**Prediction of Knee Replacement** *Prediction of Knee-Replacement using*

*Deep-Learning.*

SINDHU N  
8th Sem CSE department  
Dr.Ambedkar Institution of Technology Bengaluru ,India  
sindhunkotegar@gmail.com

SG ANUSHKA  
8th Sem CSE department  
Dr.Ambedkar Institution of Technology Bengaluru ,India  
anushkasg2017@gmail.com SHRIMAD MISHRA  
8th Sem CSE department  
Dr.Ambedkar Institution of Technology Bengaluru ,India  
snehanandaha2402@gmail.com

SNEHANANDA HA  
8th Sem CSE department  
Dr.Ambedkar Institution of Technology Bengaluru ,India  
[snehanandaha2402@gmail.com](mailto:snehanandaha2402@gmail.com)

***Abstract***

**Knee Osteoarthritis (OA) is a common musculoskeletal disorder found in young adults and older generations. When diagnosed at early stages, lifestyle meddling such as weight loss and exercise can slow the progression of OA, but at later stages, only an intrusive option is available: Knee Replacement (KR). Though, only 2/3 of patients who undergo the KR report their knees feeling “REGULAR” post-operation in a successful procedure, and complications can arise that require emendation. This demands a predictive model to identify a population at higher risk of KR, especially at less advanced stages of OA, such that applicable treatments can be implemented that slow OA progression and delay KR. Here, we present a deep learning predictive model that formulates the X-ray images and clinical and demographic information to predict TKR with AUC 0.86 ± 0.56(p<0.05). In addition, we have developed this predictive model using both Xception and Inception models of CNN, where we compare the graphs for Training and validation Accuracy for both Training and Testing datasets and also the confusion Matrix for both the models used in CNN, thereby identifying accurate KR predicting models for clinical utility.**

***Keywords — OA: Osteoarthritis, KR: Knee Replacement, CNN: Convolitional Neural Network, LR: Logistic Regression. BMI: Body Mass Index, WOMAC: Western Ontario & McMaster Universities osteoarthritis index.***

# Introduction

Knee Osteoarthritis (OA) is one of the most common musculoskeletal disorders in the World, with estimates of its incidence rate ranging from 14 to 30 million. Annual arthritis-related medical expenditures are nearly $150 million, and knee OA together is the 11th highest contributor to global disability. The propensity of knee OA to induce eventual disability can be attributed to structural changes in the joint that characterize the disease, as well as symptoms that can include inflammation, debilitating pain, and functional limitations . Progression of the full-joint disease is typically assessed using the Kellgren-Lawrence (KL) scale, a 0-4 scale in which a higher score is associated with narrowing of the tibiofemoral joint (TFJ) space and other radiographic changes, and thus, a more advanced stage of knee OA When diagnosed at early stages (KL=0, 1), knee OA can be managed through nonsurgical treatment options, including exercise and/or weight loss, oral medications such as acetaminophen or NSAIDs, or intra-articular injections such as corticosteroids and hyaluronic acid, all of which have varying degrees of success in reducing pain . At late stages (KL=4), however, no noninvasive option exists; here, the only option is knee replacement (KR).

KR is an elective procedure in which the knee joint is resurfaced with a metal or plastic implant intended to restore function, provide pain relief, and improve quality of life.

In the United States, estimates of KR incidence lie at 400,000 each year, a figure expected to grow 143% by 2050 even through conservative projections. While TKR is considered one of the most effective procedures in orthopedic surgery, electing for it is far from straightforward: noninvasive alternatives such as weight loss, physical therapy, and NSAIDs are first exhausted. If unsuccessful, a patient will undergo a thorough examination of clinical history and comprehensive imaging of the joint to determine if a TKR is feasible and if so, the desired implant design and size. The procedure is also imperfect: only 66% of patients report their knees feeling “Regular,” and 33% of patients report some degree of pain post-implant.

Furthermore, the implant can fail under some circumstances: periprosthetic joint infection and wound complications can be observed, and implant instability can occur due to aseptic loosening, malpositioning of the implant, and wear of joint components. It is thus much preferable to prolong the good health of the knee, particularly in patients where OA has not advanced to the most severe stages, thereby delaying KR as long as possible. This necessitates a model to identify patients at higher risk of KR such that appropriate treatment options can be followed.

In this study, we formulate a DL-based predictive model that incorporates knee joint images in addition to clinical and demographic information to predict the onset of KR (Fig. 1). We demonstrate that the predictive model uses solely X-ray images, prediction with especially high sensitivity. Furthermore, we show the comparison of performance of models when two different models of Deep Learning (DL) are used for prediction.

