Abstract #02859 ISMRM ABSTRACT

Influence of Site Effects on Radiomics-Based Knee Injury Diagnosis

Primary: Musculoskeletal - Tendon/Ligament 〉 **Secondary:** Analysis Methods - Radiomics 〉 **Presentation:** Oral, PowerPitch Oral, Digital Poster 〉 **Keywords:** RADIOMICS MAGNETIC FIELD STRENGTH GAUSSIAN MIXTURE MODEL SITE EFFECTS KNEE INJURY

Jiqing Huang[1], **Yi Chen**[2,3], **Antoine Jacquemin**[1], **Evgenios N Kornaropoulos**[1], **Mohamed Ali Bahri**[1], **Christophe Phillips**[1]

[1]CRC-Human Imaging, GIGA-Institute, University of Liège, Liège, Belgium
[2]Key Laboratory of Advanced Medical Imaging and Intelligent Computing of Guizhou Province, Guizhou University, Guizhou, China
[3]D-Lab, Maastricht University, Maastricht, Netherlands

Presenting Author: *Jiqing Huang (jiqing.huang@uliege.be)*

Impact

Magnetic field strength (MFS) substantially affects feature distributions and may introduce bias into machine learning-based disease classification. These findings highlight the need to balance field strength variations in multi-site MRI datasets to enhance model generalization and clinical applicability.

Synopsis

Motivation: MRNet is a widely used database for knee injury diagnosis. While several classifiers achieve high accuracy on internal datasets, their performance often declines on external data. Researchers primarily attributed this degradation to site effects, though the role of MFS variations has likely been underestimated.

Goals: To disentangle the effects of MFS and other site-related factors in knee injury detection.

Approach: MFS was inferred using a Gaussian mixture model. Radiomic features were evaluated through one-way ANOVA to quantify site effects.

Results: Preliminary results suggest that MFS substantially influences radiomic feature distributions and introduces bias, particularly in imbalanced datasets.

INTRODUCTION

MRI remains the imaging modality of choice for assessing knee injuries. Recent studies have reported that radiomics- and deep learning–based methods can achieve high diagnostic accuracy in classifying specific knee pathologies, particularly anterior cruciate ligament (ACL) and meniscal tears. The MRNet 1 challenge dataset, including over 1,000 subjects, has been extensively used and continues to yield improving internal accuracy in the classification tasks. However, when these well-tuned models are applied to external validation cohorts, performance often drops markedly, limiting clinical utility1-3. This degradation is commonly attributed to site-related factors (e.g., scanner manufacturer, acquisition protocol,)4. However, these explanations may be incomplete: heterogeneity in magnetic field strength (MFS) may also play a significant role5. In this study, we analyse how MFS and non-MFS effects drive systematic variation in radiomic feature representations.

METHODS

Among the publicly available datasets, the MRNet and KneeMRI cohorts[6] have become benchmarks for knee injury diagnosis, as listed in Table 1. We analyzed the effect of MFS on MRNet tasks. Since MRNet does not provide MFS metadata, we inferred MFS using a Gaussian Mixture Model (GMM) with a slice-count prior, retaining subjects with >95% clustering confidence. The impact of the inferred 1.5 T and 3 T groups was then evaluated across the abnormality, meniscal tear, and ACL tear tasks. To isolate manufacturer effects from MFS, we compared 1.5 T scans from MRNet and KneeMRI on the ACL task. To ensure comparable between the KneeMRI and MRNet cohorts, we apply the same histogram equalization procedure and combine partial and complete ACL tear into ACL-injured class in the KneeMRI cohort. Cartilage was segmented using nnU-Net[7], after which a peri-cartilaginous rectangular-voxel region of interest (ROI) was defined for feature extraction. Radiomic features were then extracted within the ROI using PyRadiomics. For injury-specific analyses, we used axial PD-weighted images for abnormalities and meniscal tears (as recommended by the original study[1]) and sagittal T1-weighted images for ACL tears to assess site and MFS effects. To mitigate redundancy, principal component analysis (PCA) was applied to the extracted radiomic features to retain 20 principal components; mutual information with the disease label was then used to select the 10 most informative components. Finally, ANOVA assessed the influence of MFS and site effects on feature separability.

RESULTS

Using the slice count information listed in Table 1, the GMM successfully split MRNet participants into 1.5 T and 3 T cohorts, as shown in Figure 1a. Figure 1b shows class prevalence stratified by MFS. Following histogram equalization for the KneeMRI cohort and nnU-Net-based segmentation for both, a rectangular-voxel ROI (three slices per participant) was defined. Representative examples are shown in Figure 2. ANOVA results are summarized in Table 2. For MFS, feature-level effects are reported across all three lesion categories (abnormality, meniscal tear, and ACL), while site effects were analyzed only for the ACL group.

