CS230

Super resolution for atrial fibrillation mapping

Francisco Sahli Costabal
Department of Mechanical Engineering
Stanford University
fsc@stanford.edu

Abstract

Atrial fibrillation is the most prevalent heart rhythm disorder. A novel procedure to treat this disease involves acquiring electrical activation maps using 64 electrodes. To interpret this kind of information, physicians usually rely on rule-based interpolation schemes. In this work, we use a convolutional neural network to enhance simulated, low resolution images of activation times in atrial fibrillation. Our models outperform bi-cubic interpolation and are able to better define sharp edges in the images. With further refinement, we expect this to be a useful tool to interpret atrial fibrillation maps.

1 Introduction

Atrial fibrillation is the most common rhythm disorder of the heart. It is associated with chaotic electrical waves that lead to rapid and irregular beating of the upper chambers. A novel procedure to diagnose and treat this disease is to insert a catheter with 64 electrodes into the chamber to identify electrical activation patterns [6]. However, this data is sparse and difficult to interpret due to its low resolution. Increasing the quality of these activation maps will help physicians to deliver more effective treatments. This problem is particularly suitable for deep learning, super resolution techniques. Here, we take simulated low resolution images of atrial fibrillation activation times and enhance them with a convolutional neural network. The input of our model are gray scale images that are downscaled and the output are the super resolution images, which we compare to the original image.

2 Related work

Traditionally, atrial fibrillation mapping has been interpolated using ruled-based approaches [6]. On the other hand, there are many approaches to do super-resolution using deep learning techniques. For example, generative adversarial networks have been used [5]. Other work focuses on the speed of the model to achieve real-time performance [7]. Additionally, there has been work regarding the loss functions to achieve better visual perception of the enhanced images [4]. One of the simplest and most effective models was introduced in [3]. This approach uses a convolutional network with 3 layers and a mean squared error loss function. We choose this approach over the other because we are not particularly interested in the speed of the model, such as in [5] or in the visual perception, such as in [7, 4]. We are interested in the accuracy of the predicted images with respect to the original images, and a mean squared error loss accomplishes this goal.

CS230: Deep Learning, Winter 2018, Stanford University, CA. (LateX template borrowed from NIPS 2017.)

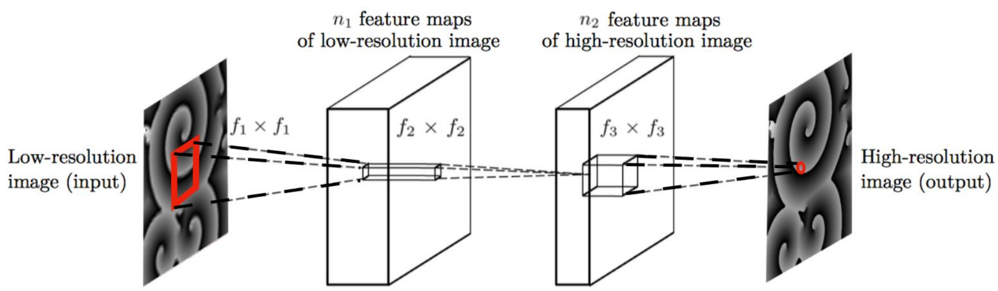

Figure 1: Convolutional neural network for super-resolution. The model is defined by layers with different number of filters and filter sizes. Adapted from [3]

3 Dataset and Features

Because is nearly impossible and unethical to collect simultaneously low resolution and high resolution data from human hearts, we use computational models that simulate the electrical activity of the heart [2]. We generate 400 images of regular activations and more than 800 images of fibrillation of 330x330, representing the activation times. Since super resolution approaches focus only on parts of the images, we obtain 100 sub-images per example of 33x33 pixels. The labels are, at the most, 21x21, depending on the specific filter sizes. Due to the nature of the electrical waves in the heart, rotated and flipped images also represent valid examples. We can augment our dataset 6 times, considering three 90 degree rotations and two flipping operations. The final dataset contains 718,200 images. We separate 10,000 images for development. For the test set, we generate 6000 more images. As features, we use the grey scale image values directly into the model and we also use this as the output.

4 Methods

We start by briefly describing the approach in [3]. This method uses 3 convolutional layers to predict the high-resolution counterpart of a low resolution image, upscaled using bi-cubic interpolation. The loss is defined as mean squared error between the pixels of the predicted high-resolution image and the training data:

$$\mathcal{L}(\mathbf{y}, \hat{\mathbf{y}}) = \frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{n_{pixel}} (y_{ij} - \hat{y}_{ij})^2 \tag{1}$$

where y represents the labels in the training set, y_{ij} is the value of the jth pixel of ith training example and $\hat{\mathbf{y}}$ represents the output of the model given the training input. Figure 1 shows a diagram of the model. We use Relu activations and He initialization for the parameters. We need 5 hyper-parameters to define the model: 3 filter sizes f_1, f_2 and f_3, and the number of filters for the intermediate layers n_1 and n_2. This model was implemented on Keras [1], based on the code found in `https://github.com/jormeli/srcnn-keras`. We use an Adam optimizer with default parameters and a mini-batch size of 32.