.

# Materials and methods

## *Data*

Data was acquired from a prospective observational study conducted by OAI. The dataset followed 4,796 patients and acquired images including X-ray from public-private partnership which had collected the data from the patients and few were also collected from nearby Government and GM hospitals. Details of data collection and study design will be represented in the flow chart. The study was carried out in accordance with all pertinent guidelines and regulations, and written and informed consent was obtained from the hospital for using the X-rays in our work. Table 1 contains the characteristics on which we have selected the data for training the models.

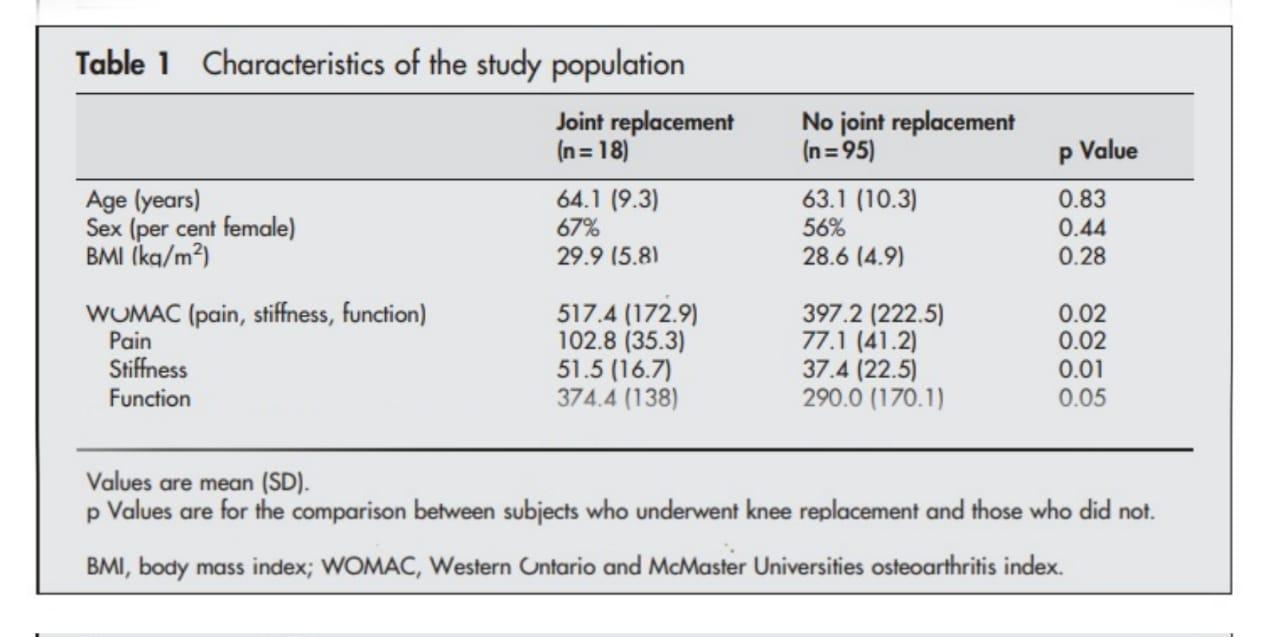
The X-ray images were preprocessed for training and model evaluation. Images were cropped to a 500 × 500 region centered on the knee joint. Briefly, 2D cross-correlation template matching was used to identify a 500 × 500 bounding box centered around the knee joint , and these cases were used to train a CNN architecture , after which the cropped images were normalized. This normalizing approach was initially tested as a strategy to accelerate training of CNN, given the large imaging volumes and large dataset on which it was being trained, believing the approach could suppress information extraneous to eventual KR. The data were then split into training, validation, and test with a 65%,20%,15% split, ensuring entries of any patient were only in one of the three datasets to prevent data leakage. Within the training set, imbalance between KR and non-KR cases was addressed with data augmentation, drawing bootstrap samples from the rare class with replacement.

.

