Detection and classification of femur fractures using deep learning

**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).***

A general problem with the available dataset is that the images are of very different size, making the pre-processing important. In this thesis the main idea is to first train a network (CNN or vision transformer) to localize the center of the fracture using deep learning regression, then cut out a patch around this center, and use the patch for classification. This will avoid any downsampling of the images, where important information is lost.

**Objectives**

Investigate how accurately a network (CNN or vision transformer) can predict the center location of the fracture (ground truth = manual annotations). Is the performance similar for AFF and NFF images? If time permits, compare CNNs and vision transformers.

Compare the performance (classification accuracy) of downsampling all images to 224 x 224, and then classifying them as AFF or NFF, to first localize the fracture using an additional network, then automatically cutting out a patch of 224 x 224 pixels and doing the classification using these patches.

**Data**

About 4300 X-ray images from some 1200 patients, of which about 20% are AFF.

**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.

**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