involves parallel model development by creating random samples from the main data, while boosting is a sequential process that corrects errors of previous models to improve performance. AdaBoost and Gradient Boost are presented as examples of boosting algorithms, each with its approach to iteratively adding weak learners to create a strong classifier.

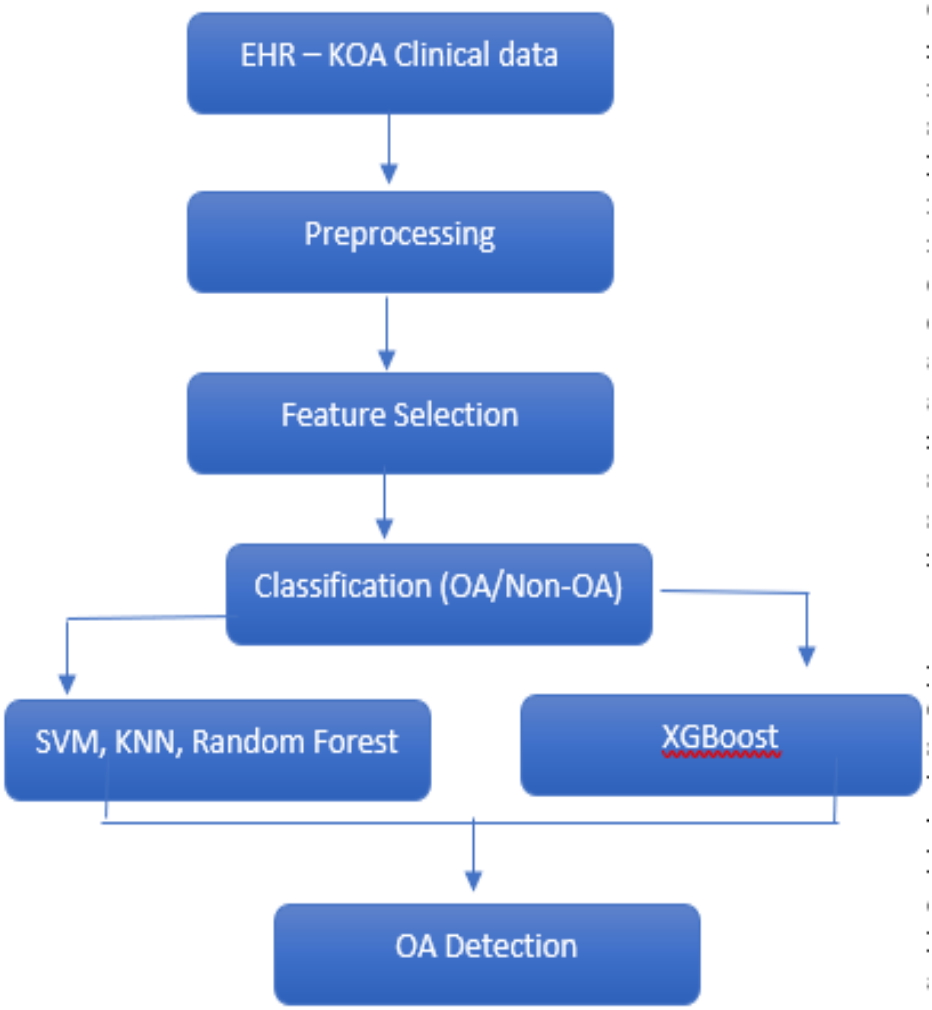


**Fig 1:** Proposed ensemble model

The paper also highlights the features and advantages of XGBoost, including its ability to handle sparse data, cache awareness, built-in cross-validation, and tree pruning. It emphasizes that Gradient Descent is a technique used by XGBoost, an ensemble learning framework and offers flexibility and high computational performance. Regularization, capability to handle sparse data, and built-in cross-validation are presented as unique features of XGBoost.


