Automatic Segmentation of Aortic Anomaly CT Scans using U-Net

Ismael Perez
Department of Biomedical Engineering
Duke University
Durham, NC 27705
ismael.perez@duke.edu

Sayan Roychowdhury
Department of Biomedical Engineering
Duke University
Durham, NC 27705
sayan.roychowdhury@duke.edu

Abstract

Studies have shown that a morphological analysis of the aorta is important for both cardiovascular diagnosis and risk assessment in patients. Manual segmentation of the aorta in CT scans is currently a slow and inefficient task. The tediousness of this process is compounded by the inclusion of input noise, such as motion blur or a lack of contrast found in low-dose CT scans. Here, we implement a U-Net to develop an automatic method of segmenting these aorta CT scans containing noise. We train the network using the original, blurred (to simulate motion artifacts), and shot noise (to simulate low contrast) with previously annotated data. We find that with the inclusion of input noise, the model still performs very well, showing that this network could be applied in larger scale studies of this nature.

1 Introduction

The aorta is the main artery which carries blood away from the heart, connected to other major arteries to deliver oxygen-rich blood throughout the body. Any type of aortic problems can put the the entire body's blood supply in jeopardy. Thus, early diagnosis is critical for the management of aortic diseases. For example, in patients with aortic aneurysms, the aortic size has a profound impact on the risk of dissection [1], while abnormalities in the aortic arch can lead to respiratory difficulties [2]. Detecting aortic conditions at an early stage enables preventive surgery, which might save lives. Due of these types of deadly consequences, screening programs using computed tomography (CT) are recommended for at-risk populations.

Accurate segmentation of the aorta in CT scans can be used to analyze morphology and detect pathology. Most importantly, the location of specific shape changes in the aorta is relevant for diagnosis and risk assessment in patients. However, manual annotation of the aorta and its segmentation is a time-consuming and inefficient process. This task is further complicated by issues of noise, such as motion blur or lack of soft tissue contrast in low-dose CT scans.

The aim of this study is to develop and validate an automatic method to segment portions of the aorta in CT scans, with minimal human interaction. We evaluate on three sets of aorta CT scans, including the addition of simulated blur and low-contrast noise.

2 Related Work

Aorta segmentation is not a trivial task because the boundaries between the aorta and the surrounding structures are often unclear due to blurriness or lack of contrast. Previous works have developed algorithms for aortic segmentation based on a priori models; however these display very limited flexibility in capturing variablity. More recent work includes Bai et al. who combine a full convolutional network with a RNN, incorporating both spatial and temporal information on aortic MR

images [3]. de Vos et al. train a single convolutional network on CT scans of the chest and abdomen for localization (i.e. not segmentation) [4].

Ronnerberger et al. proposed the U-Net, a fully connected U-shaped convolutional neural network consisting of contracting and expanding pathways [5]. In this work, we apply the U-net to our aorta CT dataset and expand more on its structure in the Methods section.

3 Methods

In this section we describe the dataset we used for our study, the U-Net architecture, the motion artifact simulation parameters, shot noise parameters, and the approach to our simulations.

3.1 Datasets

In this study, we have a total of 1319 X-ray slices of 3 aortas (referred from now on as aorta6, aorta7 and aorta8) and only used axial slices (table for breakdown of each aorta CT scan in Table 3.1). All aortas were purchased from the Open Source Software Corporation and each one of them has an anomaly (aorta6 and aorta8: abnormal aortic arch , aorta7: aneurysm). Scans labeled as ground truth were segmented using MIMICS software. An example of the original and the annotated CT scans are shown in Figure 1.

Geometry	Total Slices	Slices removed
aorta6	342	0
aorta7	420	45
aorta8	348	164

Table 1: Breakdown of slices per aorta and how many were removed due to not having enough annotated pixels.