DISCUSSION

The GMM separated MRNet data into 1.5 T and 3 T cohorts. Disease prevalence differed into abnormality (65.52% vs. 89.55%), ACL (24.60% vs. 18.87%), and meniscus (24.83% vs. 44.85%). These findings indicate clear distribution shifts between the 1.5 T and 3 T subsets. In contrast, ACL prevalence in MRNet and KneeMRI 1.5 T remained roughly balanced (24.60 % vs 24.75 %). In the ACL analysis, no clear association was observed between radiomic features and non-MFS effects, indicating that when factors such as manufacturer or acquisition protocol are balanced, radiomic features exhibit minimal domain-related bias. Across the three diagnostic tasks, the influence of MFS increased progressively from ACL to abnormality to meniscus detection. For abnormality detection, two of ten disease-related features were correlated with MFS, suggesting mild but detectable field-strength sensitivity. In contrast, the meniscus analysis revealed a pronounced MFS effect, with five of ten features associated with MFS. Among these, the most responsive feature reached an F-value of 34.95 ($p < 0.001$), indicating substantial differences between the 1.5 T and 3 T cohorts. These results highlight the importance of explicitly distinguishing MFS and non-MFS effects, as MFS imbalance can systematically bias radiomic feature representations and should be modeled or corrected during preprocessing and model development. This study has certain limitations. Future work will use more precise ROI to better characterize the influence of different factors. We also plan to explore alternative data splits and conduct a more continuous assessment of site effects to gain a deeper understanding of their impact on radiomic feature variability.

CONCLUSION

This study demonstrates that while radiomics features remain largely domain-agnostic under balanced condition. Moreover, data imbalance can introduce systematic biases related to site effect, challenging the robustness and reproducibility for clinical translation and implementation.

References

1. Bien, Nicholas, et al. "Deep-learning-assisted diagnosis for knee magnetic resonance imaging: development and retrospective validation of MRNet." PLoS medicine 15.11 (2018): e1002699. doi: 10.1371/journal.pmed.1002699

2. Wong S, Steinbach L, Zhao J, Stehling C, Ma CB, Link TM. Comparative study of imaging at 3.0 T versus 1.5 T of the knee. Skeletal Radiol. 2009 Aug;38(8):761-9. doi: 10.1007/s00256-009-0683-0.

3. Tran, Alexia, et al. "Deep learning to detect anterior cruciate ligament tear on knee MRI: multi-continental external validation." European Radiology 32.12 (2022): 8394-8403. doi: 10.1007/s00330-022-08923-z

4. Nikolaou, Vassilios S., et al. "MRI efficacy in diagnosing internal lesions of the knee: a retrospective analysis." Journal of trauma management & outcomes 2.1 (2008): 4. doi: 10.1186/1752-2897-2-4

5. Hajianfar G, Hosseini SA, Bagherieh S, Oveisi M, Shiri I, Zaidi H. Impact of harmonization on the reproducibility of MRI radiomic features when using different scanners, acquisition parameters, and image pre-processing techniques: a phantom study. Med Biol Eng Comput. 2024 Aug;62(8):2319-2332. doi: 10.1007/s11517-024-03071-6.

6. Štajduhar I, Mamula M, Miletić D, Ünal G. Semi-automated detection of anterior cruciate ligament injury from MRI. Comput Methods Programs Biomed. 2017;140:151–64. doi: 10.1016/j.cmpb.2016.12.006

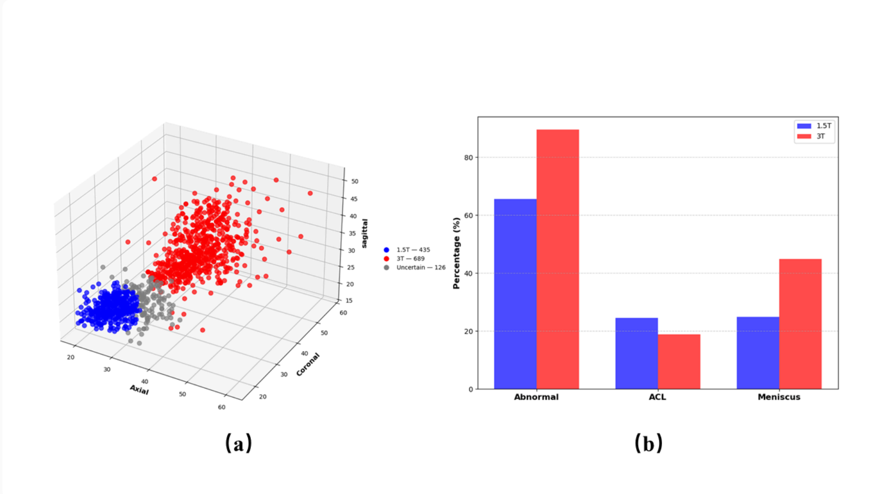
7. Isensee, Fabian, et al. "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation." Nature methods 18.2 (2021): 203-211.doi: 10.1038/s41592-020-01008-z

