

3D Convolutional Networks for prediction of Total Knee Replacement using Structural MRI

Tianyu Wang¹, Kevin Leung^{2,3}, Kyunghyun Cho^{1,2}, Gregory Chang^{4,5} and Cem M. Deniz^{4,6,7}

¹Center for Data Science, ²Courant Institute of Mathematical Sciences, ³Leonard N. Stern School of Business, New York University. ⁴Department of Radiology, ⁵Center for Musculoskeletal Care, ⁶Center for Advanced Imaging Innovation and Research (CAI2R) and Bernard and Irene Schwartz Center for Biomedical Imaging, New York University Langone Medical Center ⁷The Sackler Institute of Graduate Biomedical Sciences, New York University School of Medicine, New York, NY, United States.

Target Audience: Researchers who are interested in osteoarthritis and modelling with deep learning.

Purpose

Osteoarthritis (OA) is the most common form of arthritis and the major cause of physical disability in the elder. The discovery of imaging-based biomarkers could lead to identification of new treatment targets and mechanisms for shorter, more efficient clinical trials of possible disease-modifying agents. In recent years, magnetic resonance imaging (MRI) has been increasingly used in OA, since it can visualize all tissues in the knee joint involved in OA, such as cartilage,^{1,2} menisci,³ bone,² and soft tissue. Current image analysis methods focus mainly on locations/features⁴ previously defined to characterize knee joint health separately, while neglecting the potential effects of a constellation of locations/features yet to be discovered. In this study, we use deep learning methods to identify OA-relevant imaging biomarkers directly from knee MR images. Our aim is to predict the outcome of OA as a subject's probability of total knee replacement (TKR) within nine years directly from structural MR images.

Dataset and Methods

Intermediate-weighted fat-saturated Turbo Spin Echo images from Osteoarthritis Initiative⁵ are used for this study. This sequence provides information about cartilage loss, ligament integrity, meniscal tears and subchondral bone marrow. Parameters of sagittal 2D sequence are: TE / TR = 30 / 3200 ms, FOV = 16 cm, slice thickness = 3 mm, in-plane spatial resolution = 0.357×0.357 mm², matrix size = 448×448×37, bandwidth = 248 Hz/ pixel. 718 case-control pairs (age: 63±12 years, BMI: 29±6.6 kg/m², 274/444 male/female split) were selected from OAI dataset by propensity score matching⁶ on individuals based on the baseline variables: age, BMI, gender and race. Cases were defined as individuals who received a medically confirmed TKR after baseline; individuals who received a partial knee replacement or have already received a TKR at baseline were not included in this study. We defined controls as individuals who did not receive a TKR in either knee on the 108-month visit.

The outline of our modelling approach is shown in Figure 1. We developed a 3D CNN model with a binary target variable (1 for cases, 0 for controls) to predict TKR using structural MRI. Our model contained 5 convolution blocks including convolution layers with batch normalization and max-pooling, followed by three fully-connected layers. Details of the network architecture is shown in Figure 2. Using the demographic and clinical data, we developed a reference logistic regression model (LR), similar to Hochberg et al.,⁷ with baseline age, gender, race, BMI, KOOS Quality of Life score and WOMAC pain score as features. Moreover, another logistic regression model (LR&CNN) combining the TKR probability estimation from 3D CNN model and clinical and demographic information - as in LR model- is compared with the LR model and 3D CNN model. In order to better understand the CNN's learning behavior, we present occlusion maps to get the discriminative image regions used by a CNN to identify cases class in the image.

Result

Receiver operating curve (ROC) analysis of three models are shown in Figure 3. The proposed 3D CNN model achieved the area under the ROC curve (AUC) of 0.84±0.02. The 3D CNN model showed significantly higher AUC than the LR model (AUC: 0.77±0.02, p<0.001). Moreover, the LR&CNN model (AUC: 0.86±0.02) has significantly higher AUC than the LR model (p<0.001).

Figure 4 shows a visualization of the regions in which the trained CNN model uses to make a decision with high impact. The heat map overlaid on an exemplary MR images were generated by systematically occluding the portions of input image and monitoring the output of a classifier.⁹ The example clearly shows the trained CNN model pays an attention to the locations where OA has immediate impact on such as cartilage as the probability of the correct class drops significantly. More interestingly, as indicated by the green arrow in the second row, trained CNN also uses the information from the front of the knee cap where prepatellar bursitis was diagnosed for this patient. For this patient, we found that both cartilage and bursa were playing a role in CNNs decision for TKR probability.

