Automated Identification of Vertebral Screws Using Artificial Intelligence

Kellen Leo Mulford, Julia Todderud, Christopher Adam Magera, Andrew David Pumford, A. Noelle Larson

INTRODUCTION: Applications of artificial intelligence in healthcare offer efficient and accurate tools for clinical and research settings. Spine surgeries constitute a large sector of healthcare but have a variety of potential procedures and devices. With this study, we aim to develop an automated radiograph classifier for vertebral screws. Characteristics such as screw thread patterns and tulips vary with each screw type and could be used to identify patient instrumentation. This has potential for use in cases of patients with unknown surgical history, need for revision procedures, registry development, and research image processing. We hypothesized the model would perform better on more unique screw types and have lower performance distinguishing screw sizes.

METHODS: Patient spine radiographs were collected from our institutional image registry. 3 human annotators labeled the individual screws for model ground truth. Images were split in a 70:20:10 split for training, validation, and testing and the YoloV8 model was used for model development. 6 screw types were included: Dynesys, Expedium, Legacy, Solera, Synapse, and USS. Results are reported as precision, recall, and MAP metrics illustrating the level of overlap between screw label location by AI and human annotations.

RESULTS:

2837 spine radiographs for 353 patients were used, with 28126 screw objects. The overall performance of the model on the test set was an MAP-50 of 0.96 across 567 images and 5829 screw objects. For individual labels, model performance was highest on Dynesys screws (mAP50=0.992, N=163 instances) and lowest on Legacy screws (mAP50=0.936, N=1020 instances).

Initial images were annotated with screw type and diameter: Expedium 4.5, 5.0, and 5.5, Legacy 5.5 and 6.35, and Solera 5.5 and 6.0, in addition to Synapse, USS, and Dynesys. Model performance was uneven stratified by screw sizes, with mAP-50 metrics ranging 0.587 (Solera 6.0) to 0.99 (Expedium 5.5) and 0.875 overall.

DISCUSSION AND CONCLUSION: We trained an AI model to a high degree of accuracy to classify a variety of vertebral screws in patient spine radiographs. The model excelled at identifying screw types, with more variability in classifying sub-types of screw diameters. This is likely associated with varying image sizes and screw lengths contributing to size confusion. With this deep learning model, we offer an efficient and effective resource for understanding and processing patient spine radiographs.