**Fig 2:** XGBoost Features

**OSTEOARTHRITIS:**

The study examines osteoarthritis (OA) and its impact on cartilage, particularly focusing on Type 2 collagen and proteoglycan aggrecan as key components. It highlights characteristics of OA tendon like aggrecan depletion, collagen matrix unfurling and type 2 collagen loss, which increase vulnerability. OA commonly affects joints like knee, hip, spine and hands with specific joints like DIP, PIP, and thumb base being frequently damaged. The research identifies joint involvement patterns as a diagnostic tool and explores inflammatory arthritis symptoms. It emphasizes the higher prevalence of OA in women and an association with advancing age. The study also emphasizes joint vulnerability and loading as major risk factors for OA. Overall, it reinforces the significance of cartilage alterations in OA, its joint-specific impact, diagnostic considerations, and key risk factors.

Age, gender, race, genetics, obesity, proprioceptive deficits, malalignment, previous damage, joint trauma, and nutritional variables are some of the factors that affect OA. The lower leg bones (tibia and fibula) and femur, which makes up the thigh bone, come together to form the knee joint. Bones frequently rub against one another as a result of osteoarthritis. The healthy and osteoarthritic knees are shown in this image. In this type of arthritis, bone spurs are frequently present. Usually after many years of use, osteoarthritis sets in. Middle-aged and elderly individuals are affected by obesity; past injuries to the afflicted joint and family history of osteoarthritis are other risk factors for the condition. Radiographs might appear normal but as the condition worsens, they may reveal osteophytes, subchondral cysts and joint space reduction. The central part of the joint surface experiences subchondral erosions, which set individuals apart from rheumatoid and psoriatic arthritis.

**Diagnosis of Osteoarthritis:**

At the University of Michigan Department of Rheumatology, we follow a thorough diagnostic process to identify the type of arthritis you have and develop a successful treatment strategy. We may need to do the following procedures in addition to a physical examination and medical history:

**Arthrocentesis**: utilizing a hollow needle to drain joint fluid

**Arthroscopy:** using a tiny camera on a narrow tube to inspect joints;

**Physical examination:** Closed Synovial Biopsy: removal of a sample of tissue lining for analysis;

**Ultrasounds:** analysis of synovial fluid. X-rays are taken in various positions to demonstrate the extent of joint degeneration, including Narrowing of joint space, Bone thinned or eroded and Bone spurs or other abnormalities.

**RHEUMATOID ARTHRITIS:**

Through damage to joints and soft tissue structures, rheumatoid arthritis is a debilitating and degenerative condition that can have an impact on the appearance and functionality of the hands and other parts of the body. It frequently causes finger joints to distort and bend, which makes mobility difficult. One in every 100 Americans suffers with rheumatoid arthritis, which is two times more prevalent in females than in men. This type of symmetric polyarthritis, the peripheral joints frequently affect PIP and MCP joints, which can result in morning stiffness, discomfort, and swelling in the affected joints. Joint abnormalities can also happen with a protracted inflammatory response. RA associated with certain conditions like a cardiovascular condition, which accounts for the majority of fatalities. More people with RA than the general population have osteoporosis. Diminished LH, DHEA, and testosterone in hypoandrogenism 2-4-fold greater lymphoma risk in individuals with RA.

**Diagnosis of Rheumatoid Arthritis:**

Examining the patient's medical history and joints meticulously required. In between 75 and 80 percent of patients, rheumatoid factor (RF) is found. RF presence is correlated with rigorous illness, nodule and extra-articular characteristics. Antibodies against anti-CCP are more prevalent in patients with severe illness and a propensity for developing bone erosions; they are similarly sensitive to RF but more specific; they may be most helpful in early RA. Additional laboratory data ESR, CBC Analyzing synovial fluid can help rule out infections and crystalline illness. On radiographs, you can see juxta articular osteopenia, shortened joint spaces, and marginal erosions. You ought to purchase a CXR.

