## Supplementary Material for Deep 3D Dual Path Nets for Automated Pulmonary Nodule Detection and Classification

Wentao Zhu University of California, Irvine

wentaoz1@ics.uci.edu

Chaochun Liu
Baidu Research
liuchaochun@baidu.com

Wei Fan
Tencent Medical AI Lab
davidwfan@tencent.com

Xiaohui Xie University of California, Irvine xhx@ics.uci.edu

# 1. Detailed network structure for 3D Faster R-CNN with Deep 3D Dual Path Net in Nodule Detection

The encoder network is adapted from DPN92 directly by changing  $7\times7$  filters into  $3\times3$  [1]. The numbers of blocks are changed from 3, 4, 20, 3 to 2, 2, 2, 2. The decoder network is to make the network symmetric. The stride 2 of 3D convolution is added in the first  $3\times3\times3$  convolution in each block.

Stage	Output	Weights
Pre-dual	96×96×96, 24	3×3×3, 24
path		
Dual path	48×48×48, 48	$[1\times1\times1,24]$
block 1		$3\times3\times3$ , 24, (stride 2) $\times2$
		$\lfloor 1 \times 1 \times 1, 32 \rfloor$
Dual path	24×24×24, 72	<sub>1</sub> /1×1×1, 48
block 2		$3\times3\times3$ , 48, (stride 2) $\times2$
		[1×1×1, 56]
Dual path	12×12×12, 96	$[1\times1\times1,72]$
block 3		$3\times3\times3$ , 72, (stride 2) $\times2$
		$\lfloor 1 \times 1 \times 1, 80 \rfloor$
Dual path	6×6×6, 120	(1×1×1, 96
block 4		$\langle 3 \times 3 \times 3, 96, (\text{stride 2}) \rangle \times 2$
		[1×1×1, 104]
Deconv. 1	12×12×12, 216	2×2×2, 216
Dual path	12×12×12, 152	(1×1×1, 128
block 5		√3×3×3, 128
		[1×1×1, 136]
Deconv. 2	24×24×24, 224	2×2×2, 152
Dual path	24×24×24, 248	(1×1×1, 224
block 6		$3\times3\times3,224$ $\times2$
		$[1\times1\times1,232]$
Output	24×24×24, 3×5	Dropout, p=0.5
		$1\times1\times1$ , $64$
		1×1×1, 15

#### 2. Detailed network structure for 3D Faster R-CNN with Deep 3D Residual Network in Nodule Detection

The encoder network is adapted from Res18 directly by changing  $7\times7$  filters into  $3\times3$  [2]. We find the latest reference for 3D Res18 network in [3], and will add it into the reference.

Stage	Output	Weights
Pre-	96×96×96, 24	3×3×3, 24
Residual		3×3×3, 24
Residual	48×48×48, 32	$3\times3\times3,32$
block 1		$(3\times3\times3, 32, (stride 2))\times2$
Residual	24×24×24, 64	$\int 3\times 3\times 3$ , 64
block 2		$(3\times3\times3, 64, (stride 2))\times2$
Residual	12×12×12, 64	∫3×3×3, 64
block 3		$3\times3\times3$ , 64, (stride 2) $5\times3$
Residual	6×6×6, 64	$\int 3\times 3\times 3$ , 64
block 4		$(3\times3\times3, 64, (stride 2))\times3$
Deconv. 1	12×12×12, 128	2×2×2, 64
Residual	12×12×12, 64	∫3×3×3, 64\
block 5		3×3×3, 64√×3
Deconv. 2	24×24×24, 128	2×2×2, 64
Residual	24×24×24, 64	∫3×3×3, 64↑
block 6		\(\frac{1}{3} \times 3 \times 3, 64 \)\(\frac{1}{5} \times 3
Output	24×24×24, 3×5	Dropout, p=0.5
		1×1×1, 64
		1×1×1, 15

#### 3. Comparison with different methods for each fold and average false positives on LUNA16 dataset

Method	Deep 3D Res18	Deep 3D DPN26
Fold 0	0.8610	0.8750
Fold 1	0.8538	0.8783
Fold 2	0.7902	0.8170
Fold 3	0.7863	0.7731
Fold 4	0.8795	0.8850
Fold 5	0.8360	0.8095
Fold 6	0.8959	0.8649
Fold 7	0.8700	0.8816
Fold 8	0.8886	0.8668
Fold 9	0.8041	0.8122

Methods	0.125	0.25	0.5	1	2	4	8
DIAG_ConvNet	0.692	0.771	0.809	0.863	0.895	0.914	0.923
ZENT	0.661	0.724	0.779	0.831	0.872	0.892	0.915
Aidence	0.601	0.712	0.783	0.845	0.885	0.908	0.917
MOT_M5Lv1	0.597	0.670	0.718	0.759	0.788	0.816	0.843
VisiaCTLung	0.577	0.644	0.697	0.739	0.769	0.788	0.793
Etrocad	0.250	0.522	0.651	0.752	0.811	0.856	0.887
Dou et al 2017	0.659	0.745	0.819	0.865	0.906	0.933	0.946
3D RES	0.662	0.746	0.815	0.864	0.902	0.918	0.932
3D DPN	0.692	0.769	0.824	0.865	0.893	0.917	0.933

### **4. Detailed network structure for Deep 3D Dual Path Net in Nodule Classification**

We design 3D dual path network with 92 layers for nodule classification.

Stage	Output	Weights
Pre-dual	32×32×32, 64	3×3×3, 64
path		
Dual path	32×32×32, 320	$[1\times1\times1,96]$
block 1		$\langle 3 \times 3 \times 3, 96, (\text{stride } 2) \rangle \times 3$
		(1×1×1, 272
Dual path	16×16×16, 672	$(1\times1\times1, 192)$
block 2		$\langle 3\times 3\times 3, 192, (\text{stride } 2) \rangle \times 4$
		[1×1×1, 544]
Dual path	8×8×8, 1528	$[1\times1\times1,384]$
block 3		$3\times3\times3$ , 72, (stride 2) $\times20$
		(1×1×1, 1048
Dual path	4×4×4, 2560	$[1\times1\times1,768]$
block 4		$\langle 3 \times 3 \times 3, 768, (\text{stride } 2) \rangle \times 3$
		\(\lambda \times 1 \times 1, 2176 \)
Output	2560	3D average pool
	2	2560×2

#### References

- [1] Y. Chen, J. Li, H. Xiao, X. Jin, S. Yan, and J. Feng. "Dual path networks." In Advances in Neural Information Processing Systems, pp. 4468-4476. 2017.
- [2] K. He, X. Zhang, S. Ren, and J. Sun. "Deep residual learning for image recognition." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778. 2016.
- [3] F. Liao, M. Li, Z. Li, X. Hu, and S. Song. "Evaluate the Malignancy of Pulmonary Nodules Using the 3D Deep Leaky Noisy-or Network." arXiv preprint arXiv:1711.08324 (2017).