Deep Learning-Based Identification Of Traditional Hip, Knee, and Shoulder Arthroplasty and Application to Alternative Arthroplasty Designs

Paul H Yi, MD, Tae Kyung Kim, BA, Jinchi Wei, BSE, Ji Won Shin, BSE, Tae Soo Kim, MSE, Gregory Hager, PhD, Jan Fritz, MD, Haris Sair, MD

Johns Hopkins Radiology Artificial Intelligence Lab (RAIL) & Johns Hopkins Malone Center for Engineering in Healthcare

Background.
The purpose of this study was to develop and test the performance of deep convolutional neural networks (DCNNs) to 1) detect the presence of hip, knee, and shoulder arthroplasty; and 2) differentiation between types of arthroplasty. A secondary goal was to test the ability of "traditional" arthroplasty DCNNs to detect "alternative" arthroplasties without additional training.

Methods.
We collected 3 datasets consisting of 156 to 240 AP pelvis, knee, or shoulder radiographs (XRs) [50% with and 50% without arthroplasty]. Amongst each joint's arthroplasty cases, 50% were "traditional" arthroplasty (total hip [THA], total knee [TKA], or total shoulder [TSA]) and 50% were "alternative" arthroplasty (hip resurfacing [HRA], unicompartmental knee [UKA], or reverse TSA [RTSA]). For each DCNN, images were split into training (70%), validation (10%), and test (20%) datasets. Training & validation datasets were augmented 22x via standard techniques. The ResNet-18 DCNN (pretrained on ImageNet) was trained and validated for 1) presence of an arthroplasty and 2) distinguishing "traditional" from "alternative" arthroplasties. We also tested DCNNs trained for "traditional" arthroplasty on a set of 34 XRs with and 34 without "alternative" arthroplasty. Receiver operating characteristic (ROC) curves with area under the curve (AUC) were generated and statistically compared between DCNNs.

Results.
DCNNs trained for presence of hip, knee, and shoulder arthroplasty all achieved AUC of 1 ($p=1$). DCNNs trained to distinguish "traditional" from "alternative" arthroplasty had AUCs of 0.96 to 1 for all 3 joints ($p=0.3$). DCNNs trained for THA, TKA, and TSA were able to identify HRA, UKA, and RTSA, respectively, all with AUC of 1, despite not being previously exposed to any of the latter prostheses.

Conclusion.
DCNNs trained on small image datasets augmented 22x can accurately identify presence of hip, knee, and shoulder arthroplasty and distinguish between specific designs. We demonstrate DCNNs' ability to identify the presence of prostheses they were not previously exposed to. The proof-of-concept of these DCNNs may set the foundation for DCNNs to identify specific prosthesis models and prosthesis-related complications.