Foundation models for orthopedics

**Background**

Drugs commonly used in the treatment of osteoporosis (bisphosphonates) inhibit cell function of one specific cell type in bone, leading to increased bone mass and reduced fracture risk. This treatment has been used successfully for decades. Long-term inhibition of these bone-specific cells has recently been shown to cause bone material insufficiency, leading to spontaneous stress fractures in the thigh bone – Atypical Femoral Fractures (AFF). These fractures show features on x-ray images that differentiate them from Normal Femur Fractures (NFF). However, these features are very subtle and can easily be overlooked if not specifically sought for (Figure 1). The detection rate of AFF on clinical plain radiographs is <7%, and reports of drug adverse reactions to the Swedish Drug Agency have an even lower detection rate. While these events are rare compared to fractures that can be prevented, they are of clinical concern and have resulted in decreased use of these medications. As these events are so rare, standard statistical models have failed to identify reliable risk factors that would allow a precision medicine approach to identifying which patients to treat and for how long.

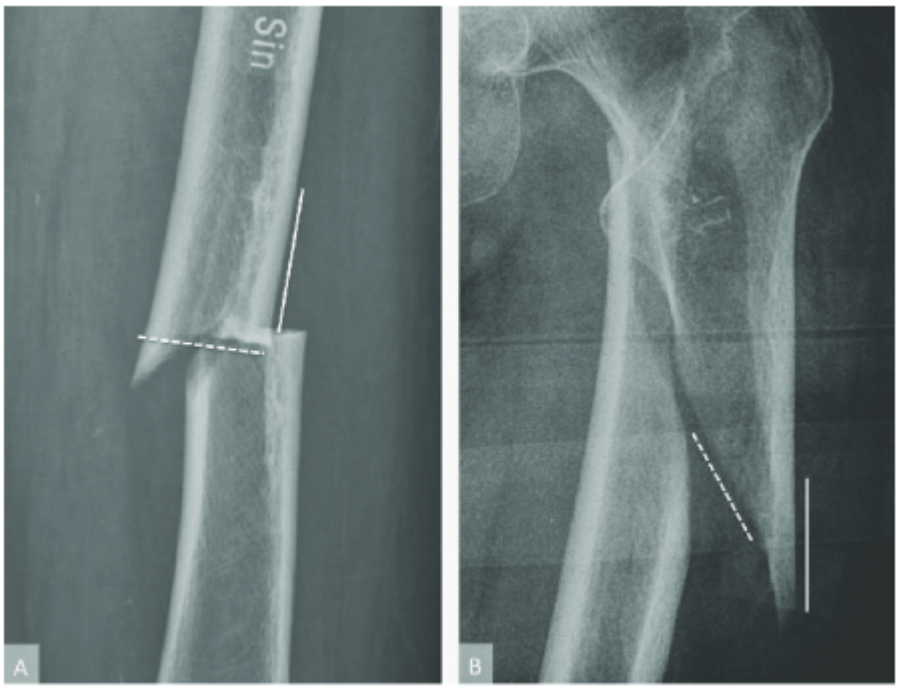


***Figure 1. Atypical femur fractures (AFF, image A) and normal femur fractures (NFF, image B).***

Convolutional neural networks (CNNs) and vision transformers (ViT) can be trained to detect AFF. A general problem in medical imaging is to create large datasets for training. Many CNNs and vision transformers are pre-trained on the large ImageNet dataset, and only a smaller dataset is then required for fine tuning. However, ImageNet does not contain any medical images. In this thesis the main idea is to train a foundation model for orthopedics, using our large dataset of 500,000 x-rays from 17,000 patients.

**Objectives**

Train and evaluate foundation models using our large dataset of femur fractures. Compare training the models from scratch and from other foundation models, e.g. DINO and RadImageNet.

How does the preprocessing of the images and the training approach affect the foundation model? How can the most informative images be found automatically?

Evaluate the foundation models using other types of medical images, how well does our foundation model generalize to other tasks in orthopedics (e.g. classification of elbow, knee or shoulder fractures) compared to other foundation models? Here other openly available datasets can be used.

Investigate how accurately a network (CNN or ViT) can classify AFF and NFF, using networks pre-trained on ImageNet, RadImageNet and our foundation models.

**Data**

About 500,000 X-ray images of femur fractures from 17,000 patients.

Several open datasets for other tasks in orthopedics, e.g. MURA which is a large dataset of musculoskeletal radiographs containing 40,561 images from 14,863 studies, where each study is manually labeled by radiologists as either normal or abnormal.

**Required background**

Machine learning, deep learning, Python programming

**Computing resources**

The student will have access to very good computing resources (graphics cards) for deep learning. Trainings can also be performed on the GPU cluster Berzelius.

**Contact persons**

Anders Eklund, anders.eklund@liu.se, Department of Biomedical Engineering, Department of Computer and Information Science, Center for Medical Image Science and Visualization

Jörg Schilcher, jorg.schilcher@liu.se, Department of Orthopaedic Surgery, University Hospital Linköping