

Diagnostic Classification of Lung CT Images using Deep 3D Multi-Scale Convolutional Neural Network

Ahmad P. Tafti^{*1}, Fereshteh S. Bashiri^{*1}, Eric LaRose¹, and Peggy Peissig¹

¹Marshfield Clinic Research Institute, WI 54449, USA

I. INTRODUCTION

LUNG cancer is the second most common cancer among men and women. It is the leading cause of cancer mortality worldwide and accounts for 1 in every 4 cancer deaths across the United States [1]. Accurate diagnosis of pulmonary lesions is of vital importance to early detection and effective treatment of lung cancer. Misclassification of the lesion can lead to additional medical costs, invasive surgery, or unnecessary lung biopsy. Low-dose computed tomography (CT) screening has been widely used to increase early diagnosis of lung cancer [2], [3], however, the differences between cancerous or non-cancerous pulmonary lesions in lung CT images are not immediately obvious or comprehensible, and make the task very challenging even for medical image analysis experts.

The primary aim of this study is to develop an advanced computer-aided diagnosis (CADx) method that extracts data from medical images efficiently and provides physicians and radiologists a precise and timely diagnosis of cancerous lung lesion. CADx methods, which are based on manually-designed features [4]–[10], are regularly incomplete, over-specified, and normally take a long time to be designed and validated. In contrast, methods based on automatically-designed features using neural networks and deep learning attempt to learn top-notch features automatically [11]–[17]. Despite the fact that the nature of CT images is three-dimensional, most of existing deep learning-based CADx methods use 2D convolutional neural networks (CNN) only, and therefore, throw off some partial information. Moreover, the existing 3D CNN models have primarily considered only single-space representation, which often discards valuable features that appear in different scales, and has been trained, built, and evaluated using a moderate-size of only publicly available data (e.g., The Cancer Imaging Archive (TCIA) [18]). Our effort addresses some of these limitations by developing a multi-scale 3D CNN for end-to-end diagnostic classification of lung CT images into cancerous or non-cancerous groups. We used both publicly and privately available CT image data sources to train and test our method.

II. METHODS

In this work, we exploited two large-scale datasets (Table I). The Data Science Bowl 2017 dataset (*DSB2017*) is

^{*}Corresponding authors: pahlavantafti.ahmad@marshfieldresearch.org & bashiri.fereshteh@marshfieldresearch.org

publicly and freely available, and the ground truth labels were provided by pathology diagnosis [19]. We also took advantage of privately held medical imaging data available at the Marshfield Clinic, named *Marshfield Clinic Lung Image Archive (MCLIA)*. With an Institutional Review Board (IRB) approval, the dataset was pooled out using the information obtained from the National Cancer Registry (NCR) and the Research Data Warehouse (RDW) of the Marshfield Clinic Health System (MCHS).

Figure 1 demonstrates the workflow of the proposed system, which consists of two tiers: 1) the scale-space tier, and 2) the 3D CNN tier.

Scale-space representation is motivated by the fact that real-world objects are composed of various structures and a wide range of scales [20]. The scale-space theory attempts to replicate the concept of physical dimension into mathematics, where objects have no scale, by introducing *scale-free* kernels. It considers representations at all scales by lowering resolution without introducing new details into the image [21]–[23].

In the current study, as the filter kernel size is predefined in the CNN network, scale-space representation is generated by providing input images in multiple scales to the 3D CNN tier. For the sake of memory usage and computing time, input images are resized into three scales including $50 \times 50 \times 20$, $100 \times 100 \times 20$, and $150 \times 150 \times 20$ (Figure 1).

To provide a large-scale dataset with sufficient variations, and to obtain a well-trained network that is able to cope with noisy images, an augmented dataset is generated by adding two Gaussian noise models of size 9 and 13 with the standard deviation of 2 and 3, respectively, in addition to the original images of each scale-space layer (Figure 1).