5 Experiments/Results/Discussion

Following the recommendations outlined in [3], we keep the number of layers to 3, as it has been shown that changing this parameter does not improve the performance significantly. Instead, we focus on the number of filters and the filter size of the second layer f_2. We set f_1 = 9 and f_3 = 5. We start by adjusting the learning rate. To do this, we select a model with parameters n_1 = 64, n_2 = 32, f_2 = 1 and a scaling of 3x. We vary the learning rate from 1e-2 to 1e-5 in a logarithmic scale. After training these models for 20 epochs, we select a learning rate of 1e-3 because it achieves the lowest training and development loss (Figure 2). We locked this value at this point.

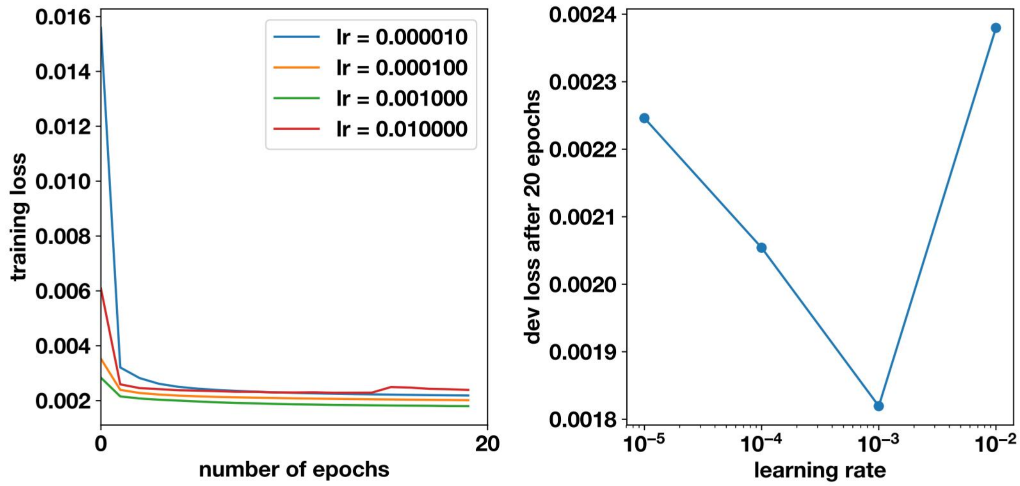

Figure 2: Learning rate tuning for n_1 = 64, n_2 = 32 and f_2 = 1. The optimal rate is 0.001

scale	n_1	n_2	f_2	training loss	development loss
3	64	32	1	1.80e-03	1.82e-03
3	64	32	3	1.68e-03	1.76e-03
3	128	64	1	1.70e-03	1.75e-03
3	128	64	3	1.46e-03	1.47e-03
3	128	64	5	1.79e-03	1.82e-03
3	bi-cubic			3.7e-03	
9	128	64	1	5.98e-03	6.22e-03
9	128	64	3	5.69e-03	5.95e-03
9	128	64	5	5.67e-03	5.89e-03
9	bi-cubic			1.21e-02	

Table 1: Summary of models evaluated

We tested our models with two levels of scaling: 3x and 9x. Our results, summarized in Table 1 show that the best models are the ones with more filters (n_1 = 128 and n_2 = 64) and that increasing the filter size f_2 above 3 either decreases the accuracy or marginally improves it. For both cases, our model performs better than bi-cubic interpolation. After selecting the model with n_1 = 128, n_2 = 64 and f_2 = 3, we evaluate them with an independent test set of 6000 images. For 3x scaling, the loss slightly higher: 1.61e-3, compared to a development loss of 1.47e-3, but this is still better than the other models. For 9x scaling, the test loss is 6.33e-3, which is higher than the training and development loss, but still close.

Qualitatively, as can be seen in Figure 3, the selected models produce images than are accurate for the smooth part of the images, but have problems in regions of sharp edges. Although, compared to bi-cubic interpolation, the models present a better approximation of the edges.

6 Conclusion/Future Work

We have used a convolutional neural network to enhance atrial fibrillation maps. Although the mean squared error was reduced with respect to a simple bi-cubic interpolation, the result are not visually pleasant. This could be caused by the mean squared error loss that we used. Despite this, the model is able to better interpret the sharp edges in the image. For the future, we would like to explore different loss functions that could improve our predictions. Additionally, we would also modify all filter sizes to find the optimal architecture.