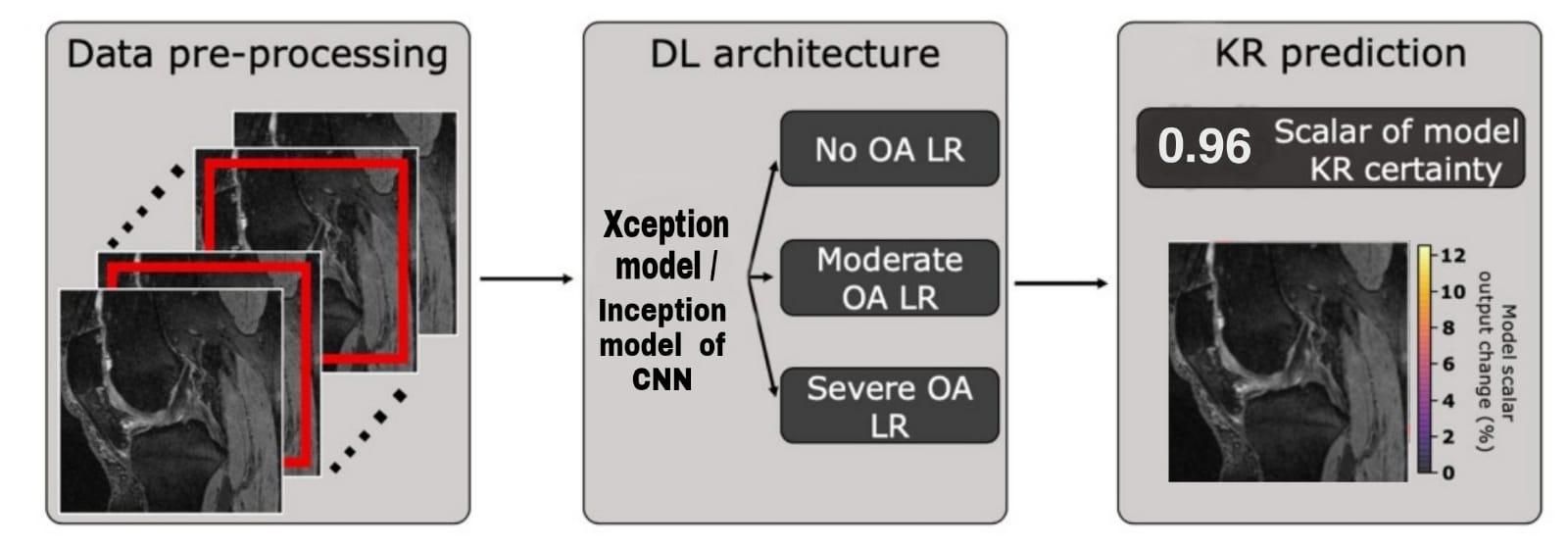
## *Training*

A Xception model was initially pre-trained to predict KR using the entire training set, assessing cross-entropy loss and accuracy on the validation set after completion of each epoch. The pre-train was stopped when validation loss began to increase. The pre-trained model was subsequently fine-tuned to predict KR. We utilized a random search to determine optimal learning rate, dropout rate, weights of the cross-entropy loss  function, and number of layers to freeze during fine-tuning. The search was carried out for 20 iterations, after which a set of parameters were selected that yielded the best combination of accuracy, sensitivity, and specificity on the validation set. Due to computational intensity, the hyper parameter search was not conducted on the entire dataset: for the Xception model.

Later the model fine-tuned using the subset of the training set was further fine-tuned on the entire training set using optimal parameters until validation loss began to increase. The test set was held out during training and predictions for it evaluated just once after fine-tuning, which marked the end of model optimization. Fig 2 shows the flow of the prediction model building and evaluation in a diagrammatic way. The Xception and Inception model algorithm working are also shown in the Fig 3 and Fig 4 respectively.

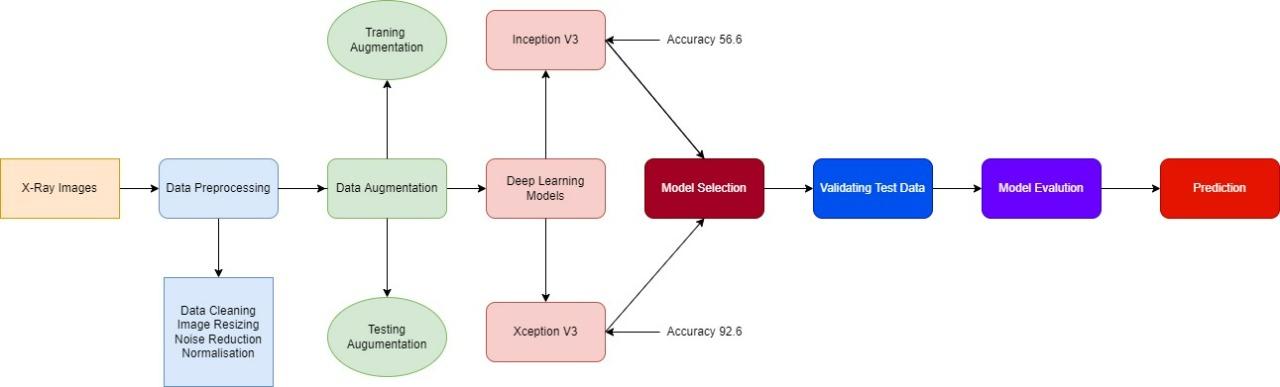


*Table 1. We have collected data based on various characters i.e., Age, Sex and BMI, WOMAC to understand the clinical conditions affecting the severity of the patients’ condition and also the effects of which will be seen in KR prediction.*

