Automated Femur Geometry Analysis Using Machine Learning in the Context of Atypical Femoral Fracture

Alyssa Rosen[1], Ruoyu Xu[1], Frank Pham[1], Stephen Falkner[1], Jonathan Gao[1], Jayden Yu[1], Maggie Jiang[1], Jason Li[1], Tania Lokesh[1], Olivia Rosen[1], Makayla Clark[1], Jason Xu[1], Iris Yang[1], Rashad Madi[1], Rebecca Willcocks[2], Ani Ural[3], Bruce Kneeland[1], Neil Sheth[1], Elena Taratuta[1], Rui Feng[1], Chamith S Rajapakse[1]

[1]University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, [2]University of Florida, Gainesville, FL, [3]Villanova University, Villanova, PA

arosen22@upenn.edu

Disclosures: Alyssa Rosen (N), Ruoyu Xu (N), Frank Pham (N), Stephen Falkner (N), Jonathan Gao (N), Jayden Yu (N), Maggie Jiang (N), Jason Li (N), Tania Lokesh (N), Olivia Rosen (N), Makayla Clark (N), Jason Xu (N), Iris Yang (N), Rashad Madi (N), Rebecca Willcocks (N), Ani Ural (N), Bruce Kneeland (N), Neil Sheth (N), Elena Taratuta (N), Rui Feng (N), Chamith S Rajapakse (N)

INTRODUCTION: Osteoporosis is a disease characterized by the deterioration of bone microarchitecture, decreased bone mass, and increased bone fragility. Bisphosphates are the standard pharmacological treatment for osteoporosis, but long-term consumption has been associated with adverse effects, including atypical femoral fracture (AFF). Femur geometry is also recognized as an important risk factor for AFF. For example, long-term bisphosphonate users with acute femoral neck-shaft angles were associated with AFF [1]. While X-ray analysis of proximal femoral geometry is valuable in the study of AFF prediction and diagnosis, manual landmark detection and measurement are labor- and time-intensive processes. To address this, we propose a convolutional neural network for fully automated landmarking and measurement of proximal femur geometry from X-ray.

METHODS: To train and evaluate the model, 382 anteroposterior pelvic radiographs from the publicly available Osteoarthritis Initiative (OAI) Dataset were used. For each image, 20 measurements were taken manually, including femoral medial and lateral cortical thicknesses, shaft width, head diameter, horizontal and vertical offsets, neck width, hip axis length, neck axis length, and neck-shaft angle for both left and right sides. The images were split into three datasets for training (n = 314), validation (n = 39), and testing (n = 40). Since the model was designed to learn the left femur measurements, each image was included in its original form in addition to a horizontally flipped version, effectively doubling the size of data. A convolutional neural network (CNN) with four convolutional layers and two fully connected layers was used. Each convolutional layer used 3x3 kernels and doubled the number of filters of the previous layer (except for the first layer, which used 32 filters), followed by a ReLU activation function and a 2x2 max pooling. The final feature maps were flattened and passed through a fully connected layer of 4096 units (with ReLU) and followed by a second fully connected layer with 11 units. Instead of directly predicting the 10 left femur measurements, the model produced 11 outputs corresponding to the coordinates of predefined landmark points in a relative coordinate system. In this system, the femoral head center (point A) served as the origin, and the femoral shaft axis defined the y-axis (Fig. 1). The ground-truth measurements were used to reconstruct the true locations of these points in the relative coordinate space, and the loss was computed as the sum of the L2 norms between each predicted point and its corresponding true point.

RESULTS SECTION: After training for 10 epochs, with the stopping point determined by early stopping based on validation loss, the model predicted each of the 10 left femur measurements with the following percent errors relative to the ground truth values for the testing set: 3.1 (femoral neck-shaft angle), 5.9 (femoral neck axis length), 6.1 (femoral head diameter), 7.1 (hip axis length), 7.9 (shaft width), 8.5 (femoral neck width), 9.3 (vertical offset), 10.9 (horizontal offset), 18.2 (lateral cortical thickness), 21.8 (medial cortical thickness). Figure 2 also shows the predicted (red) and true (blue) points in the relative coordinate space for one image.

DISCUSSION: The proposed model is able to efficiently extract the 10 measurements of interest. However, its accuracy is limited by the absence of explicit locational data. Because landmark coordinates were not manually annotated in the images, we had to rely on a relative coordinate system that preserved the geometric relationships among points but did not specify their actual positions within the radiograph. Future work could explore unsupervised keypoint detection models to enable the network to better capture the structural geometry of the hip and femur or incorporate explicit landmark annotations during training to provide locational guidance. Additionally, future work could involve a larger dataset to potentially improve model accuracy and contribute to its viability in clinical AFF prediction.

SIGNIFICANCE/CLINICAL RELEVANCE: This convolutional neural network automates femoral geometry landmark detection and measurement for potential use in AFF prediction.

REFERENCES: [1] Taormina DP, Marcano AI, Karia R, Egol KA, Tejwani NC. Symptomatic atypical femoral fractures are related to underlying hip geometry. Bone. 2014 Jun;63:1-6. doi: 10.1016/j.bone.2014.02.006. Epub 2014 Feb 21. PMID: 24565751.

IMAGES AND TABLES:

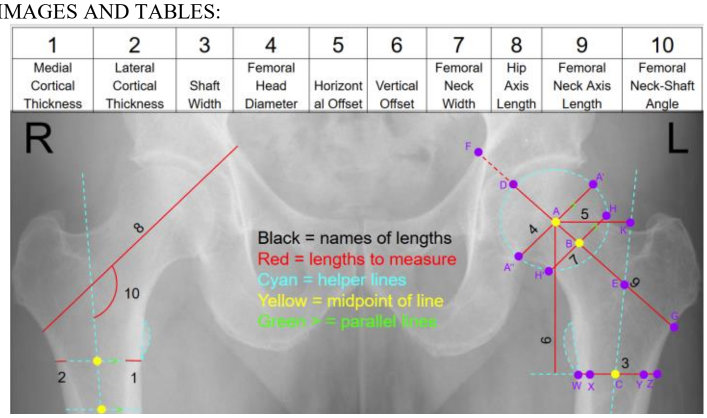

Figure 1: example of landmark points in the relative coordinate system

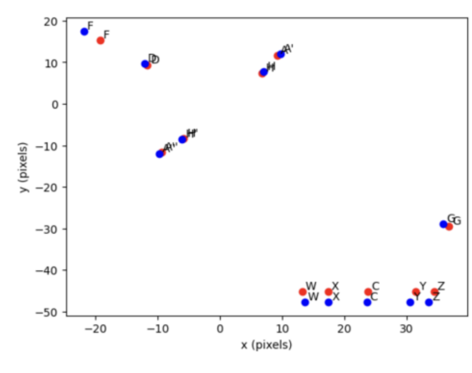

Figure 2: the predicted (red) and true (blue) points in the relative coordinate space for an image in the testing set