Title: Efficacy of 2D and 3D Convolutional Neural Networks in Severity Staging of Anterior Cruciate Ligament Injuries

Authors: Nikan K. Namiri, BS,[1] Io Flament, MS,[1] Bruno Astuto, PhD,[1] Rutwik Shah, MD,[1] Radhika Tibrewala, MS,[1] Francesco Caliva, PhD,[1] Thomas M. Link, MD, PhD,[1] Valentina Pedoia, PhD,[1] Sharmila Majumdar, PhD[1]

Affiliations: [1]Musculoskeletal and Quantitative Imaging Research Group, Department of Radiology and Biomedical Imaging

Introduction: Severity staging of osteoarthritis (OA) in the anterior cruciate ligament (ACL) is dependent on clinician expertise and can be subject to inter-rater variability. Deep learning may serve as a means of standardizing ACL grading and reducing this variability. The purpose of this study is to evaluate the diagnostic utility of two convolutional neural network (CNNs) for severity staging of intact, partially torn, fully torn, and reconstructed ACLs.

Methods: This retrospective analysis was conducted on 1252 knee MR images collected between 2011 and 2014 from 202 subjects, with mean age of 46.50±13.55 (SD) years, mean body mass index 24.58±3.60 (SD) kg/m^2, and 46% female subject proportion. Images were acquired with a 3.0T MR scanner using 3D fast spin echo CUBE-sequences. The ACL was bounded in each volume and CNN architectures with either three-dimensional (3D) or two-dimensional (2D) convolutional kernels classifying the ACL grade (**Figure 1**). Performance metrics included two sample t tests to compare the sensitivity, specificity, weighted Cohen's kappa, and overall accuracy of the two CNNs.

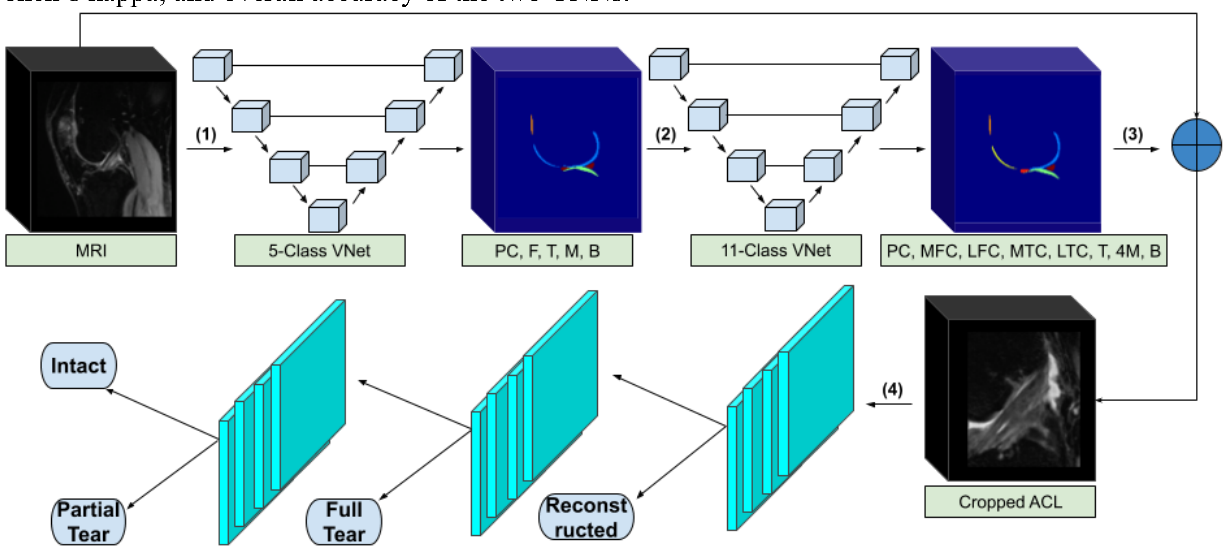

Figure 1: The segmentation and classification pipeline begins with (1) the input of a full MR volume into a 3D VNet which segments the knee into patellar cartilage (PC), femur (F), tibia (T), meniscus (M), and background (B). (2) The 5-class segmentation is then input to a second VNet which further categorizes the knee into 11 compartments including patellar cartilage (PC), medial and lateral femoral and tibial condyles (MFC, LFC, MTC, LTC), and four meniscal horns (4M). (3) The 11-class segmentation is used to determine the ACL boundaries of the original input MRI. (4) The cropped ACL volume is input to three hierarchical CNNs (either 2D or 3D), which each detect reconstructed, fully torn, partially torn, and intact ACLs.

Results: The overall accuracy (92%) and weighted Cohen's kappa (.88) reported for ACL injury classification were significantly greater using the 2D CNN than the 3D CNN. However, the 3D CNN outperformed the 2D CNN in the partial tear classification task, achieving a sensitivity of 74%, compared to 26% for the 2D CNN. The 2D CNN and 3D CNN performed similarly in assessing intact ACLs (2D CNN: 93% sensitivity and 90% specificity, 3D CNN: 89% sensitivity and 88% specificity). Classification of full tears by both networks were also comparable (2D CNN: 83% sensitivity and 94% specificity, 3D CNN: 77% sensitivity and 100% specificity). The 2D CNN classified all reconstructed ACLs correctly.

Conclusion: Deep learning methods applied to ACL lesion classification tasks demonstrate both high sensitivity and high specificity. This technique may potentially be used as a diagnostic tool reducing radiologist workload.

Highlights of Abstract: We developed a deep learning-based pipeline to detect ACL abnormalities and stage lesion severity of fully torn, partially torn, reconstructed, and intact ACLs. Both 2D and 3D convolutional kernels achieved a high overall accuracy (2D: 88%, 3D: 82%) and high Cohen's kappa (2D: .92 and 3D: .84).