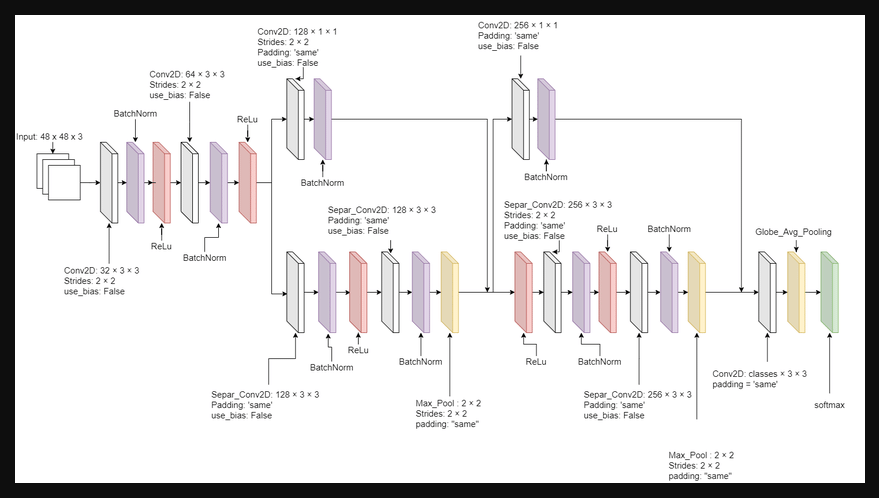


*Figure 1. Models predicting if a patient will undergo KR from X-ray images . X-ray images are center-cropped and cropped to a region centered around the joint, and normalized. Xception / Inception Models are pre-trained to predict KR and fine-tuned to predict KR. Image Based predictions and clinical information are fed to a logistic regression (LR) ensemble based on OA severity. Each ensemble, whose hyper parameters were optimized for Youden’s index in a hyper parameter search, averages predictions of LR models in its OA severity for final KR prediction.*

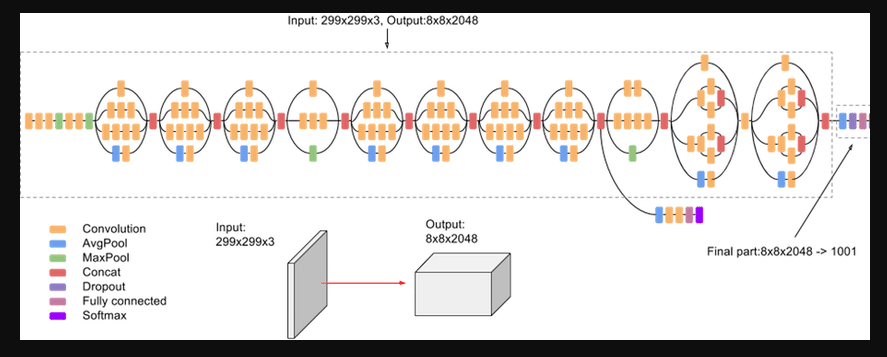


**

*Figure 2. Above Flowchart speaks of the workflow of collecting data, processing and cleaning , augmentation of both training and test sets. Building the Xception and Inception models , validating the test data along with model evaluating which conclusively predicts the risk of KR.*



*Figure 3.Xception is a 71 layer deep CNN where we feed the pre trained data sets for it to classify the data into OA severity and predict KR.*



*Figure 4. Inception v3 is a CNN model that assists in Image analysis and object detection which plays a major role in data set normalization.*

## *Litreture Reiew*

**Prediction of Total Knee Replacement and Diagnosis of Osteoarthritis by Using Deep Learning on Knee Radiographs -September 2020**

–By Kevin Leung, BS • Bofei Zhang, BS • Jimin Tan, BS • Yiqiu Shen, MS • Krzysztof J. Geras, PhD • James S. Babb, PhD • Kyunghyun Cho, PhD • Gregory Chang, MD, MBA • Cem M. Deniz, PhD, Study supported by National Institutes of Health (R01 AR074453).

