Denoising of CT Images using Deep Learning

Juan I. Pisula.

Supervisors: Claudio Delrieux PhD, Felix Thomsen PhD

Abstract

The ionizing radiation produced by the CT scanners in the acquisition process has biological effects that can result in the incidence of cancer. Therefore, medical studies that make use of this imaging method posses an associated risk of causing this disease. Some sources estimate that in the USA, between 1.5% and 2% of cancer patients were caused by CT scanning, and that 29000 new cases will be related to CT scans performed in 2007.

For this reason, the radiation dose of the scan needs to be minimized as reasonable as possible. This can be achieved by shortening the exposure time, as well as reducing the tube current of the scanner. Nevertheless, even though reducing the radiation dose results in a less harmful medical study, the effect produced in the reconstructed CT images is an increase in noise levels, so a diagnostic procedure could be compromised.

The objective of this thesis is to study the application of Convolutional Neural Networks (CNNs) for noise reduction in Low Resolution Quantitative Computed Tomography (LRQCT). More concisely, on the particular case of CTs of human vertebrae. The final purpose of this concept would be to enhance the diagnosis process of osteoporosis, a skeleton pathology characterized by a low bone density and low bone structure quality, predisposing the person to a higher fracture risk. Having CT images of high quality, obtained with a low radiation dose would allow to perform a better analysis of the patient's bone structure, without subjecting them to a radiation level that could compromise their health.

Based on previous works on the topic, a CNN with an Encoder-Decoder architecture was designed for this task. The network is composed by 8 layers in total: 4 convolution layers followed by 4 transposed convolution layers. Every layer has 32 input channels, except for the first layer which has 1 input channel; and 32 output channels, except for the last layer which has 1 output channel. The Rectified Linear Unit function was chosen as activation, and is present at the output of every layer, except for the last one.

The dataset comes from five different human vertebrae. These vertebrae were scanned with two different devices, in order to have a set of clinical quality images to feed the network, and a set of high resolution images used as ground truth. The latter was acquired using a High Resolution Peripheral Quantitative

Tomography (HRPQCT) scanner. Furthermore, the clinical quality set was scanned with three different tube currents: 100mAs, 250mAs, and 360mAs; making three repetitions for each tube current.

Given the reduced number of CT scans of the dataset, small patches were extracted from the original volumes to augment the available data. Every LRQCT patch was recorded with its corresponding HRQCT patch, obtaining three sets of 6141 LRQCT-HRPQCT pairs (one for each tube current). A CNN per tube current was trained, using 85% of the data as training set, and 15% as testing set.

Before the final training of the models, a hyperparameter validation stage was done to select the values of learning rate and batch size. For this, the CNN corresponding to the 100mAs tube current was trained multiple time using different combinations of these values. These trainings were done using a subset of 10% of the data of the training set, and another subset of same size was used for testing. Finally, the values of learning rate $= 10^{-5}$ and batch size $= 32$ were chosen. The loss function used was L1, and the models were optimized with Adam.

The results show an improvement in the metrics of MAE, RMSE, and PSNR, and top the ones produced by a BM4D filter, a state of the art denoising filter that does not use Deep Learning. However, by comparing the values of Structural Similarity Index (SSIM), the only network that performs better than BM4D is the one trained with the 100mAs LRQCT. Having in mind these results, it is proposed to study the use of SSIM as loss function in a following iteration. Also, different networks architectures could be studied, like a conditional Generative Adversarial Network, where the current model could be used as the Generative Network.

