Using Tissue-Energy Response to Generate Virtual Monoenergetic Images from Conventional CT for Computer-aided Diagnosis of Lesions

Shaojie Chang[a], Yongfeng Gao[a], Marc J. Pomeroy[a], Ti Bai[b], Hao Zhang[c], Zhengrong Liang[*d]

[a] Department of Radiology, Stony Brook University, Stony Brook, NY 11794, USA;[b] Department of Radiation Oncology, University of Texas Southwestern Medical Centre, Dallas, TX 75390, USA; [c]Department of Medical Physics, Memorial Sloan Kettering Cancer Center, NY 10065, USA; [d]Departments of Radiology and Biomedical Engineering, Stony Brook University, Stony Brook, NY 11794, USA

ABSTRACT

Based on the X-ray physics in computed tomography (CT) imaging, the linear attenuation coefficient (LAC) of each human tissue is described as a function of the X-ray photon energy. Different tissue types (i.e. muscle, fat, bone, and lung tissue) have their energy responses and bring more tissue contrast distribution information along the energy axis, which we call tissue-energy response (TER). In this study, we propose to use TER to generate virtual monoenergetic images (VMIs) from conventional CT for computer-aided diagnosis (CADx) of lesions. Specifically, for a conventional CT image, each tissue fraction can be identified by the TER curve at the effective energy of the setting tube voltage. Based on this, a series of VMIs can be generated by the tissue fractions multiplying the corresponding TER. Moreover, a machine learning (ML) model is developed to exploit the energy-enhanced tissue material features for differentiating malignant from benign lesions, which is based on the data-driven deep learning (DL)-CNN method. Experimental results demonstrated that the DL-CADx models with the proposed method can achieve better classification performance than the conventional CT-based CADx method from three sets of pathologically proven lesion datasets.

Keywords: CT image analysis, Computer-aided diagnosis, Machine learning, Malignant and benign differentiation

1.INTRODUCTION

IN computed tomography (CT) imaging, different tissue types can be represented by the linear attenuation coefficients (LACs) [1]. Based on the well-established X-ray physics inside the human tissues [2], Fig. 1(a) shows the LACs as a function of the X-ray energy for four important human body tissues, i.e., bone, muscle, fat, and lung, and water as the reference. Fig. 1(b) shows the difference between these tissues in terms of CT values. Fig. 1(c) is a zoomed version of Fig. 1(b) to emphasize the differences among muscle, fat, lung and water. It is clearly seen that different tissue types have their energy responses along the energy axis, which we call tissue-energy response (TER) in this study. The different TERs shall bring more tissue contrast distribution information in a series of monoenergetic images than that of the single image reconstructed from the conventional CT, which uses a wide spectrum and could not take the response into consideration. More tissue contrast distribution information shall enhance tissue characterization and, therefore, improve lesion diagnosis. If the energy-independent fraction of each tissue can be obtained, a series of virtual monoenergetic images (VMIs) is able to be generated by the TER curve. Hence, photon-counting CT (PCCT) [3] and multiple energy spectral CT (MECT) [4-6] are recently developed to reconstruct the energy-independent fractions with multiple energy measurements and then generate a series of VMIs. However, they require the use of expensive photon counting detection technology and/or sophisticated image reconstruction methods. Following the physics behind the TER, this study explores an alternative approach to use the TER to generate a series of VMIs from a conventional CT image to further extract contrast textures for lesion diagnosis. An assumption is made that the conventional CT image is obtained from the effective energy of the energy spectrum used by the
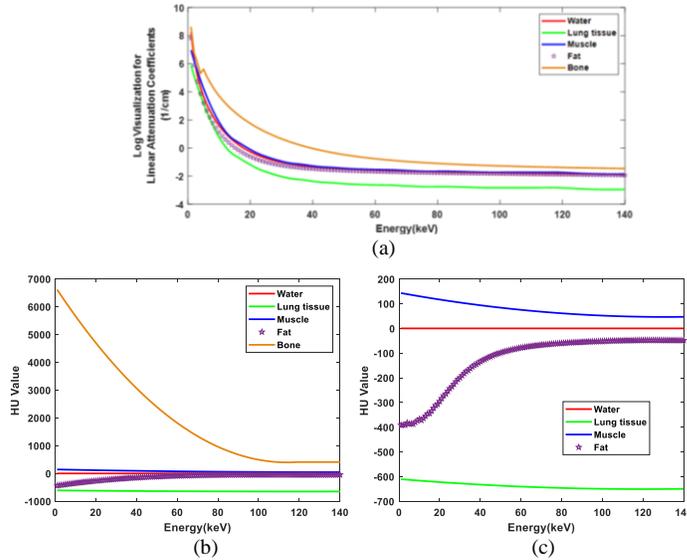
Fig. 1: Tissue energy response curve. (a) Log visualization for linear attenuation coefficients curve; (b) HU curve; (c) Zoomed HU curve.

