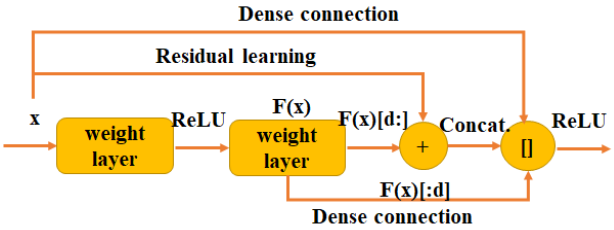
Update and Supplementary Material for Deep 3D Dual Path Nets for Automated Pulmonary Nodule Detection and Classification

Anonymous WACV submission

Paper ID 293

1. Fix the error in the manuscript

In the manuscript, Fig. 2 is a little hard to understand. The fixed figure is as follows.



Equation 1 is fixed as
$\mathbf{y} = \mathbf{G}([\mathbf{H}(\mathbf{x})[:d], \mathbf{F}(\mathbf{x})[:d], \mathbf{F}(\mathbf{x})[d:] + \mathbf{H}(\mathbf{x})[d:]])$,
where $\mathbf{H}(\mathbf{x})$ is the short cut activation employing 1×1×1 convolution. The 3D DPN has more feature maps and uses less filters for 3×3×3 convolution. More detailed network structure comparison is in Section 2 and Section 3.

2. 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 7×7 filters into 3×3 [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 3×3×3 convolution in each block.

Stage	Output	Weights
Pre-dual path	96×96×96, 24	3×3×3, 24
Dual path block 1	48×48×48, 48	{1×1×1, 24 3×3×3, 24, (stride 2) 1×1×1, 32} ×2
Dual path block 2	24×24×24, 72	{1×1×1, 48 3×3×3, 48, (stride 2) 1×1×1, 56} ×2
Dual path block 3	12×12×12, 96	{1×1×1, 72 3×3×3, 72, (stride 2) 1×1×1, 80} ×2
Dual path block 4	6×6×6, 120	{1×1×1, 96 3×3×3, 96, (stride 2) 1×1×1, 104} ×2
Deconv. 1	12×12×12, 216	2×2×2, 216
Dual path block 5	12×12×12, 152	{1×1×1, 128 3×3×3, 128 1×1×1, 136} ×2
Deconv. 2	24×24×24, 224	2×2×2, 152
Dual path block 6	24×24×24, 248	{1×1×1, 224 3×3×3, 224 1×1×1, 232} ×2
Output	24×24×24, 3×5	Dropout, p=0.5 1×1×1, 64 1×1×1, 15

3. 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 7×7 filters into 3×3 [2]. We find the latest reference for 3D Res18 network in [3], and will add it into the reference.

Stage	Output	Weights
Pre-Residual	96×96×96, 24	3×3×3, 24 3×3×3, 24
Residual block 1	48×48×48, 32	{3×3×3, 32 3×3×3, 32, (stride 2)} ×2
Residual block 2	24×24×24, 64	{3×3×3, 64 3×3×3, 64, (stride 2)} ×2
Residual block 3	12×12×12, 64	{3×3×3, 64 3×3×3, 64, (stride 2)} ×3
Residual block 4	6×6×6, 64	{3×3×3, 64 3×3×3, 64, (stride 2)} ×3
Deconv. 1	12×12×12, 128	2×2×2, 64
Residual block 5	12×12×12, 64	{3×3×3, 64 3×3×3, 64} ×3
Deconv. 2	24×24×24, 128	2×2×2, 64
Residual block 6	24×24×24, 64	{3×3×3, 64 3×3×3, 64} ×3
Output	24×24×24, 3×5	Dropout, p=0.5 1×1×1, 64

		1×1×1, 15

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 path	32×32×32, 64	3×3×3, 64
Dual path block 1	32×32×32, 320	{1×1×1, 96 3×3×3, 96, (stride 2) 1×1×1, 272} ×3
Dual path block 2	16×16×16, 672	{1×1×1, 192 3×3×3, 192, (stride 2) 1×1×1, 544} ×4
Dual path block 3	8×8×8, 1528	{1×1×1, 384 3×3×3, 72, (stride 2) 1×1×1, 1048} ×20
Dual path block 4	4×4×4, 2560	{1×1×1, 768 3×3×3, 768, (stride 2) 1×1×1, 2176} ×3
Output	2560 2	3D average pool 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).