3D CNN automatically extracts a set of concise and discernible features from a given 3D CT scan. It normally consists of at least four internal layers, i.e., a 3D convolutional layer, activation function, 3D pooling, and a fully-connected neural network layer [24]. To overcome the problems associated with linear classifiers, we employed a non-linear layer, called the Rectified Linear Unit (ReLU), as the activation function. Besides the 3D pooling layer, which aims to reduce the spatial dimensions, we utilized dropout on the fully-connected layer to avoid over-fitting problem [25]. In the end, the softmax function calculates the probabilities of every target class over all possible target classes (Cancerous vs. Non-Cancerous) [26].

From the implementation perspective, three convolutional layers with a sequence of 32, 64, and 128 feature maps provided by 3D convolutional filters of size $3 \times 3 \times 3$, along

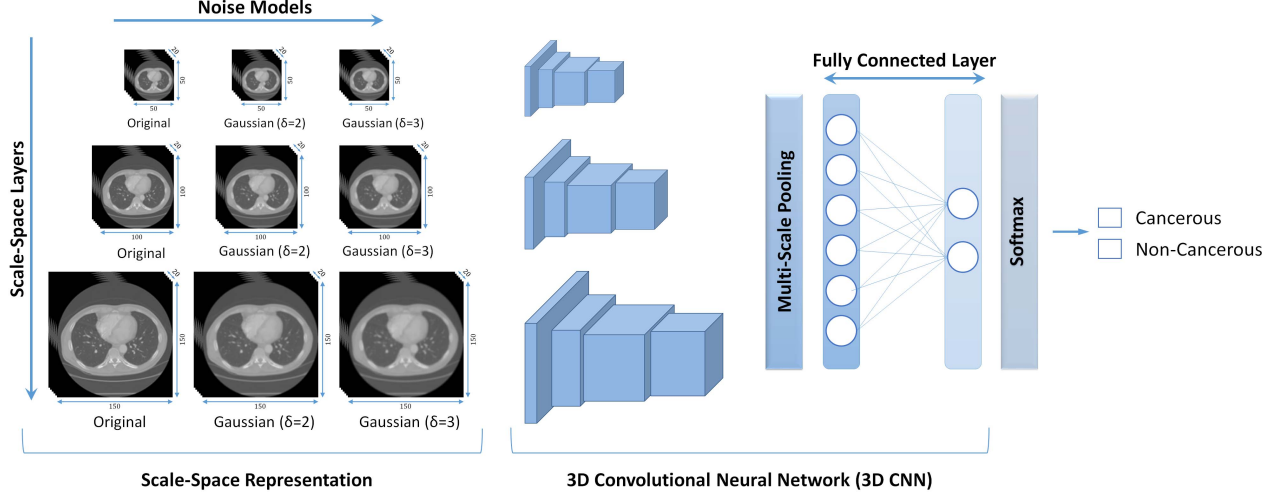


Fig. 1. The proposed multi-scale 3D convolutional neural network to diagnostic binary classification of lung CT images.

TABLE I

DATASETS' NAMES AND ATTRIBUTES. FOR EACH CLASS OF EACH DATASET, TOTAL 9 VERSION OF IMAGES WERE GENERATED USING COMBINATION OF 3 DIFFERENT SCALES AND 2 DIFFERENT LEVEL OF GAUSSIAN NOISE ALONG WITH ORIGINAL SET.

Dataset	Number of instances	
	Original	Augmented
DSB2017 [19]	Cancerous: 362	Cancerous: 3258
	Non-Cancerous: 1035	Non-Cancerous: 9315
	Total: 1397	Total: 12573
MCLIA	Cancerous: 1000	Cancerous: 9000
	Non-Cancerous: 1000	Non-Cancerous: 9000
	Total: 2000	Total: 18000

with two max-pooling layers with a filter of size $2 \times 2 \times 2$ were used. Padding has been utilized to manage the appropriate size of the feature map. Between layers, batch normalization was also performed [27].

III. RESULTS

We examined the performance of the proposed system on all datasets illustrated in (Table I). The accuracy, precision, recall, and AUC are shown in (Table II). Our benchmark was the basic 3D CNN which is indicated in (Table II) as “Original” datasets. The proposed 3D multi-scale CNN model, indicated in (Table II) as “Augmented” datasets, outperformed the basic 3D CNN. The results are statistically significant at ($p < 0.04$). The evaluation was performed using 4-fold cross-validation with maximum iteration set to 100 epochs.

