Automated Landmark Detection in Functional Lateral Radiographs Using Deep Learning

Joseph Kavolus, Sanjeev Gupta[1], Gerard Hayden Smith[2], Edgar A Wakelin, Christopher Plaskos[2], Jim Pierrepont[2]
[1]RPAH Medical Centre, [2]Corin

INTRODUCTION:

In total hip arthroplasty, precise assessment of patient-specific spinopelvic mobility is required for optimal implant alignment and postoperative outcomes. The evaluation process necessitates the detection of bony landmarks in lateral functional radiographs. Regrettably, current manual landmarking methods are inefficient, subjective, and ill-equipped for large-scale application, which can restrict the throughput of clinical assessment especially in high-volume healthcare settings. This study proposes using deep learning models to automate landmark detection and derivation of spinopelvic measurements (SPM). Given the complexity of the Spinopelvic interactions and multiple datapoints to create a patient specific plan for hip replacement, demonstration of Artificial Intelligence or Deep Learning models to assist surgeons in surgical planning is essential to allow for broad implementation and consideration of this complex concept into preoperative planning for hip replacement patients.

METHODS:

Utilizing an international multicenter imaging database comprising 25,628 landmarked preoperative and postoperative lateral functional radiographs (HREC: Bellberry: 2020-08-764-A-2), we developed an array of deep learning models for anatomic landmark detection. An object detection model was applied to isolate the L1 vertebra initially, after which the L1 and sacral endplate, pubic symphysis, and anterior superior iliac spine landmarks were identified using a second keypoint detection model, Figure 1. Three functional positions were analyzed: 1) standing, 2) contralateral step-up, and 3) flexed seated. Landmarks were manually captured and independently verified by qualified engineers during preoperative planning with additional assistance of 3D computed tomography derived landmarks. From both the ground truth and predicted landmark coordinates, we derived the pelvic tilt (PT), sacral slope (SS), lumbar lordotic angle (LLA), and the L1 endplate angle (relative to the horizontal). In a pilot study, we examined interobserver variability with nine qualified engineers annotating three functional images, blinded to additional 3D data. The SPM values were tested on 2,127 images and the L1 object detection algorithm was tested using 555 instances of vertebrae across 40 images.

RESULTS:

The object detection algorithm used to isolate the L1 vertebra achieved a mean average precision at an intersection of union of 0.5 (mAP@50) of 99.3%, precision of 97.9%, and recall of 98.6%. The model produced a mean absolute error (MAE) ± standard deviation for PT, SS, and LLA of 1.6°±2.1°, 3.2°±2.6°, 4.2°±3.2°, respectively. Furthermore, the L1 endplate angle error relative to the horizontal was determined to be 2.5°±2.4°. The interobserver 95% confidence interval (CI) for engineer measured PT, SS, and LLA (1.9°, 1.9°, 3.1°, respectively) was comparable to the MAE values generated by the model. The mean model prediction time was 1.3 seconds per image.

DISCUSSION AND CONCLUSION:

The model MAE reported comparable performance to the gold standard when blinded to additional 3D information. LLA prediction produced the lowest SPM accuracy, potentially due to error propagation from the SS and L1 landmarks. Our models demonstrate excellent performance when compared against the current gold standard manual annotation process. This indicates the potential for accurate and efficient spinopelvic mobility assessments in healthcare environments, especially when broad-scale deployment is required. Given the complexity of inputs and need for quality control regarding data point inputs from images a deep learning or artificial intelligence aid for surgeons will be essential for broader implementation into clinical practice. Future work could include exploration into other measurements of interest, such as pelvic incidence and pelvic femoral angle, which could provide further insights and enhance the model's utility in orthopaedic assessments.