conventional CT. For example, if conventional CT operates on a 120kVp X-ray tube voltage, the effective photon energy would be 75keV. With the identified location in the TER curve, a series of VMIs are obtained by the relative factor along the energy axis.

With the above VMIs, a machine learning (ML)-based CADx model is proposed to differentiate malignant from benign lesions by exploiting the energy-enhanced material features, which uses automatically extracted features by deep learning (DL) technics. DL-based CADx algorithms, more specifically convolutional neural network (CNN), have achieved noticeable successes in the differentiation of malignant and benign lesions [7-11]. Due to the advanced feature learning power, a multi-channel 3D CNN-based CADx model is developed in this study to help recognize energy-specific features to differentiate malignant lesions from benign ones. The final classification result will reflect the CADx performance with the explored spectral information.

The remainder of this paper is organized as follows. Section II will describe the proposed computer-aided diagnosis framework and the overall workflow. Section III presents the experiment design and results. Discussion and conclusions are drawn in Sections IV and V.
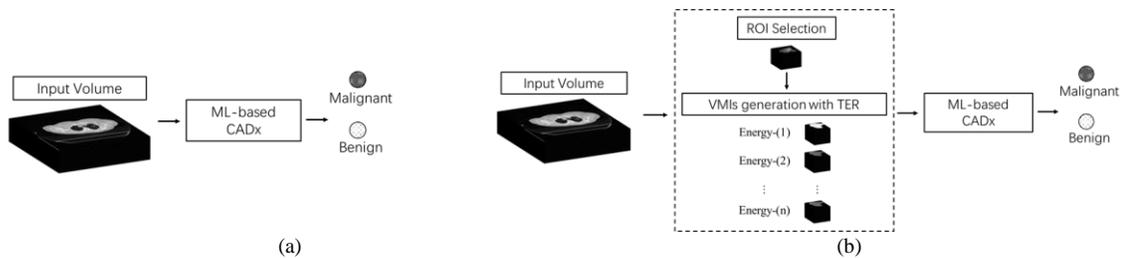
Fig. 2: Illustration and comparison between traditional and our proposed method for lesion diagnosis. (a) Traditional ML-based CADx pipeline. (b) Our proposed ML-based CADx pipeline with the VMIs generated by the use of TER.

2. METHODS

An overall illustration of our proposed lesion diagnosis pipeline can be found in Fig.2. Fig.2(a) shows the traditional diagnosis pipeline, the conventional CT data is directly used in the ML-based CADx model. As shown in Fig.2(b), our proposed framework is composed of two components: one is virtual monoenergetic images (VMIs) generation by the use of TER and the other is ML-based CADx model for classification. The details of each are described as follows.