In this retrospective analysis that used data from the OA Initiative, a DL model on knee radiographs was developed to predict both the likelihood of a patient undergoing TKR within 9 years and Kellgren-Lawrence (KL) grade. Study participants included a case-control matched subcohort between 45 and 79 years. Patients were matched to control patients according to age, sex, ethnicity, and body mass index. The proposed model used a transfer learning approach based on the ResNet34 architecture with sevenfold nested cross-validation. Receiver operating characteristic curve analysis and conditional logistic regression assessed model performance for predicting probability and risk of TKR compared with clinical observations and two binary outcome prediction models on the basis of radiographic readings: KL grade and OA Research Society International.

As a result The prediction model based on DL achieved an area under the receiver operating characteristic curve of 0.87 , outperforming a baseline prediction model by using KL grade with an AUC of 0.74. The risk for TKR increases with probability that a person will undergo TKR from the DL model, KL grade, and OARSI grade. We can conclude that the proposed deep learning model better predicted the risk of total knee replacement in osteoarthritis than did binary outcome models by using standard grading systems.

**A Lightweight CNN and Joint Shape-Joint Space (JS 2 ) Descriptor for Radiological Osteoarthritis Detection - July 2020**

*–By Neslihan Bayramoglu, Miika T. Nieminen, and Simo Saarakkala.*

*Research Unit of Medical Imaging, Physics and Technology, University of Oulu, Finland   Department of Diagnostic Radiology, Oulu University Hospital, Oulu, Finland*.*Medical Research Center, University of Oulu and Oulu University Hospital, Oulu, Finland.*

 In this study, they propose a fully automated novel method, based on combination of joint shape and convolutional neural network (CNN) based bone texture features, to distinguish between the knee radiographs with and without radiographic osteoarthritis. Moreover, they have  reported the first attempt at describing the bone texture using CNN. Knee radiographs from Osteoarthritis Initiative (OAI) and Multicenter Osteoarthritis (MOST) studies were used in the experiments. These models were trained on 8953 knee radiographs from OAI and evaluated on 3445 knee radiographs from MOST. The results demonstrate that fusing the proposed shape and texture parameters achieves the state-of-the art performance in radiographic OA detection yielding area under the ROC curve (AUC) of 95 .21%. They also proposed a simple knee joint shape descriptor JS2 for OA detection from plain radiographs. In addition to this demonstration of a lightweight CNN for extracting bone texture features together with the proposed joint shape joint space descriptor achieves SOTA performance in radiographic OA detection. Also a CNN model, which utilized the most informative ROI to build a network with fewer parameters. Compared to the heavy deep CNN models that use the whole joint image, this tiny CNN model has the ability to recognize OA with high accuracy.

**The primary total knee arthroplasty: a global analysis 2020**

 –By Jiaxiang Gao, Dan Xing, Shengjie Dong and Jianhao Lin

 In this study, we can conclude that Total knee arthroplasty (TKA) is considered as the most common treatment for end-stage knee osteoarthritis (OA). The researchers have evaluated the research at a global level, by selecting the Web of Science (WoS) and Science Citation Index-Expanded (SCIE) for this analysis The search strategy was the exclusion criteria should cover the most common terms related to the arthroplasty on other joints, and other types of knee surgeries .VOS viewer was used for bibliographic coupling, co authorship, cocitation, and co-occurrence analysis. The current status and global trends of TKA research were delineated through bibliometric analysis. The USA has made tremendous contributions to this area, establishing its leading position in global research on TKA. It can be predicted that an increasing number of papers would be published in the next few years, which indicates an underlying vigorous development of TKA research.

**Deep Learning Predicts Total Knee Replacement from Magnetic Resonance Images - April 2020**

-By Aniket.Tolpadi, Jinhee J. Lee, Valentina Pedoia & Sharmila Majumdar

In this work, a pipeline that integrates MR imaging and non-imaging features to attain strong TKR prediction performance, where the data was acquired from a prospective observational study conducted by OAI. The dataset followed 4,796 patients and acquired images including 2D posteroanterior radiographs and 3D Sagittal Double Echo SteadyState (DESS) MRI images over the course of 10 years, which reports accuracy of 78.5 ± 0.134%, sensitivity of 81.8 ± 0.643%, and specificity of 78.4 ± 0.138% has been presented. Comparisons of AUCs showed that the MRI pipeline outperformed the X-ray pipeline for patients without OA and with severe OA, thereby showing the MRI model to have a better combination of sensitivity and specificity in these OA classifications.The results included the OA pretrain utility in TKR prediction-Ray pipeline optimization and performance, MRI pipeline optimization and performance, Comparison of MRI and radiograph pipeline performances and Biomarker identification and analysis.