Figures and Tables

Dataset	KneeMRI	MRNet
Scanner Field Strength	Siemens Avanto 1.5 T	GE 1.5 T and 3.0 T systems (3 T: 56.6%)
Population	917 knee MRI cases: • ~76% (717) ACL non-injured • ~19% (182) ACL partially injured • ~5% (45) ACL completely ruptured	1,370 knee MRI cases: • ~81% (1,104) abnormal exams • ~23% (319) ACL tears • ~37% (508) meniscal tears *[120 tuning cases inaccessible]*
Sequence	Sagittal: PD-FS weighted	Sagittal: PD-FS weighted Coronal, Axial: PD weighted
Number of Slices	21 – 45	1.5 T: Sagittal 24 / Coronal 20 / Axial 30 3 T: Sagittal 42 / Coronal 44 / Axial 44

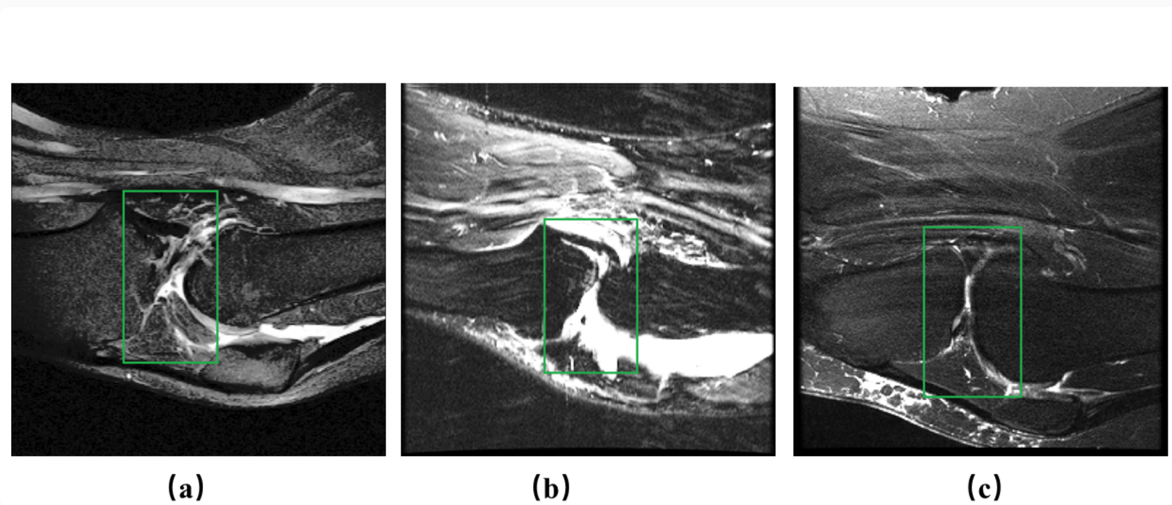
Scan for high-resolution version

Figure 1: *Table 1. Comparison of Knee MRI Datasets: KneeMRI vs. MRNet (ACL: anterior cruciate ligament; PD: Proton Density; FS: Fat-Suppressed)*



Scan for high-resolution version

Figure 2: *Figure.1 Cohort separation and magnetic field strength distribution in MRNet. (a) GMM-based classification of 1.5 T and 3 T cohorts (posterior probability > 0.95). (b) Sample distribution by inferred field strength.*



Scan for high-resolution version

Figure 3: *Figure 2. Representative ACL MRI images: (a) KneeMRI 1.5 T, (b) MRNet 1.5 T, and (c) MRNet 3 T.*

Effect analysis	Disease	Number of significant features	F-value range	p-value range
Non-MFS	ACL tear	0/10	0.5 – 1.8	0.25 – 0.78
MFS	ACL tear	0/10	0.6 – 2.0	0.22 – 0.65
	abnormality	2/10	0.5-3.2	0.04 – 0.68
	Meniscal tear	5/10	5.8 – 34.95	<0.001-0.4

Scan for high-resolution version

Figure 4: *Table 2. Statistical analysis (ANOVA) of site effects on radiomic features.*