Appendix: Figures

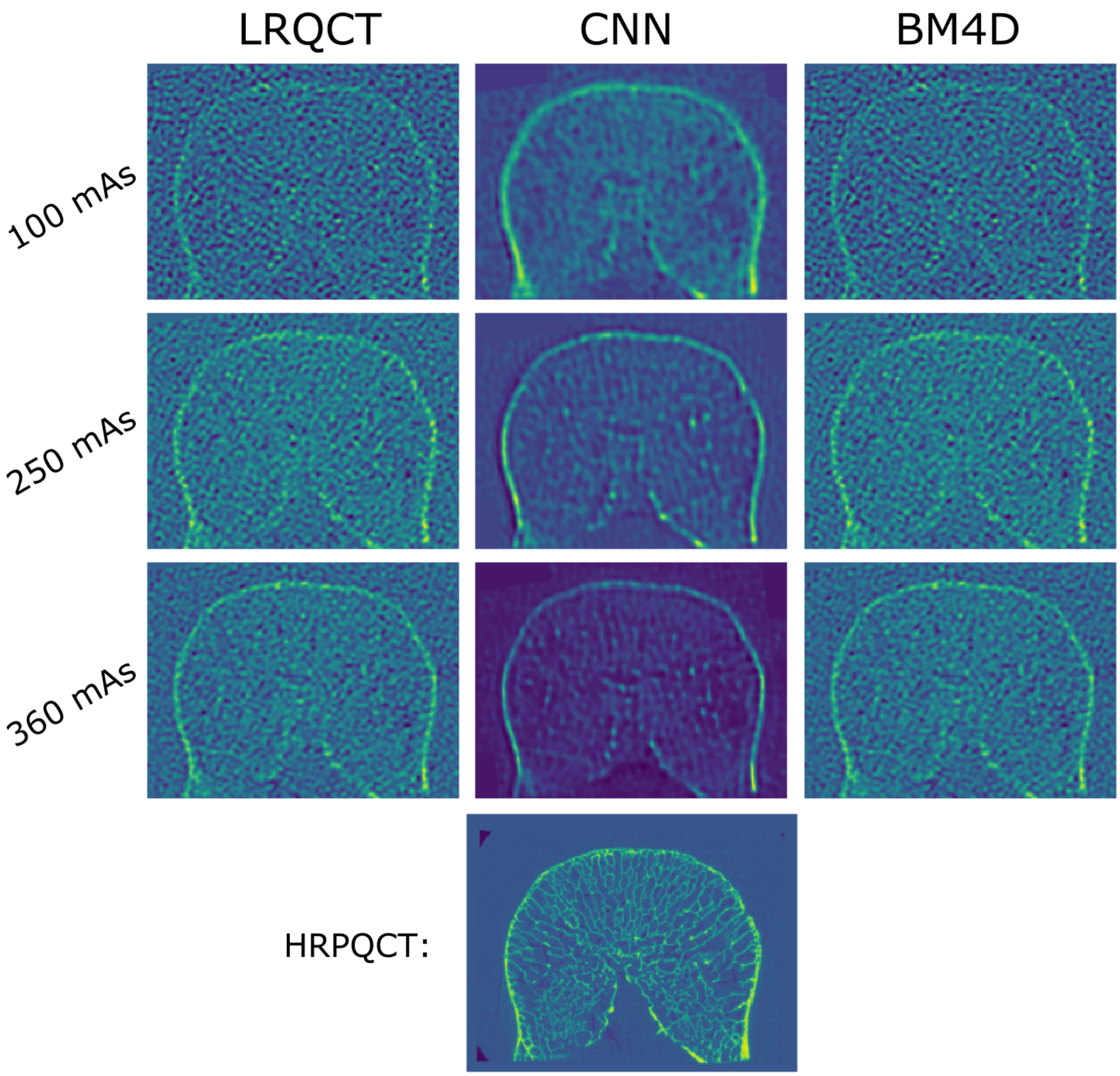

Figure 1: Comparison between vertebrae slices.

