RSS08

Can artificial intelligence compete with radiographers in characterizing radiographs of the upper limb?

C. Chênes[1], D. Locarnini[1, 2], D. Perréard[1, 3], B. Sierra Paulino[1, 2], J. Schmid[1]
[1]Haute école de santé de Genève, HES-SO, Genève, Switzerland, [2]CHUV Centre hospitalier universitaire vaudois, Service de radiodiagnostic, Lausanne, Switzerland, [3]Hirslanden Clinique des Grangettes, Institut de radiologie, Genève, Switzerland

Purpose: Artificial intelligence (AI) is transforming the landscape of radiology, including the curriculum and clinical practice of radiographers. This study focuses on the development of an AI system to characterize radiographs of the upper limb based on deep learning – by investigating its performances and applicability in a teaching environment.

Methods and Materials: 40'561 radiographs from the public MURA dataset were organized into main anatomical parts (shoulder, humerus, elbow, forearm, wrist, and hand). Furthermore, we categorized 8'379 shoulder radiographs into 5 standard projections: axial, Neers and AP internal/external/neutral rotations. A variant of the NASNet architecture was trained to classify the limb parts and shoulder projections. Part of the dataset was kept aside to compare the performances of 20 last year BSc radiographer students against the AI, during a hands-on workshop.

Results: In general, the AI correctly recognized anatomical parts (mean accuracy of 96.5%), and reached a good accuracy in classifying shoulder projections (84.2%) – excelling with axial radiographs (99.5%) while facing major difficulties with external rotations (59.4%).

Using the workshop dataset, AI performance in shoulder classification dropped to 75% while students collectively performed better (87.5%). Actually, AI and students were on par for axial, Neers and neutral rotation radiographs but the AI struggled with internal and external rotations. Interestingly, a class activation map technique showed that AI focused like humans on the humeral head to distinguish shoulder radiographs. The AI was clearly faster than students (~100x speedup).

Conclusion: The developed AI showed very promising classification capabilities, not for diagnosis support, but for automatic post-processing of radiographs (e.g., check/propose image description, support of billing systems). More importantly, by engaging radiographer students to compete against an AI, we raised student interest and awareness on the capabilities and limitations of AI.