2.1 CT Image-based Virtual Monoenergetic Images Generation by the Use of TER

In a CT image, a linear attenuation coefficient can be represented by R types of tissues, e.g. muscle, fat, lung, bone, and so on. The LAC function $\mu_j(\varepsilon)$ at the jth pixel of the image is decomposed as:

$$\mu_j(\varepsilon) = \sum_{r=1}^{R} \mu_r(\varepsilon) f_{rj}, \qquad (1)$$

where $\mu_r(\varepsilon)$ denotes linear attenuation coefficient of tissue r at energy ε, as shown in Fig.1(a). Notation f_{rj} is a unitless tissue fraction that quantifies the contribution of tissue r to attenuation in pixel j. Once the tissue fractions are identified, a series of VMIs can be generated.

1) Region of interest (ROI) selection
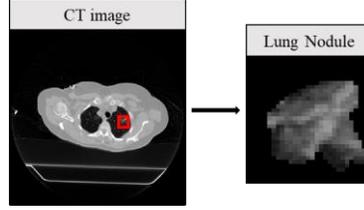
Fig. 3: An example of ROI selection from the CT image for lung nodule diagnosis. The display window is [0,0.35] cm^{-1}.

For lesion diagnosis, an ROI containing the lesion is firstly selected from the whole CT image. An example of ROI selection from one patient slice for lung nodule diagnosis is shown in Fig. 3.

2) Tissue fractions identification

Due to the polychromatic X-ray source, a CT ROI image μ^{E_1} reconstructed by filtered backprojection (FBP) method reflects the attenuation coefficients at the effective energy with the tube voltage E_1 kVp as follows,

$$\mu_j^{E_1} = \sum_{r=1}^{R} \mu_r(E_1^{eff}) f_{rj}, \qquad (2)$$

where E_1^{eff} denotes the effective energy the selected tube voltage E_1. To identify the tissue fractions, the CT image is first segmented into four tissue types including lung tissue, fat, bone and muscle by a threshold method. And then, for each tissue region, the tissue fraction f_{rj} can be obtained by $\frac{\mu_j^{E_1}}{\mu_r(E_1^{eff})}$. In this work, we assumed $E_1^{eff} = 75$ when a conventional CT scan at $E_1 = 120$ kVp. And $\mu_r(E_1^{eff})$ can be directly found with TER as shown in Fig.1(a).

3) Virtual monoenergetic CT images generation

Based on the identified tissue fractions, a series of VMIs at selected n energies are generated with the corresponding tissue LACs as follows.

$$\mu_j^{virtual}(\varepsilon) = \sum_r \mu_r(\varepsilon) f_{rj}, \qquad (3)$$

In this work, $n = 10$, the energy values are used by 5, 8, 10, 12, 15, 20, 25, 30, 35, 40 and 45 keV as an example. It is because that as shown in Fig. 1(c), the HU values of different tissues have the maximum differences in this energy range and the contrast features could be efficiently enhanced.

2.2 Machine Learning (ML)-based CADx
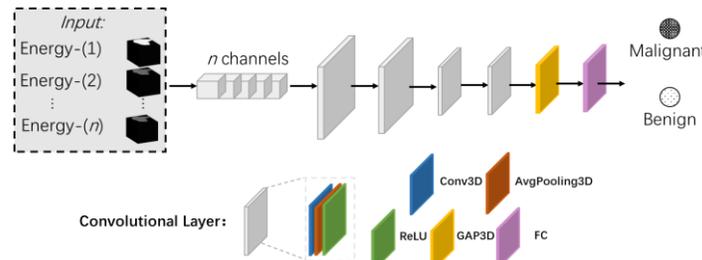
Fig. 4: The 3D CNN architecture of CADx for lesion diagnosis.

4) Deep learning (DL)-based CNN model:

For the DL-based model, a 3D CNN architecture with a multi-channel input is developed, which uses each energy image as one input channel as shown in Fig. 4. First, four convolutional layers are used to extract the features, which are then pooled together with a global average pooling (GAP3D) layer such that the final features have a global receptive field. At last, a fully connected (FC) layer-based classifier is adopted to distinguish the malignant and benign lesions. Specifically, each convolutional layer consists of three operators: 3D convolution (Conv3D), 3D average pooling (Avgpolling3D), and rectified linear unit (ReLU). And the binary cross-entropy loss is utilized to train the model. The details of the CNN model are listed in Table I.