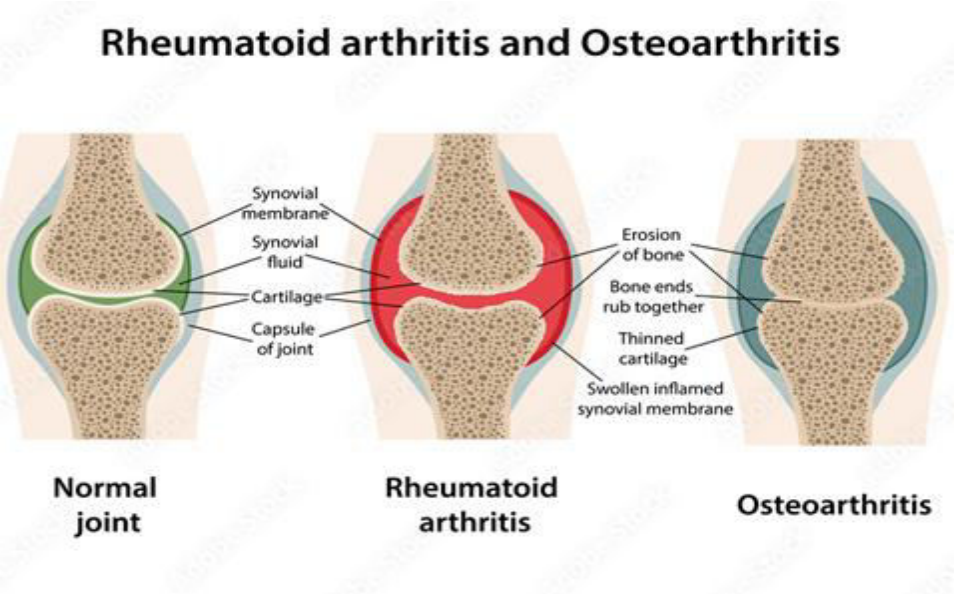


**Fig 3:** Comparison of OA and RA (Referenced from google)

**Table 1:** Feature comparison of Osteoarthritis and Rheumatoid Arthritis

| Feature | Osteoarthritis | Rheumatoid Arthritis |
|---|---|---|
| Prime Factor | Cartilage degradation | Synovium Inflammation |
| Age | >50 | 30-50 |
| Pain | Worsen after activity | Worsen in the Morning / on inactivity |
| Stiffness | <30 minutes | >30 minutes |
| Lab Results | Inverse RF<br>Anti-CCP antibody negative | Favorable RF<br>+anti-CCP antibody<br>Raised ESR and CRP |

| | | |
|---|---|---|
| | Normal CRP and ESR with negative anti-CCP antibodies | |
| Joint Characteristics | Hard and Bony | Soft, Warm and Tender |
| Radiograph Features | Osteopytes<br>JSN<br>Subchondral Sclerosis, cysts | Swelling around the joint<br>Localized Osteoporosis<br>Osteolytic subchondral erosions<br>JSN<br>Cysts and Subchondral Sclerosis |

**Table 2:** OA/RA Categorization Measures

| Field | Type | Point |
|---|---|---|
| 1 | Gender | M/F |
| 2 | Age >50 | 0-1 |
| 3 | BMI | 0-1 |
| 4 | Small Joint Count | 0-14 |
| 5 | Rheumatoid Factor(u/ml) | 0-3 |
| 6 | Anti. CCP(u/ml) | 0-3 |
| 7 | ESR (mm/h) | 0-3 |
| 8 | CRP (md/dl) | 0-3 |
| 9 | ANA | -1,1 |
| 10 | HLA-DRB1 | -1,1,2 |
| 11 | Symptom duration in weeks (>6) | 0-1 |

## IV RESULTS AND DISCUSSION

Knee osteoarthritis is diagnosed by using clinical data which encompasses both physical examination data and laboratory test data. First the features which influences the target variable is identified using Correlation map referred in Fig 4, we can also use logistic regression to identify key features.