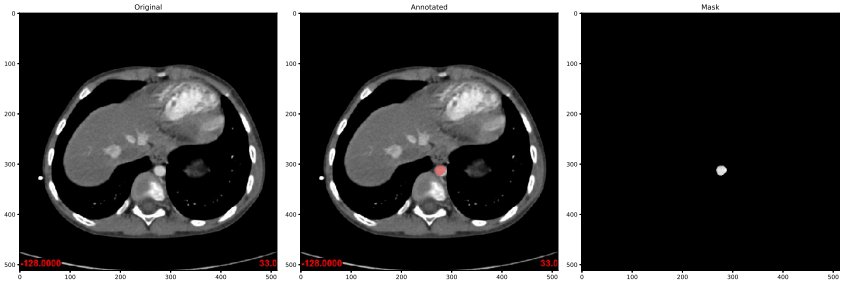

Figure 1: Left: Original image, Middle: Annotated image, Right: Final binary mask

3.2 U-Net architecture

The network architecture contains two paths, the contraction and expansion paths. The contracting path consists of the traditional stack of two 3x3 convolutions, then a rectified linear unit (ReLU), and a 2x2 max pooling operation with stride 2. The expansion path includes steps of an upsampling of the feature map, followed by a 2x2 convolution, along with two 3x3 convolution each followed by a ReLU. At the final layer, a 1x1 convolution is used to map each 64-component feature vector to the desired number of classes. In total the network has 23 convolutional layers [5].

3.3 Motion artifact

To simulate a motion artifact, we used a Gaussian kernel with three different standard deviations (σ = 0.01, 0.10, 1.00),

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{1}$$

The Gaussian kernel is convolved with the original image, which is shown in Figure 2.

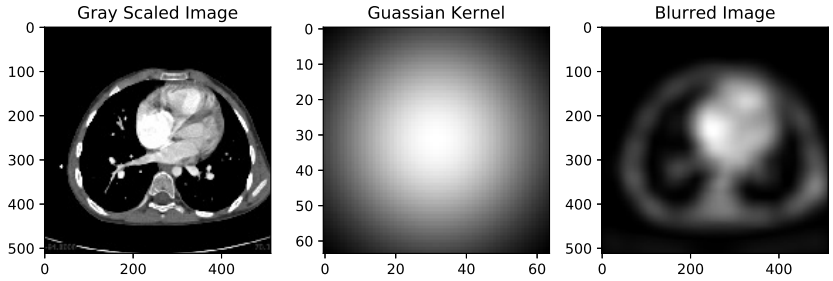

Figure 2: Left: original image, Middle: Gaussian kernel ($\sigma = 1.00$), Right: blurred image

3.4 Shot noise

We aim to model low-dosage CT scans by introducing shot noise into the images. This is motivated by the high usage of CT scans performed and the related risk associated with radiation [6]. Shot noise is also known as Poisson noise, which is signal dependent and so it is not an additive noise. In Figure 3, we show one image from the aorta6 dataset after the poisson noise is introduced (right image).

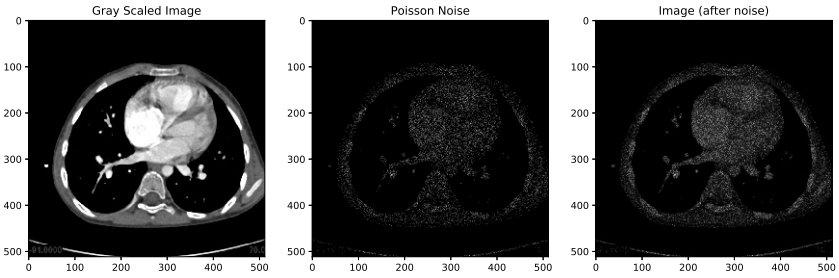

Figure 3: Left: original image, Middle: poisson noise (signal dependent), Right: image with noise

3.5 Approach

We implemented the UNET in Keras. Our hyper-parameters for the UNET are the following:

- Learning rate = 0.0001
- EPOCHS = 200
- Loss function = binary cross-entropy
- Batch size = 4
- Input size = 512 X 512 X 1

4 Results & Discussion

In this section we investigate the effects of introducing the noise to the images and changing the size of the training set. We resorted to using other metrics, structural similarity index, structural similarity index multi-scale, and peak signal to noise ratio.

4.1 Effects of Noise

In this section, we see the effect of the different noises described in the previous section. Figure 4 shows results for aorta6 for three cases: no noise, motion artifact with $\sigma = 1.00$, and shot noise. For this aorta, the blurring kernel performs the worst, which is also the case for aorta8 shown in Figure 5.