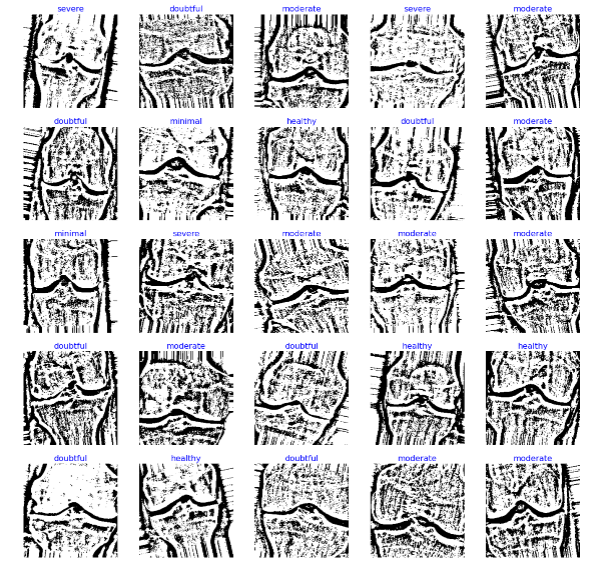
**Rate of cartilage loss at two years predicts subsequent total knee arthroplasty: a prospective study – April 2004**

-By F M Cicuttini, A Forbes, A E Wluka, Department of Epidemiology and Preventive Medicine, Monash University Medical School, Alfred Hospital, Melbourne, Australia G Jones, Menzies Center for Population Health Research, Hobart, Tasmania, Australia

By this study one can conclude that the rate of tibial cartilage loss over two years was an independent predictor of knee replacement at four years. For every 1% increase in the rate of tibial cartilage loss there was a 20% increased risk of undergoing a knee replacement at four years. Those in the highest tertile of tibial cartilage loss had 7.1 higher odds of undergoing a knee replacement than those in the lowest tertile. WOMAC score at baseline, female sex, and tibial bone size (but not age and radiographic score) were also predictors of knee replacement. The data suggest that treatment targeted at reducing the rate of knee cartilage loss in subjects with symptomatic osteoarthritis may delay knee replacement. This has important implications in terms of prevention and therapeutic interventions in osteoarthritis. To perform this as a methodological view  123 subjects with mild to moderate symptomatic radiographic knee osteoarthritis were recruited by either advertising, the Victorian branch of the Arthritis Foundation of Australia, treating general practitioners, orthopedic surgeons, or rheumatologists; 113 completed the study.

Magnetic resonance imaging was carried out at baseline and at 2 years on the symptomatic knee. Rate of change in tibial cartilage volume was calculated. Subjects were then followed up at year 4 to determine whether they had undergone a knee replacement. the first time that the rate of structural change at the knee—articular cartilage loss—is an independent risk factor for subsequent replacement of that knee. This was independent of age, sex, baseline level of pain, and the radiological severity of osteoarthritis.

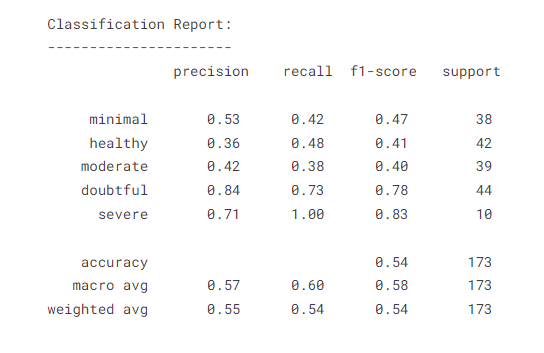
# Results

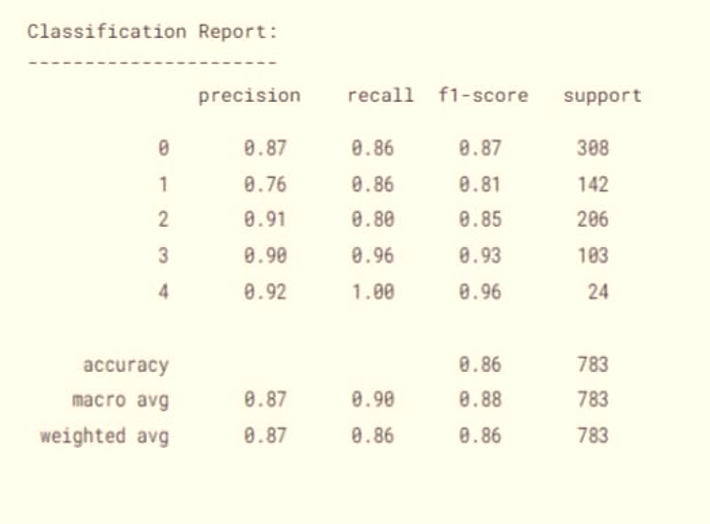


*Figure 5. This image shows the knee severity with stages as minimal, healty, moderate, doubtful and severe.*