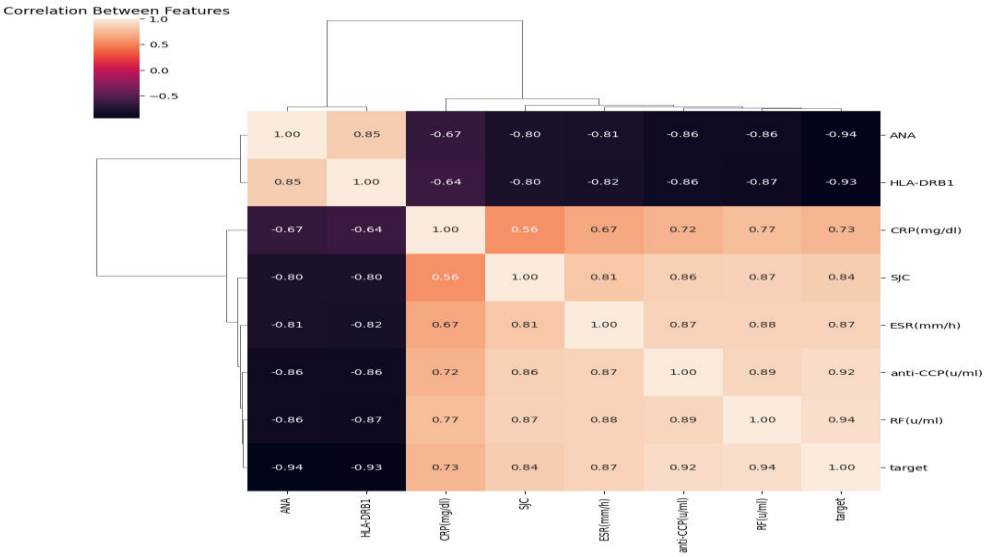


**Fig 4:** Correlation between feature

The research paper focuses on identifying key features for diagnosing Rheumatoid Arthritis (RA) using machine learning approaches. The investigation initially involves in calculating correlation between various features to determine the key ones for RA diagnosis.

**Table 3.** Performance Evaluation

| Classifier | Performance Evaluation Methods | | |
|---|---|---|---|
| | Accuracy % | Precision% | Recall % |
| KNN | 93.3 | 92.8 | 93.3 |
| SVM | 94.63 | 94.21 | 94.51 |
| Random Forest | 94.33 | 94.21 | 94.33 |
| Naive Bayes | 95.61 | 94.89 | 95.62 |
| XGBoost | 96.13 | 96.01 | 96.13 |

Following this, different machine learning models are trained and tested using clinical samples to evaluate their performance. The evaluation is based on metrics such as Accuracy, Precision, and Recall. The paper then introduces an ensemble machine learning algorithm that entails two layers. Several machine learning techniques are applied in the first layer to create models, and output of these models is evaluated with second layer XGBoost algorithm to enhance overall model performance.

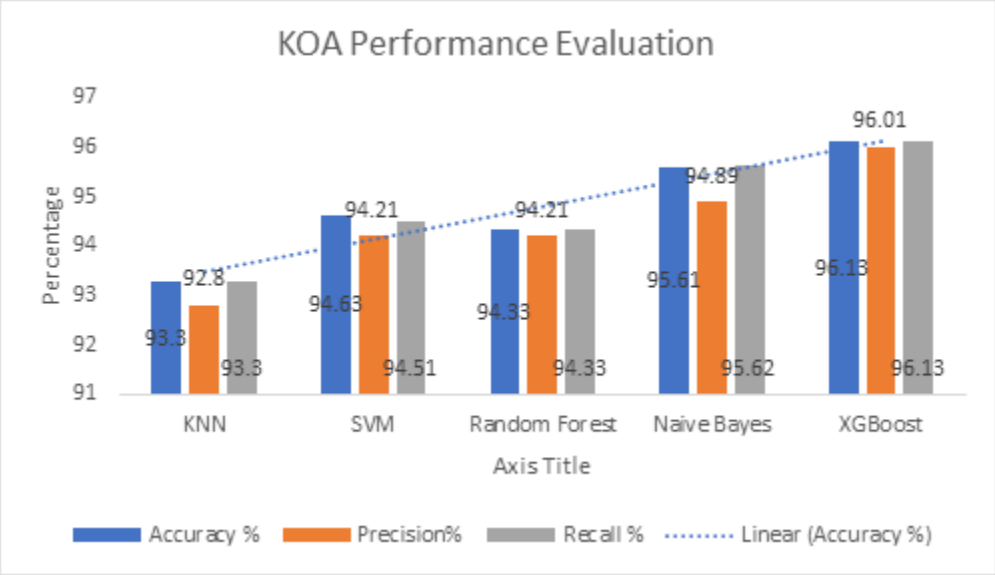


**Fig 5:** Model performance comparison Graph

By applying methods from machine learning, the study intends on the way to increase efficacy and accurateness of RA diagnosis and provides insights into the specific features and models that are most relevant for this purpose.