IV. DISCUSSION

In this pilot study, a multi-scale 3D CNN is developed and studied for diagnostic classification of lung CT images. Preliminary results suggest that more accurate diagnostic classification model can be constructed with the use of combined publicly and privately held medical images, rather than each

TABLE II

ACCURACY, PRECISION, RECALL, AND AUC OF THE PROPOSED MULTI-SCALE 3D CNN ACROSS ALL DATASETS ILLUSTRATED IN TABLE I.

Dataset	Accuracy	Precision	Recall	AUC
DSB2017 (Original)	73.58%	73.19%	69.81%	0.774
DSB2017 (Augmented)	75.27%	73.80%	74.11%	0.803
MCLIA (Original)	73.29%	71.45%	70.38%	0.763
MCLIA (Augmented)	76.55%	77.04%	77.81%	0.832
DSB2017 + MCLIA (Original)	79.47%	77.47%	80.66%	0.872
DSB2017 + MCLIA (Augmented)	83.75%	85.17%	86.19%	0.926

data sources individually. Performing a t-test on AUC matched by those models shows statistically significant differences at ($p < 0.03$). With the use of augmented data, which covers multiple scale-space layers and noise conditions, the present model yields to higher accuracy. This contribution also demonstrates the potential of combined medical images integrated from multiple data sources in making a more accurate diagnostic model.

V. ACKNOWLEDGMENT

The authors of the paper wish to thank Marshfield Clinic. Our special thanks goes to Elham Sagheb at Marshfield Clinic Research Institute and Sawyer G Butterfield at Marshfield Clinic Information Services for their help and valuable contributions in providing the MCLIA dataset. We would like to express our very great appreciation to our colleagues Dr. Aaron R Dagit, DO, and Dr. William G Hocking, MD, from Marshfield Clinic who provided insight and expertise that greatly assisted the research. The authors gratefully acknowledge the support from Clinical and Translational Science Award (CTSA) program through the NIH National Center for Advancing Translational Sciences (NCATS) grant UL1TR000427 and the NIH BD2K Initiative grant U54 AI117924. Furthermore, we gratefully acknowledge the support of NVIDIA Corporation with the donation of the Quadro M5000 GPU used for this research.