As expected, introducing noise images does affect the performance, but not by a big margin for this aorta.

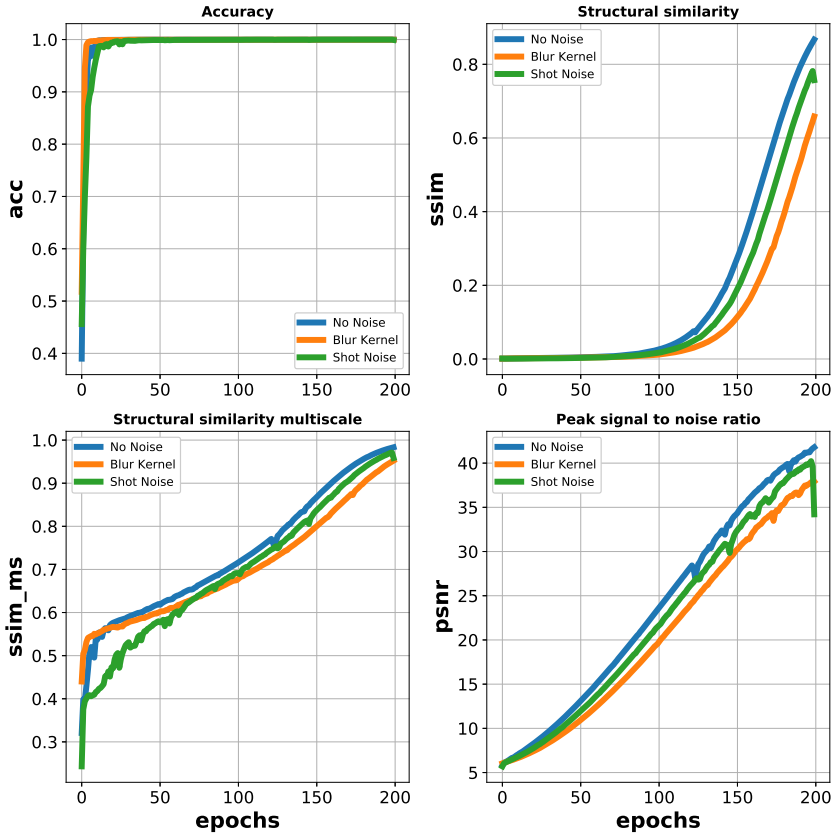

Figure 4: Training results for aorta6 and four metrics (accuracy,structural similarity, structural similarity multi-scale, and peak signal to noise ratio,).

For all the cases explored in Figure 4, we observe that after a couple of epochs we reach desirable accuracies if we define accuracy using a simple pixel-by-pixel comparison, but not for the other metrics. This explained why we were not seeing good predictions with our initial model using simply pixel-to-pixel.

In Figure 5, we show training results for aorta7 and aorta8 for only structural similarity index metric. As mentioned before, for aorta6 and aorta8 we see that the blurring kernel performs the worse. In the case for aorta7, the Gaussian kernel outperforms the case with no noise and shot noise. In other words, shot noise does worse than the blurring kernel. Note that aorta7 is the case with an aneurysm while the other aortas have an abnormal aortic arch.

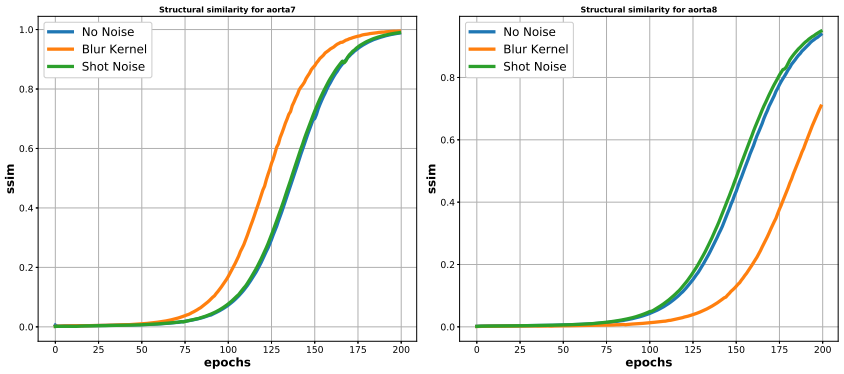

Figure 5: Training results for aorta7 (left) and aorta8 (right) and structural similarity index.