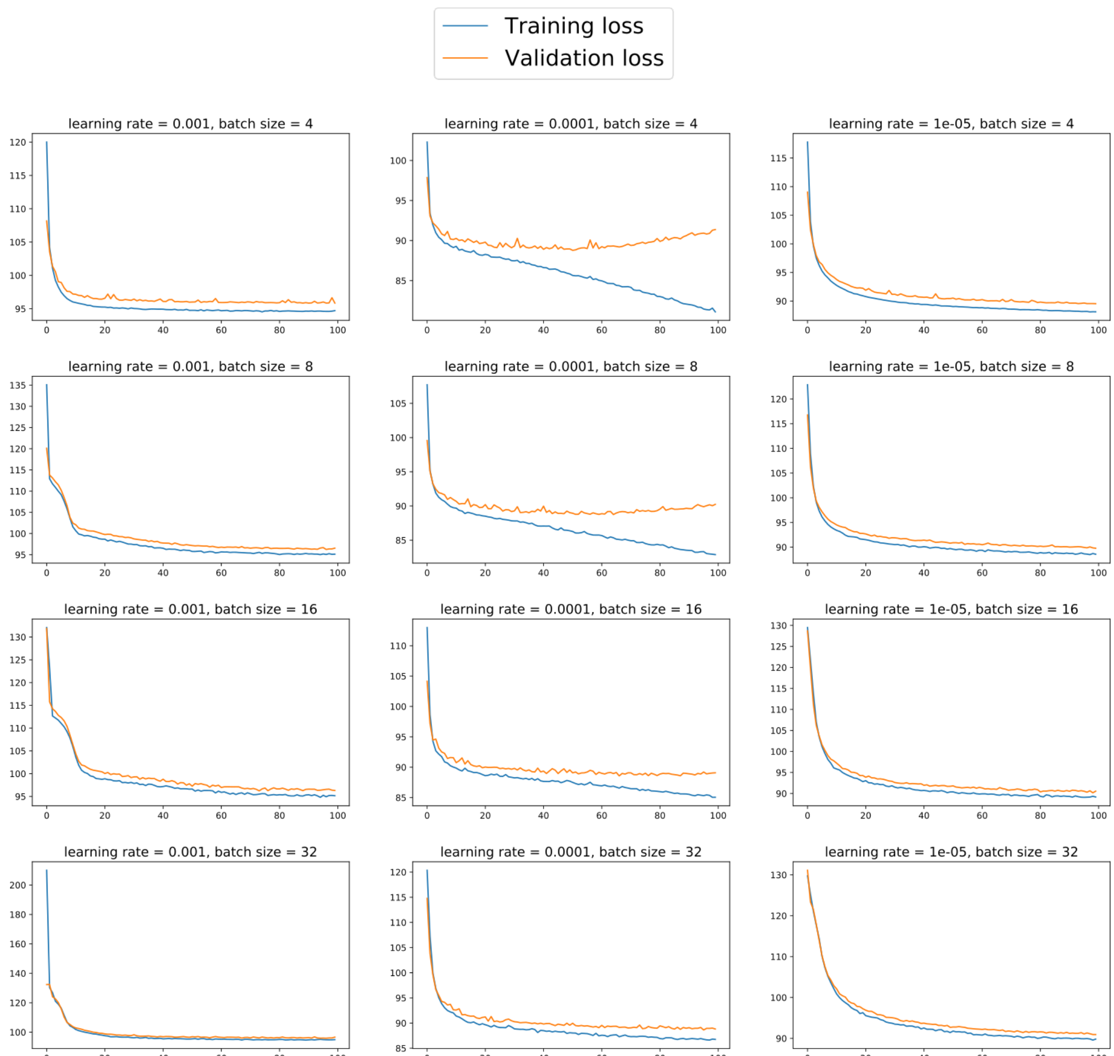

Figure 2: Learning curves for the hyperparameter validation stage.

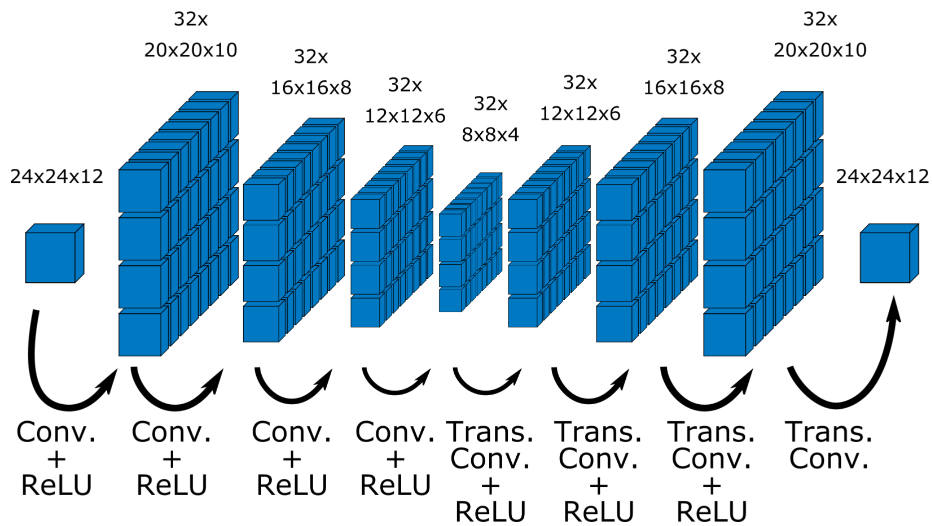

Figure 3: Architecture of the Encoder-Decoder CNN.