Using deep learning to diagnose knee injuries on magnetic resonance images: current potential and limitations

Nicolai Sandau BSc MD, Stig Brorson MD PhD DMSc

Centre for Evidence-Based Orthopedics, Department of Orthopedic Surgery, Zealand University Hospital, Køge, Denmark

# Introduction

Magnetic resonance imaging (MRI) is the primary non-invasive intervention for diagnosing soft tissue knee injuries. Correct interpretation of knee MRI can be time-consuming. Deep learning assisted diagnosis in knee MRI has shown potential for improving diagnostic speed and accuracy, but the results have been inconsistent. Recent advances in deep learning methods have resulted in improvements in both performance and interpretation for other medical image diagnosis tasks. We aim to study these advances and develop a deep learning model capable of diagnosing anterior cruciate ligament (ACL) and meniscal tears in knee MRI.

# Methods

We used the MRNet dataset consisting of 1250 knee MRI scans from Stanford University Medical Center. The dataset is split into a training set (1130 cases) and a validation set (120 cases). Each case has been labeled with the presence or absence of an ACL tear and/or meniscal tear as diagnosed by a radiologist. We developed a deep learning model based on the EfficientNet architecture and trained it on the training set. The trained model then diagnosed the scans in the validation set. Using the diagnostic labels as ground-truth we calculated the sensitivity and specificity. We calculated the area under the receiver operating characteristic curve (ROC) as an overall measure of performance. Lastly, we used Gradient-weighted Class Activation Mapping to visualize which regions of each scan the model used for diagnosis.

# Results

The model achieved ROC (95% CI) values of 0.99 (0.97 - 1) for ACL tears and 0.88 (0.82 - 0.94) for meniscal tears. The average ROC was 0.93, the currently highest reported for the MRNet dataset. The sensitivity (95% CI) was 0.87 (0.76 - 0.94) for ACL tears and 0.83 (0.7 - 0.91) for meniscal tears. The specificity (95 CI) was 0.94 (0.85 - 0.98) for ACL tears and 0.76 for meniscal tears (0.65 - 0.85). The sensitivity and specificity is comparable to previously reported values achieved by trained radiologists on the MRNet dataset.

# Conclusions

Our findings indicate that deep learning has the potential to aid radiologists and orthopedic surgeons when interpreting ACL and meniscal tears in knee MRI, but generalization to other institutions with different patient populations and imaging practices needs to be improved.