Discussion and Conclusion

We present an automated method to identify OA-relevant imaging biomarkers for the prediction of TKR probability using 3D CNN and structural MR images. Our preliminary results suggest that the use of 3D CNNs provides a significantly better estimation of TKR within 9 years compared to a model using demographic and clinical information. This was achieved by extracting an ensemble of features directly from MR images without specifying the locations/features from the knee. In the future, we plan to incorporate other MR imaging sequences from OAI's dataset into our CNN-based modelling approach. Meanwhile, we would like to see how the 3D CNN performs against WOMBS score¹⁰, which is a semi-quantitative metric that radiologists define the degree of joint damage from MR images.

References:

[1] Carballido-Gamio J et al. Magn Reson Med. 2011;65(4):1184-1194. doi:10.1002/mrm.22693 [2] Kazakia GJ et al. Osteoarthritis Cartilage. 2013;21(1):94-101. doi:10.1016/j.joca.2012.09.008, [3] Wang L and Regatte RR. J Magn Reson Imaging. 2014;00. doi:10.1002/jmri.24677 [4] Neogi T et al. Arthritis Rheum. 2013;65(8):2048-2058. doi:10.1002/art.37987 [5] Peterfy CG et al. Osteoarthritis Cartil. 2008;16(12):1433-1441. doi:10.1016/j.joca.2008.06.016 [6] Rosenbaum PR and Rubin DB. Biometrika (1983) 70:4155 [7] Hochberg, M.C. et al. Osteoarthritis and Cartilage (2013) Volume 21, S11. [8] Simonyan K. and Zisserman A. In ICLR, 2015. [9] Zeiler, M. D., & Fergus, R. (2014). In European conference on computer vision (pp. 818-833). Springer, Cham. [10] Peterfy, C.G et al. Osteoarthritis and Cartilage, Volume 12, Issue 3, 177 - 190

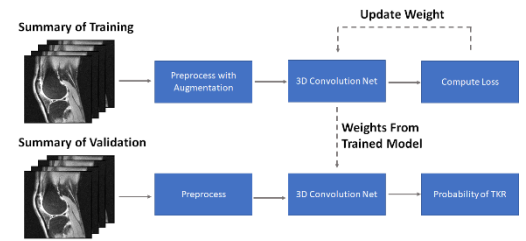


Figure 1: Outline of the proposed approach. During training, we random crop image to size 384×384×32 and random flip the image horizontally in every batch. During validation, central crop was used to obtain the images with the same size as training. 3D CNN was trained using the binary cross-entropy as a loss function and Adam as an optimizer. We performed four-fold cross validation.

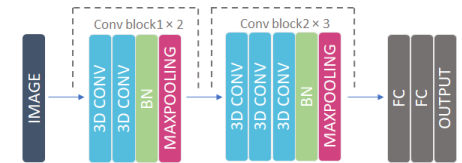


Figure 2: Model's Architecture: Developed 3D CNN architecture. This architecture was inspired by the VGG⁸ net that is extensively used in 2D image recognition.

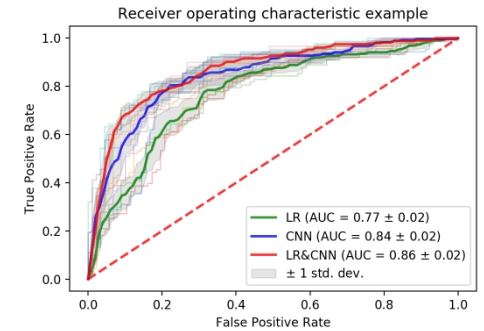


Figure 3: The ROC curves of the proposed 3D CNN model, Hochberg's LR model and the LR model combining demographic and clinical data with CNN result (LR&CNN). ROC curves indicate the improved predictive power of TKR using 3D CNN with structural MRIs.

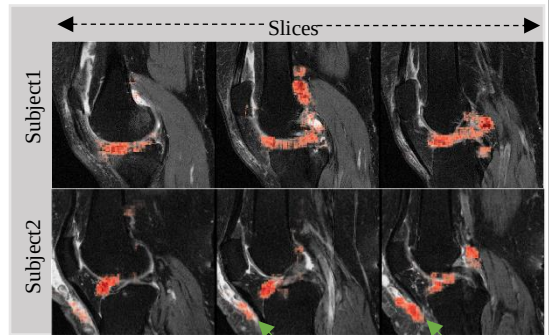


Figure 4: The occlusion heatmap overlaid on MR images from two subjects. Multiple slices from each patient are shown in each row. Heatmaps are generated by one minus the prediction probability by 3D CNN model with blocking 32x32x3 regions sequentially with stride 3 in each direction. The occlusion size for each region is 11.4 mm × 11.4 mm × 9 mm.