Toward Limitless Implant Recognition: Morphology-Based AI for Known and Unseen Legacy Implants

Michael Murphy[1], Michael J Murphy
[1]Orthopaedic Surgery

INTRODUCTION: Existing artificial intelligence (AI) models for orthopaedic implant identification from radiographs are limited in scope, typically recognizing nine or fewer implant types. Even the most expansive published models only distinguish among 18 implants, often having distinct morphologies, and all require extensive labeled examples for training. These limitations constrain applicability, especially for distinguishing among subtly different, older, or discontinued implant models.

METHODS: A multi-institutional total joint database comprising over 50,000 radiographs was utilized, supplemented with publicly available online images, to train an AI model on 173 unique total hip arthroplasty (THA) and total knee arthroplasty (TKA) implant models. Unlike traditional classification-based approaches, this model generates embedding vectors to capture implant similarity, enabling recognition beyond the training set. To evaluate generalizability, 15 implant models (and all associated radiographs) were excluded from training. Model performance was assessed both on unseen images of known implants and on the excluded implant types.

RESULTS: On the hold-out test set comprising previously unseen images of the 173 trained implants, the model achieved an area under the curve (AUC) of 0.97, indicating strong discriminatory power. For the 15 excluded implant types, the model achieved an AUC of 0.91, demonstrating robust generalization to unseen implants. Confidence scores generated alongside predictions were significantly higher for correct identifications ($p < 0.0001$). The system processed approximately 28 radiographs per second, supporting real-time clinical applicability.

DISCUSSION AND CONCLUSION: This study presents the first AI-based orthopaedic implant identification model designed to recognize implants outside its training set. By using this approach, the model can scale to an unlimited number of implant types without retraining. While performance declined slightly on unseen implants, the high AUC and interpretable confidence scores suggest strong potential for clinical deployment in identifying both current and legacy implants.