Anatomical Priors for Image Segmentation via Post-Processing with Denoising Autoencoders

Supplementary material

Submission 1542

1 Detailed description of the CNN and DAE architectures

UNet details: The UNet model (see Table 1) receives a 1024x1024 gray image as input and was trained using the soft Dice loss [1], batch size of 4, Adam optimizer with learning rate 1e-5 and the other parameters as by Keras default. We also used dropout for regularization, including a dropout layer after layer L_5 with keep probability p=0.5.

Post-DAE: Post-DAE (see Table 2)receives a 1024x1024 binary segmentation as input. The network was also trained to minimize the Dice loss function using Adam Optimizer. The best performance was achieve with learning rate 0.0001; batch size 15 and 150 epochs.

References

1. Milletari, F., Navab, N., Ahmadi, S.A.: V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: Proc. of Fourth International Conference on 3D Vision (3DV). (2016)

Table 1: Detailed description of the UNet architecture used as baseline model segmentation

		Kernel	Stride	#Kernels	NonLin
L1	Conv	(f:3,3)	(s:1,1)	(N:16)	ReLu
	Conv	(f:3,3)	(s:1,1)	(N:16)	ReLu
	Max Pooling	(f:2,2)	(s:2,2)		
L2	Conv	(f:3,3)	(s:1,1)	(N:32)	ReLu
	Conv	(f:3,3)	(s:1,1)	(N:32)	ReLu
	Max Pooling	(f:2,2)	(s:2,2)		
L3	Conv	(f:3,3)	(s:1,1)	(N:64)	ReLu
	Conv	(f:3,3)	(s:1,1)	(N:64)	ReLu
	Max Pooling	(f:2,2)	(s:2,2)		
L4	Conv	(f:3,3)	(s:1,1)	(N:128)	ReLu
	Conv	(f:3,3)	(s:1,1)	(N:128)	ReLu
	Max Pooling	(f:2,2)	(s:2,2)		
L5	Conv	(f:3,3)	(s:1,1)	(N:256)	ReLu
	Conv	(f:3,3)	(s:1,1)	(N:256)	ReLu
L6	UpConv	(f:3,3)	(s:1,1)	(N:128)	ReLu
	Conv	(f:3,3)	(s:1,1)	(N:128)	ReLu
	Conv	(f:3,3)	(s:1,1)	(N:128)	ReLu
L7	UpConv	(f:3,3)	(s:1,1)	(N:64)	ReLu
	Conv	(f:3,3)	(s:1,1)	(N:64)	ReLu
	Conv	(f:3,3)	(s:1,1)	(N:64)	ReLu
L8	UpConv	(f:3,3)	(s:1,1)	(N:32)	ReLu
	Conv	(f:3,3)	(s:1,1)	(N:32)	ReLu
	Conv	(f:3,3)	(s:1,1)	(N:32)	ReLu
L9	UpConv	(f:3,3)	(s:1,1)	(N:16)	ReLu
	Conv	(f:3,3)	(s:1,1)	(N:16)	ReLu
	Conv	(f:3,3)	(s:1,1)	(N:16)	ReLu
L10	Conv	(f:3,3)	(s:1,1)	(N:2)	ReLu
	Conv	(f:1,1)	(s:1,1)	(N:1)	Sigmoid

Table 2: Detailed architecture of the simple denoising auto encoder model used to implement the proposed Post-DAE.

		Kernel	**Stride**	**#Kernels**	**NonLin**
L_1	Conv	(f:3,3)	(s:2,2)	(N:16)	ReLu
	Conv	(f:3,3)	(s:1,1)	(N:16)	ReLu
L2	Conv	(f:3,3)	(s:2,2)	(N:32)	ReLu
	Conv	(f:3,3)	(s:1,1)	(N:32)	ReLu
L3	Conv	(f:3,3)	(s:2,2)	(N:32)	ReLu
	Conv	(f:3,3)	(s:1,1)	(N:32)	ReLu
L4	Conv	(f:3,3)	(s:2,2)	(N:32)	ReLu
	Conv	(f:3,3)	(s:1,1)	(N:32)	ReLu
L5	Conv	(f:3,3)	(s:2,2)	(N:32)	ReLu
L6	FC	-	-	(N:512)	None
L6	FC	-	-	(N:1024)	Relu
L8	UpConv	(f:3,3)	(s:1,1)	(N:16)	ReLu
	Conv	(f:3,3)	(s:1,1)	(N:16)	ReLu
L9	UpConv	(f:3,3)	(s:1,1)	(N:16)	ReLu
	Conv	(f:3,3)	(s:1,1)	(N:16)	ReLu
L10	UpConv	(f:3,3)	(s:1,1)	(N:16)	ReLu
	Conv	(f:3,3)	(s:1,1)	(N:16)	ReLu
L11	UpConv	(f:3,3)	(s:1,1)	(N:16)	ReLu
	Conv	(f:3,3)	(s:1,1)	(N:16)	ReLu
L12	UpConv	(f:3,3)	(s:1,1)	(N:16)	ReLu
	Conv	(f:3,3)	(s:1,1)	(N:1)	Sigmoid