REFERENCES

- [1] "Key statistics for lung cancer," website, last checked: 01.16.2018. [Online]. Available: <https://www.cancer.org/cancer/non-small-cell-lung-cancer/about/key-statistics.html>
- [2] W. G. Hocking, P. Hu, M. M. Oken, S. D. Winslow, P. A. Kvale, P. C. Prorok, L. R. Ragard, J. Commins, D. A. Lynch, G. L. Andriole *et al.*, "Lung cancer screening in the randomized prostate, lung, colorectal, and ovarian (plco) cancer screening trial," *JNCI: Journal of the National Cancer Institute*, vol. 102, no. 10, pp. 722–731, 2010.
- [3] N. L. S. T. R. Team *et al.*, "Reduced lung-cancer mortality with low-dose computed tomographic screening," *N Engl J Med*, vol. 2011, no. 365, pp. 395–409, 2011.
- [4] J. Kuruvilla and K. Gunavathi, "Lung cancer classification using neural networks for ct images," *Computer methods and programs in biomedicine*, vol. 113, no. 1, pp. 202–209, 2014.
- [5] B. Ganeshan, S. Abaleke, R. C. Young, C. R. Chatwin, and K. A. Miles, "Texture analysis of non-small cell lung cancer on unenhanced computed tomography: initial evidence for a relationship with tumour glucose metabolism and stage," *Cancer imaging*, vol. 10, no. 1, p. 137, 2010.
- [6] O. S. Al-Kadi, D. Watson *et al.*, "Texture analysis of aggressive and nonaggressive lung tumor ce ct images," *IEEE transactions on biomedical engineering*, vol. 55, no. 7, pp. 1822–1830, 2008.
- [7] H. Zhou, D. Dong, B. Chen, M. Fang, Y. Cheng, Y. Gan, R. Zhang, L. Zhang, Y. Zang, Z. Liu *et al.*, "Diagnosis of distant metastasis of lung cancer: Based on clinical and radiomic features," *Translational oncology*, vol. 11, no. 1, pp. 31–36, 2018.
- [8] H. M. Orozco, O. O. V. Villegas, V. G. C. Sánchez, H. d. J. O. Domínguez, and M. d. J. N. Alfaro, "Automated system for lung nodules classification based on wavelet feature descriptor and support vector machine," *Biomedical engineering online*, vol. 14, no. 1, p. 9, 2015.
- [9] A. Farag, A. Ali, J. Graham, A. Farag, S. Elshazly, and R. Falk, "Evaluation of geometric feature descriptors for detection and classification of lung nodules in low dose ct scans of the chest," in *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*. IEEE, 2011, pp. 169–172.
- [10] J. Yuan, X. Liu, F. Hou, H. Qin, and A. Hao, "Hybrid-feature-guided lung nodule type classification on ct images," *Computers & Graphics*, vol. 70, pp. 288–299, 2018.
- [11] W. Sun, B. Zheng, and W. Qian, "Computer aided lung cancer diagnosis with deep learning algorithms," in *Medical Imaging 2016: Computer-Aided Diagnosis*, vol. 9785. International Society for Optics and Photonics, 2016, p. 97850Z.
- [12] W. Alakwaa, M. Nassef, and A. Badr, "Lung cancer detection and classification with 3d convolutional neural network (3d-cnn)," *Lung Cancer*, vol. 8, no. 8, 2017.
- [13] K.-L. Hua, C.-H. Hsu, S. C. Hidayati, W.-H. Cheng, and Y.-J. Chen, "Computer-aided classification of lung nodules on computed tomography images via deep learning technique," *OncoTargets and therapy*, vol. 8, 2015.
- [14] D. Kumar, A. Wong, and D. A. Clausi, "Lung nodule classification using deep features in ct images," in *Computer and Robot Vision (CRV), 2015 12th Conference on*. IEEE, 2015, pp. 133–138.
- [15] J.-Z. Cheng, C.-M. Chen, and D. Shen, "Deep learning techniques on texture analysis of chest and breast images," in *Biomedical Texture Analysis*. Elsevier, 2018, pp. 247–279.
- [16] K. Kuan, M. Ravaut, G. Manek, H. Chen, J. Lin, B. Nazir, C. Chen, T. C. Howe, Z. Zeng, and V. Chandrasekhar, "Deep learning for lung cancer detection: Tackling the kaggle data science bowl 2017 challenge," *arXiv preprint arXiv:1705.09435*, 2017.
- [17] G. Litjens, C. I. Sánchez, N. Timofeeva, M. Hermesen, I. Nagtegaal, I. Kovacs, C. Hulsbergen-Van De Kaa, P. Bult, B. Van Ginneken, and J. Van Der Laak, "Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis," *Scientific reports*, vol. 6, p. 26286, 2016.
- [18] "THE CANCER IMAGING ARCHIVE (TCIA)," website, last checked: 01.10.2018. [Online]. Available: <http://www.cancerimagingarchive.net/>
- [19] Data Science Bowl, "Can you improve lung cancer detection?" website, January 2017, last checked: 02.15.2018. [Online]. Available: <https://www.kaggle.com/c/data-science-bowl-2017>
- [20] "Scale space," website, last checked: 02.09.2018. [Online]. Available: https://en.wikipedia.org/wiki/Scale_space
- [21] B. M. ter Haar Romeny, "Scale-space theory for multiscale geometric image analysis," in *CVPR99 IEEE International Conference on Computer Vision and Pattern Recognition*, 1999.
- [22] B. M. T. H. Romeny, "Introduction to scale-space theory: Multiscale geometric image analysis," in *Verlag. First International Conference on Scale-Space theory*. Citeseer, 1996.
- [23] T. Maintz, "Scale space," in *Digital and medical image processing*. Universiteit Utrecht, 2005, ch. 9, pp. 247–272.
- [24] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [25] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," *arXiv preprint arXiv:1207.0580*, 2012.
- [26] M. B. Christopher, *PATTERN RECOGNITION AND MACHINE LEARNING*. Springer-Verlag New York, 2016.
- [27] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *International conference on machine learning*, 2015, pp. 448–456.