## V. CONCLUSION

The study discusses the impact of autoantibodies on rheumatology and the potential of AI in the field. The research suggests that rheumatologists could classify patients based on the likelihood of disease progression using clinical, radiological, and biologic parameters. While several prospective biomarkers have been discovered, their accuracy as predictors of therapy response is not yet confirmed. Early identification of arthritis, particularly osteoarthritis (OA) or rheumatoid arthritis (RA), is challenging, demanding the routine of machine learning models progression for quick and effective decision-making by medical professionals and researchers. our

proposed model examine the use of stacked ensemble prediction models to rapidly and precisely identify individuals with early arthritis and those at risk for rapid deterioration. The research suggests that artificial intelligence (AI) can significantly develop rheumatology and improve patient care by facilitating early detection and forecasting of disease progression.

**REFERENCES**

[1]. Chen Et Al,Fully Automatic Knee Osteoarthritis Severity Grading Using Deep Neural Networks With A Novel Ordinal Loss,Published In Final Edited Form As:Comput Med Imaging Graph. 2019 July ; 75: 84–92. Doi:10.1016/J.Compmedimag.2019.06.

[2]. M. R. Karim Et Al.: Deepkneeexplainer: Explainable Knee OA Diagnosis From Radiographs And Magnetic Resonance Imaging,Date Of Publication February 26, 2021, Date Of Current Version March 16, 2021.Digital Object Identifier 10.1109/ACCESS.2021.3062493

[3]. Shamir L, Ling SM, Scott WW Jr, Bos A, Orlov N, Macura TJ, Eckley DM, Ferrucci L, Goldberg IG, Knee X-Ray Image Analysis Method For Automated Detection Of Osteoarthritis, TMBE 56 (2) (2009) 407–415.

[4]. Antony J, Mcguinness K, O'Connor NE, Moran K, Quantifying Radiographic Knee Osteoarthritis Severity Using Deep Convolutional Neural Networks, In: ICPR, 2016, Pp. 1195–1200.

[5]. Dr. J.Deny, P. Nagaraj, V.Maneesha, Dr. B.Perumal, K.Alekhya, S.Amarnath Reddy,Detection Of Osteoarthritis By Using Multiple Edge Detections, IEEE Xplore Part Number: CFP22K74-ART; ISBN: 978-1-6654-1035-9

[6] T. K. Yoo, S. K. Kim, S. B. Choi, D. Y. Kim, And D. W. Kim, ''Interpretation Of Movement During Stair Ascent For Predicting Severity And Prognosis Of Knee Osteoarthritis In Elderly Women Using Support Vector Machines,'' In Proc. 35th Annual. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Jul. 2013, Pp. 192–196.

[7] A. Khatami, A. Khosravi, T. Nguyen, C. P. Lim, And S. Nahavandi, ''Medical Image Analysis Using Wavelet Transform And Deep Belief Networks,''Expert Syst. Appl., Vol. 86, Pp. 190–198, Nov. 2017.

[8]. Minciullo And T. Cootes, ''Fully Automated Shape Analysis For Detection Of Osteoarthritis From Lateral Knee Radiographs,'' In Proc. 23rd Int. Conf. Pattern Recognit. (ICPR), Dec. 2016, Pp. 3787–3791.

[9] A. Tiulpin, J. Thevenot, E. Rahtu, P. Lehenkari, And S. Saarakkala, ''Automatic Knee Osteoarthritis Diagnosis From Plain Radiographs: A Deep Learning-Based Approach,'' Sci. Rep., Vol. 8, No. 1, P. 1727, Dec. 2018.