Figure 6 shows results for when we use the model with trained weights from all the aortas. The prediction was done on a subset of slices in aorta6 when adding shot noise. Although it does not predict one part of the annotation, we see an example of what is attributing to the bad performance for this type of noise.

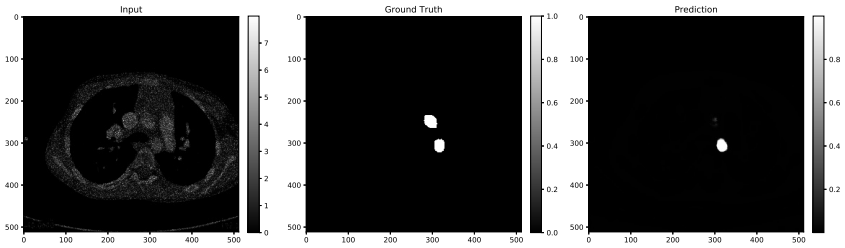

Figure 6: Left: input image, Middle: ground truth, Right: prediction

4.2 Effects of size of training set

To further validate the model, we decrease the the size of the training size to see how our model performs. Table 2 shows results for different training size (90%, 75%, 50%), and using all the aortas datasets while testing on a subset of slices in aorta6.

Training Size	acc	ssim	psnr	ssim_ms
90%	1.00	0.9996	46.4183	0.9997
75%	1.00	0.9998	51.8410	0.9997
50%	0.99	0.9912	40.8818	0.9983

Table 2: Train on all aortas with different training size (90%, 75%, 50%) and test on a subset of aorta6.

Our results clearly show that our model perform well for all the different metrics even when decreasing the size of the training set down to 50%. A follow up simulation for these results would be to find the minimum number of slices for each aorta in the training set to get similar results.

5 Conclusion

Overall, our results show that we can predict the masks with high accuracy, structural similarity(and multi-scale), and peak signal to noise ratio. Although during our training, we reach high pixel-to-pixel accuracy in a few epochs, our other metrics do not reach ideal values. We find that due to the sparsity of the the final mask, a simple pixel-by-pixel comparison did not prove to be a proper accuracy metric.

During training, the model perform the worse for the blurred kernel for aorta6 and aorta7 while it outperformed all the other cases in aorta8. Furthermore, aorta6 and aorta8 (with abnormal aortic arch), Gaussian blurred images performed worst. For aorta7 (with an aneurysm), the images with shot noise performed worst. For our results for the effects of size of training set we see that we are not affected by decreasing down to 50% but start to see a marginal decrease from the previous training sets size (90% and 75%).

Acknowledgments

We gratefully thank Bradley Feiger and Madhurima Vardhan and Professor Amanda Randles for the datasets. We also thank Professor Roarke Horstmeyer, Ouwen Huang, and Kevin Zhou for helpful advice and guidance on our project.

References

[1] R. R. Davies, et al., "Novel Measurement of Relative Aortic Size Prdicts Rupture of Thoracic Aortic Aneurysms," Ann. Thorac Surg, **81**, 169-177 (2006).

[2] S. Priya, et al., "Congenital Anomalies of the Aortic Arch," Cardiovasc Diagn Ther., **8**, S26-S44 (2018).

[3] W. Bai, et al., "Recurrent Neural Networks for Aortic Image Sequence Segmentation with Sparse Annotations," MICCAI, 586-594 (2018).

[4] B. de Vos, et al., "ConvNet-based Localization of Anatomical Structures in 3D Medical Images," IEEE Trans. Med. Imaging., **36**, 1470-1481 (2017).

[5] O. Ronnerberger, et al., "U-Net: Convolutional Networks for Biomedical Image Segmentation," MICCAI, 234-241 (2015).

[6] X. Duan, et al., "Electronic Noise in CT Detectors: Impact on Image Noise and Artifacts," AJR **201**, 626-632 (2013).