An X-ray-based model was fine-tuned; the resulting predictions were fed into an LR ensemble, where averaging predictions of the best 2  models in each OA category optimized validation performance. Performance of the resulting CNN architecture on test data is reported.  KL scale is also used to understand the condition of the knee . The severity of the knee is further used to predict the risk of KR with the accuracy of 86 for the Xception model and 56.6 for the Inception model. With all the medical aids the clinical staff and now use the prediction to perform necessary treatments.

Both Fig 6 and Fig 7 displays the classification report that shows the precision,recall,f1-score and support for the classification stages minimal , healthy, moderate, doubtful and severe for both Xception and Inception models respective with accuracy , macro average and weighted average.

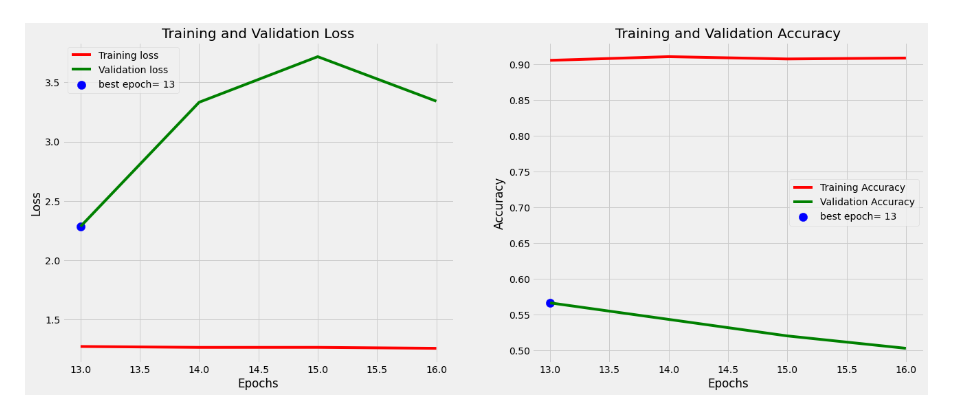
 *Figure 6. Classification Report with Xception Model*



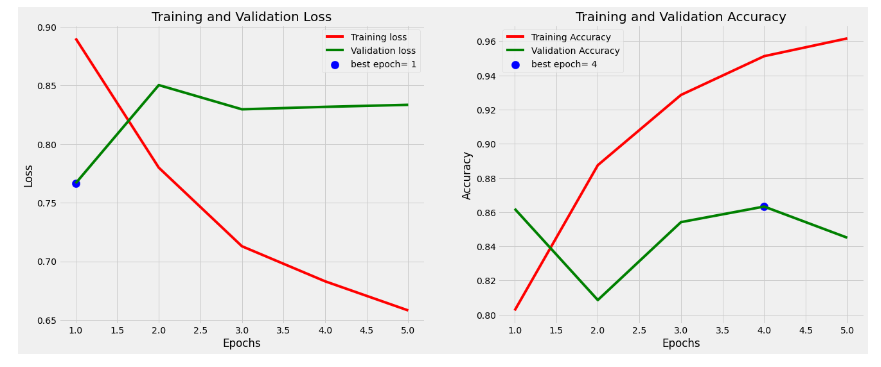
*Figure 7. Classification Report with Inception Model*

## *Comparision of Xception and Inception Model performance using ROC curve.*

The overall performance of both models have been compared here with their ROC curve in the test set traing and validation ,where we can see best epoch point for loss and Accuracy. Testing and validation loss and accuracy of testing data set of Xception model has 13 for each. Whereas for Inception model it is 1 and 4 respectively.