[10]. Lukas Folle,Deep Learning-Based Classification Of Inflammatory Arthritis By Identification Of Joint Shape Patterns -How Neural Networks Can Tell Us Where To "Deep Dive" Clinically Front Med (Lausanne) 2022 Mar 10;9:850552. Doi: 10.3389/Fmed.2022.850552.

[11]. Neelambuj Chaturvedi,"Deepra: Predicting Joint Damage From Radiographs Using CNN With Attention"May 2022.

[12]. Ahmed, S.M.; Mstafa, R.J. Identifying Severity Grading Of Knee Osteoarthritis From X-Ray Images Using An Efficient Mixture Of Deep Learning And Machine Learning Models. Diagnostics 2022, 12, 2939. Https://Doi.Org/10.3390/ Diagnostics12122939.

[13]. Haider Masood,Eisha Hassan "Osteo-Doc: Kl-Grading Of Osteoarthritis Using Deep-Learning"2022 Ieee.

[14]. S. Aladhadh, R. Mahum: Knee Osteoarthritis Detection Using An Improved Centernet With Pixel-Wise Voting Scheme,Date Of Publication 22 February 2023, Date Of Current Version 8 March 2023.Digital Object Identifier 10.1109/ACCESS.2023.3247502.

[15]. Wang Et Al.: Learning From Highly Confident Samples For Automatic Knee Oa Severity Assessment,Ieee Journal Of Biomedical And Health Informatics, Vol. 26, No. 3, March 2022.

[16] Sharon Ho, *I. Elamvazuthi And CK Lu :Classification Of Rheumatoid Arthritis Using Machine Learning Algorithms ,2018 IEEE 4th International Symposium On Robotics And Manufacturing Automation (ROMA).

[17]. Andersen JKH, Et Al.,Neural Networks For Automatic Scoring Of Arthritis Disease Activity On Ultrasound Images,RMD Open 2019;5:E000891. Doi:10.1136/Rmdopen-2018-000891.

[18]. Gitanjali S. Mate Et Al,An Efficient CNN For Hand X-Ray Classification Of Rheumatoid Arthritis,Hindawi Journal Of Healthcare Engineering Volume 2021, Article ID 6712785, Https://Doi.Org/10.1155/2021/6712785

[19] Q. Zhou, X. He:Broad Learning Model Based On Enhanced Features Earning,Publication March 19, 2019,IEEE.

[20]. Jie Yang∗,Broad Learning With Attribute Selection For Rheumatoid Arthritis,2020 IEEE International Conference On Systems, Man, And Cybernetics (SMC) October 11-14, 2020. Toronto, Canada

[21]. Jun Fukae1*Convolutional Neural Network For Classification Of Two-Dimensional Array Images Generated From Clinical Information May Support Diagnosis Of Rheumatoid Arthritis Scientific Reports (2020) 10:5648

[22]. Sevda Gharehbaghi*,A Feasibility Study On Tribological Origins Of Knee Acoustic Emissionsieee TRANSACTIONS ON BIOMEDICAL ENGINEERING, VOL. 69, NO. 5, MAY 2022. [23].RICHARDSON Et Al.:Quantifying Rheumatoid Arthritis Disease Activity Using A Multimodal Sensing Knee Brace,IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, VOL. 69, NO. 12, DECEMBER 2022

[24]. Zhuang Et Al.: Knee Cartilage Defect Assessment by Graph Representation and Surface Convolution, Ieee Transactions On Medical Imaging, Vol. 42, No. 2, February 2023

[25] S. Thongsuwan Et Al. “Convxgb: A New Deep Learning Model For Classification Problems based On CNN And Xgboost ”Elsevier-Nuclear Engineering And Technology 53 (2021) 522-531.

[26]. Sunday O. Olatunji”A Novel Ensemble-Based Technique For The Preemptive Diagnosis Of Rheumatoid Arthritis Disease In The Eastern Province Of Saudi Arabia Using Clinical Data”,Hindawi Computational And Mathematical Methods In Medicine Volume 2022, Article ID 2339546, 22 Pages.