Lung Nodule Classification with Multi-level Contextual 3D Convolutional Neural Networks

Qi Dou, Hao Chen, Pheng Ann Heng

Dept. of Computer Science and Engineering, The Chinese University of Hong Kong

Method Description

For the *false positive reduction* track of the LUNA16 Challenge, we formulate it as a binary classification problem and have developed 3D convolutional neural network (3D CNN) models to solve it. In order to meet the large size variations of the lung nodules [1], we leverage multi-level contextual 3D CNNs to confront the challenge and effectively reduce the false positive rate.

The 3D CNN is equipped with the 3D convolution and 3D max-pooling operations described in [2]. Based on the analysis of the distribution of pulmonary nodule sizes (see Fig. 1), we set three network architectures with different levels of contextual information taken into account. The model structures are shown in Table 1. The receptive files of Archi-a, Archi-b and Archi-c are designed as $20\times 20\times 6$, $30\times 30\times 10$ and $40\times 40\times 26$, respectively. Considering the relatively lower resolution in the third dimension, we used smaller kernel sizes in this dimension. Each of the model assigns a prediction probability to a candidate. To aggregate multi-level contextual information for the final classification, we fused the prediction probabilities from all architectures. The learned 3D kernels in the 1st layer of Archi-a are visualized in Fig. 2.

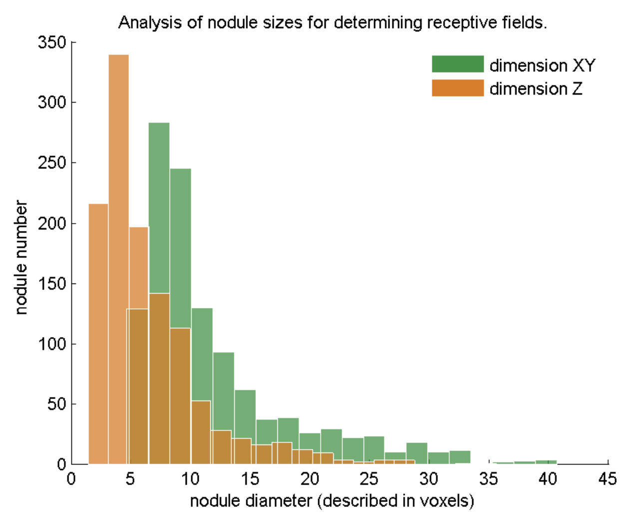

Fig. 1. Analysis of the sizes of pulmonary nodules, with diameters measured in voxels across different dimensions.

Table 1. Architectures of the multi-level contextual 3D CNNs.

Archi-a			Archi-b			Archi-c		
layer	kernel	channel	layer	kernel	channel	layer	kernel	channel
C1	5×5×3	64	C1	5×5×3	64	C1	5×5×3	64
M1	1×1×1	64	M1	2×2×1	64	M1	2×2×2	64
C2	5×5×3	64	C2	5×5×3	64	C2	5×5×3	64
C3	5×5×1	64	C3	5×5×3	64	C3	5×5×3	64
FC1	-	150	FC1	-	250	FC1	-	250
FC2	-	2	FC2	-	2	FC2	-	2

C: convolution, M: max-pooling, FC: fully-connected

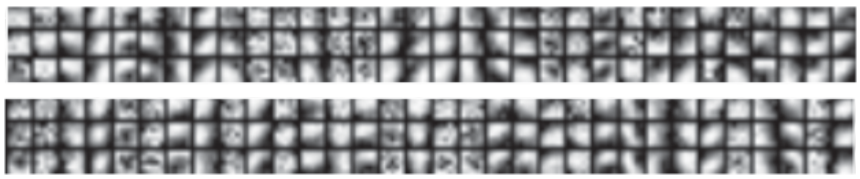

Fig. 2. The 64 learned 3D kernels of the 1st layer in Archi-a. Each $5 \times 5 \times 3$ kernel is embedded as three 5×5 maps presented in a column.

Importantly, to deal with the class imbalance, we conducted translation and rotation augmentations for the true nodule positions. In the cross-validation process, for example, when subset0-8 for training and subset9 for validation, 26.17% positive and 73.83% negative training samples were extracted. The rectifier nonlinearity was used in the network and the softmax function was used for logistic regression to yield the prediction probabilities. The weights of the 3D CNNs were initialized from Gaussian distribution and trained by minimizing the cross-entropy loss with stochastic gradient descent. The learning rate was initialized as 0.3 and decayed by 5% every 5000 iterations. Dropout and momentum were used during the training procedure. The method was implemented with Python based on the deep learning library of Theano [3].

References

1. Setio, A.A.A., Ciompi, F., Litjens, G., Gerke, P., Jacobs, C., van Riel, S., Wille, M.W., Naqibullah, M., Sanchez, C., van Ginneken, B.: Pulmonary nodule detection in ct images: false positive reduction using multi-view convolutional networks. IEEE Transactions on Medical Imaging (2016)
2. Dou, Q., Chen, H., Yu, L., Zhao, L., Qin, J., Wang, D., Mok, V., Shi, L., Heng, P.A.: Automatic detection of cerebral microbleeds from mr images via 3d convolutional neural networks. IEEE Transactions on Medical Imaging (2016)
3. Bastien, F., Lamblin, P., Pascanu, R., Bergstra, J., Goodfellow, I.J., Bergeron, A., Bouchard, N., Bengio, Y.: Theano: new features and speed improvements. Deep Learning and Unsupervised Feature Learning NIPS 2012 Workshop (2012)