Deep learning for automatic quantification of AVN of the femoral head on 3D MRI in patients eligible for joint preserving surgery: A pilot study

DS applications and challenges in Medicine, Natural Sciences, and Engineering

Ruckli Adrian
sitem Center for Translational Medicine and Biomedical Entrepreneurship, Personalised Medicine research group. University of Bern, Bern, Switzerland

Size of necrosis is an important prognostic factor in the management of femoral head necrosis (AVN), usually estimated on radiographs and MRI which is subjective and requires experienced physicians. Ideally, a fast-volumetric assessment of necrosis size would be desirable for a more objective standardized evaluation. Thus, we evaluated a deep-learning method to automatically quantify the necrotic bone in AVN.
IRB-approved retrospective study of 34 patients (mean age 30 years, 14 women) with AVN according to the commonly recommended 2019 ARCO grading: I (negative x-rays): 3 hips; II (no fracture): 5 hips; IIIA (head collapse 2 mm): 12 hips. Patients underwent preoperative 3T hip MRI including 0.8 mm^3 3D T1VIBE on which manual ground truth segmentation of necrosis and the vital bone from the femoral head was performed by an expert reader and then used to train a set of convolutional neural networks (nnU-Net [1]). The raw data had a median image shape and spacing of 104×384×384 voxels and 1×0.44×0.44 mm, respectively. The highest in-plane resolution was oriented axial-oblique parallel to the femoral neck. As a preprocessing step, the images were resampled to the medial spacing and volume cropped around the femoral head center to the shape of 80×160×160 voxels. Volume cropping reduced the background complexity and accelerated the network training time. A 5-fold cross-validation was performed between manual and automatic volumetric analysis of absolute/relative necrosis volume. The mean difference between manual and automatic segmentation was compared with paired t-tests and correlation was assessed with Pearson correlation coefficients. We compared the absolute and relative size of the necrosis between early and advanced stages of AVN (ARCO I/II versus IIIA/B) using Mann-Whitney U tests. A p-value <0.05 determined statistical significance.
The best performing configuration was the ensemble of the 2D and 3D U-net. The mean Dice coefficient for the vital femoral head bone and necrosis was 89±9% and 69±25%, respectively. The individual 2D (89±9%, 67±23%) and 3D (89±10%, 69±26%) networks were performing very similarly on both vital and necrotic bone ($p>0.05$). Mean absolute and relative AVN volume was comparable between manual (8.2±7.4cm^3, 17±15%) and automatic (7.3±6.7cm^3, 15±14%) segmentation (both $p>0.05$) and showed a strong correlation (r_p=0.90 and r_p = 0.92, both $p<0.001$), respectively. Manual and automated segmentation detected a difference (both $p<0.05$) in relative necrosis volume between early and advanced AVN: 8±8% vs 20 ±16% and 7±8% vs 18±14%, respectively.
Applying a deep learning method for volumetric assessment of AVN is feasible and showed very strong agreement and enabled to distinguish early vs advanced disease stages which paves way for evaluation in larger datasets, with the goal to determine its prognostic value.

[1] Isensee F., Jaeger P. F., Kohl S. A. A., Petersen J., Maier-Hein K. H. (2020) “nnU-Net: A Self-Configuring Method for Deep Learning-Based Biomedical Image Segmentation.” Nature Methods: 1–9.