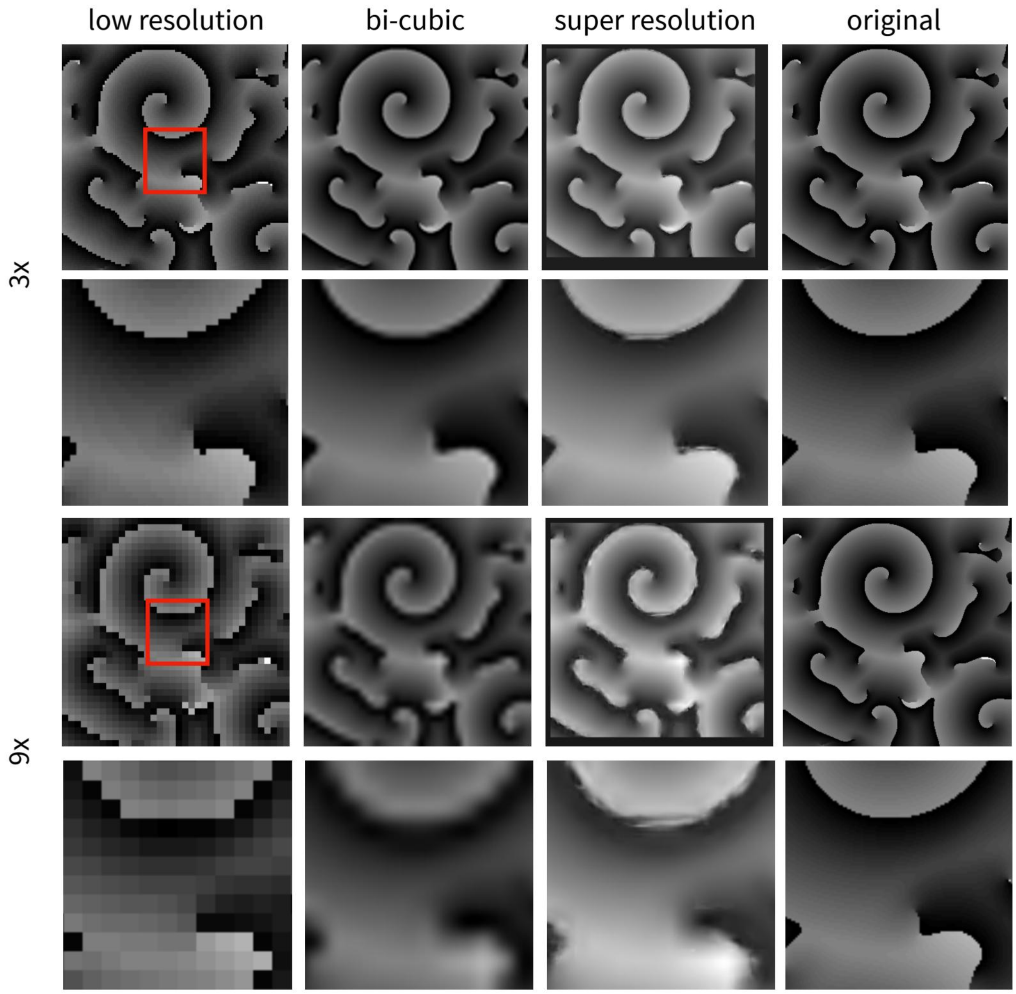

Figure 3: An example from the test set for 3x and 9x scaling where an entire image is reconstructed.

7 Contributions

Francisco Sahli Costabal did the entire project by himself, from the image generation to writing the report.

References

[1] François Chollet et al. Keras. `https://keras.io`, 2015.

[2] Francisco Sahli Costabal, Junaid AB Zaman, Ellen Kuhl, and Sanjiv M Narayan. Interpreting activation mapping of atrial fibrillation: A hybrid computational/physiological study. *Annals of biomedical engineering*, 46(2):257–269, 2018.

[3] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 38(2):295–307, 2016.

[4] Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In *European Conference on Computer Vision*, pages 694–711. Springer, 2016.

[5] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. *arXiv preprint*, 2016.

[6] Sanjiv M Narayan, David E Krummen, Kalyanam Shivkumar, Paul Clopton, Wouter-Jan Rappel, and John M Miller. Treatment of atrial fibrillation by the ablation of localized sources: Confirm (conventional ablation for atrial fibrillation with or without focal impulse and rotor modulation) trial. *Journal of the American College of Cardiology*, 60(7):628–636, 2012.

[7] Wenzhe Shi, Jose Caballero, Ferenc Huszár, Johannes Totz, Andrew P Aitken, Rob Bishop, Daniel Rueckert, and Zehan Wang. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1874–1883, 2016.