Large Cone Beam CT SCan Image Quality Improvement Using a Deep Learning U-Net Model

Joel Ruhe[1], Valeriu Codreanu[2], and Pascal Wiggers[1]

[1] Amsterdam University of Applied Sciences
[2] SURFsara, Amsterdam

Abstract. Cone beam CT scanners use much less radiation than to normal CT scans. However, compared to normal CT scans the images are noisy, showing several artifacts. The UNet Convolutional Neural Network may provide a way to reconstruct the a CT image from cone beam scans.

1 Introduction

Many people die annually from the effects of cancer. Most common and one of the deadliest type of cancers is lung cancer [1]. Primarily, lung cancer is being monitored using a CT scanner. The regular dose of a lung cancer ct scan is 1.5 millisieverts. This amount of millisievers forces the patient to take long brakes between the CT scans [2]. Cone beam CT scans are CT scans with much less radiation that can achieve relatively high image quality. However, there is more noise in the image than a regular CT scan [3].

CT image quality depends on the following factors: image contrast, spatial resolution and image noise. A Cone Beam Computed Tomography (CBCT) its ray source is cone-shaped and a two-dimensional detector is used that makes one rotation so that a 'volume' of data is obtained. The main advantage of the cone beam CT scanner is that it uses much less radiation compared to normal CT scans. The main disadvantage is that because of this, the quality of the image goes down (black 'stains', or group of pixels, appear in the image, on top of the general CT image noise). Artificial neural networks may provide a solution to this problem. In particular, we developed a U-NET model to improve the cone beam CT images so that they are 'restored' as good as possible to the quality of a normal CT scan. The question we addressed in the research described in this paper is: *How can a U-NET model, consisting of convolutional and deconvolutional layers, be used to improve 3D cone beam image quality of lung CT scans?*

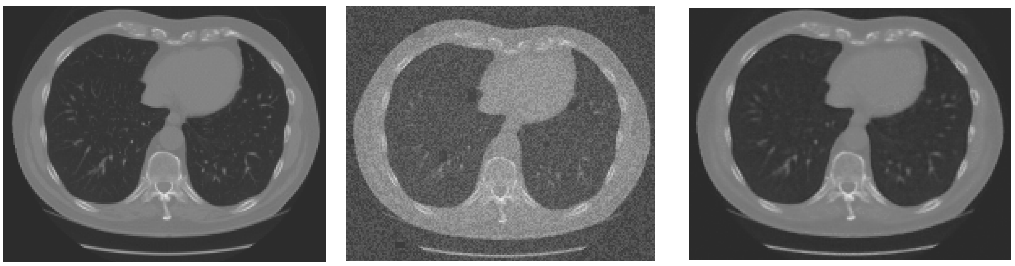
2 Model

Typical Convolutional Neural Networks (CNN's) consists of a convolution operation, Non-linearity (activation) function, pooling and fully connected layers. The convolutions create feature maps from the original input image. Because of this it can find recognizable patterns in the image. Different types of filters can be used depending on the type of feature maps one wants.

The U-NET model [4], developed to tackle the noise problem that occurs in CBCT images, consists of convolutional layers and transpose convolutional layers. This way, patterns can be found on 'what' is in the image, but also 'where' it is. When it is in the decoding path, it performs concatenation operations with the encoding blocks. This way, high resolution feature maps from the encoding blocks are being concatenated with the upsampled features. This will better learn representations with following convolutions and is also the main contribution of a U-NET model.

3 Experiments

To prove the principle, in this research a DICOM data set has been used that first had to be decoded to only get the raw pixel data on which the model can learn. Horovod has been used for data parallelism during the training. Later on, this might be extended to model parallelism to train on even larger image sizes if necessary. The U-NET model has been trained on 32x32x8, 64x64x16, 128x128x32, 256x256x128 and 512x512x128 image sizes. The images below show the results for an original image of size 256x256x64.



(a) Original image. (b) Cone beam image. (c) Reconstructed image.

From the image results, it is clearly noticeable that the U-NET model does a pretty good job in recreating the image with added noise. Only the small white matter in the middle of the lungs it has trouble with. We also found that the loss decreases as the image size increases. This also means that the U-NET model has greater precision on the 512x512x128 image due to the amount of feature maps it creates with each convolutional layer. The downside to this, is that it leaves a large memory footprint on the GPU when using large image size.

4 Conclusion

In short, it can be concluded that the U-NET model is suitable for recreating cone beam images and performs best at the image size 256x256x64. Furthermore, the U-NET may be directly applied to the CBCT images acquired from a commercial CBCT scanner after decoding the images and can directly be applied to real world problems.

References

1. World Health Organization Cancer. World Health Organization, September 12, 2018.
2. radiologyinfo. Radiation Dose in X-Ray and CT Exams. radiologyinfo, March 20, 2019.
3. Lee W. Goldman. Principles of CT: Radiation Dose and Image Quality. Journal of Nuclear Medicine Technology, November 15, 2007.
4. Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.