*Figure 8. Shows the ROC curve for Testing and validation loss and accuracy for Test data set of Xception model.*

**

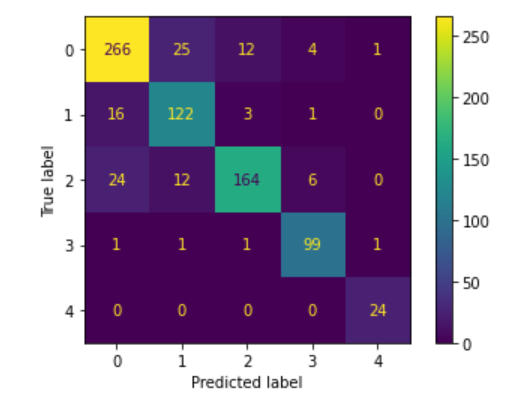
*Figure 9. Shows the ROC curve for Testing and validation loss and accuracy for Test data set of Inception model.*

## *Comparision of Xception and Inception Model performance using Confusion Matrix.*

Confusion matrix is simple tabular description of performance of the classification model. Fig 10 shows the confusion matrix of Xception model where as the Fig 11 shows the Inception model. Where the matrix is drawn on Actual/True vs. Predicted values of the severity classification of the patients. Where we can see that the performance of Inception plays well compared to Xception.

## 

*Figure 10.Confusion matrix of Xception model.*



*Figure 11. Confusion matrix of Inception model.*

## *Conclusion*

In this work, we present a CNN architecture that integrates X-ray images featuring to attain strong KR prediction performance, reporting accuracy of 86.6 ± 0.134%, sensitivity of 81.8 ± 0.643%, and specificity of 86.4 ± 0.138% (intervals calculated with standard error of measurement (s.e.m.), p<0.05). Comparison of ROC curves for Xception and Inception models and Confusion Matrix is done showing the accuracy for Xception as 86.6 ± 0.138% and 56.6 ± 0.243% for Inception for KR prediction model. Thereby showcasing the better model for predicting,the utility of X-ray in predicting through DL is certainly worth further investigation.

## *Discussion*

We showed that in 4,796 the prediction minimum 47% of people has risk of KR due to their age, BMI and lifestyle characterization. The closest analog to our work was conducted by Wang, T., who trained independent residual networks to predict TKR from both DESS and Turbo Spin Echo (TSE) MRI images, integrating both predictions with non-imaging variables in an LR model to yield a final TKR prediction. This yielded a model with AUC of 0.86 ± 0.01 (p<0.05). This performance marks progress towards a model that identifies patients at risk for KR such that nonsurgical treatment strategies can be implemented to delay KR.

##### Acknowledgment

This study was supported by ….

##### References

1. Kremers, H. M. et al. Prevalence of total hip and knee replacement in the United States. J. Bone Joint Surg. Am. 97, 1386–1397 (2015).
2. Murphy, L. B., Cisternas, M. G., Pasta, D. J., Helmick, C. G. & Yelin, E. H. Medical expenditures and earnings losses among US adults with arthritis in 2013. Arthritis Care Res. (Hoboken) 70, 869–876 (2018).
3. Lewis, J. R., Dhaliwal, S. S., Zhu, K. & Prince, R. L. A predictive model for knee joint replacement in older women. PLoS One 8, e83665, https://doi.org/10.1371/journal.pone.0083665 (2013).
4. Huang, G., Lui, Z., van der Maaten, L. & Weinberger, K. Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2261-2269 (2017).
5. Will C, Seo GS, Gale D, Totterman S, Gale ME, Felson DT: Cruciate ligament integrity in osteoarthritis of the knee. Arthritis Rheum 2005, 52:794-799.
6. The primary total knee arthroplasty: a global analysis Jiaxiang Gao, Dan Xing,, Shengjie Dong3 and Jianhao Lin.Gao . Journal of Orthopedic Surgery and Research (2020).
7. Experimental Validation of a 3D Dynamic Finite-Element Model of a Total Knee Replacement Joel L. Lanovaz1 and Randy E. Ellis(2005).
8. Stulberg, S.D., Loan, P., Sarin, V.: Computer-assisted navigation in total knee replacement: Results of an initial experience in thirty-five patients. Journal of Bone and Joint Surgery 84 (2002) 90–98.
9. Halloran, J.P., Petrella, A.J., Rullkoetter, P.J.: Explicit finite element modeling of total knee replacement mechanics. Journal of Biomechanics 38 (2005) 323–331.
10. Taylor, M., Barrett, D.S.: Explicit finite element simulation of eccentric loading in total knee replacement. Clinical Orthopaedics and Related Research 414 (2003) 162–171 7. Halloran, J.P., Petrella, A.J., Rullkoetter, P.J.: Explicit finite element modeling of total knee replacement mechanics. Journal of Biomechanics 38 (2005) 323–331.