Based on the above model, the final classification results will show the lesion diagnosis performance with the VMIs generated by the TER.

TABLE I: DETAILS OF NETWORK DESIGN

Layer	Type	Channels	Kernel Size	Padding	Stride	Activation
1	Conv3D	32	(7,7,7)	(1,1,1)	(1,1,1)	ReLU
2	Avgpooling3D	-	(2,2,2)	-	(2,2,2)	-
3	Conv3D	64	(5,5,5)	(1,1,1)	(1,1,1)	ReLU
4	Avgpooling3D	-	(2,2,2)	-	(2,2,2)	-
5	Conv3D	128	(3,3,3)	(1,1,1)	(1,1,1)	ReLU
6	Avgpooling3D	-	(2,2,2)	-	(2,2,2)	-
7	Conv3D	128	(3,3,3)	(1,1,1)	(1,1,1)	ReLU
8	Avgpooling3D	-	(2,2,2)	-	(2,2,2)	-
9	GAP3D	-	-	-	-	-
10	FC	128	-	-	-	-

3. EXPERIMENTS AND RESULTS

3.1 Datasets

In this study, three pathologically proven clinical datasets consisting of colon polyp and lung nodules are used to evaluate our proposed method. All the patients were recruited to this study under informed consent after approval by the Institutional Review Board. Details are presented in Table II.

TABLE II: DATASETS INFORMATION

Dataset	Total Number	Benign (0)	Malignant (1)	Pathological Report
Dataset 1	63	31	32	✓
Dataset 2	67	18	49	✓
Dataset 3	114	50	64	✓

1) Dataset 1

In dataset 1, 59 patients were scanned by a conventional CT at 120 kVp with automatic exposure control at the University of Wisconsin, USA. A total of 63 colon polyp masses were found and resected by the clinical examination. The pathological reports indicate 31 benign and 32 malignant polyps.

2) Dataset 2

In dataset 2, 66 patients were scheduled for CT-guided lung nodule needle biopsy at 120 kVp with automatic exposure control at Stony Brook University Hospital, USA. With the pathological reports, a total of 67 lung nodules with 18 benign and 49 malignant were confirmed.

3) Dataset 3

In dataset 3, 114 patients were scheduled for CT-guided lung nodule needle biopsy with X-ray exposure of clinical dose at 120 kVp, 100 mAs in Stony Brook University Hospital, USA. With the pathological report, a total of 114 lung nodules with 50 benign and 64 malignant were confirmed.

3.2 CNN Training Implementation

For the input to the CNN-based implementation, we first converted each *n*-energy data with the resolution of 64×64×64 voxels. And these converted energy volumetric images were fed into the multi-channel 3D CNN as shown in Fig. 4 for training. And the target is the results from the pathological reports of the malignant and benign lesions. The k-fold (k=5) cross-validation was implemented to test the robustness and avoid data bias. The procedure is as follows. We firstly shuffled the dataset randomly and split it into 5 folds. For each fold, we randomly divided the dataset into training and testing datasets. And then we trained a model on the training dataset and evaluated it on the testing dataset. Finally, we retained the evaluation score for each fold and the average score was calculated. In this study, the CNN model was trained for 100 epochs with a learning rate of 0.001 and batch size of 8 using Adam optimizer [12].

3.3 Classification Performance

For the DL-based CNN model, the conventional CT and the enhanced 10-energy VMI data generated by the use of TER were incorporated into our 3D-CNN network, respectively. We calculated the mean values of AUC scores, which are shown in Table III. The results demonstrated that the VMIs data with the TER achieve higher mean AUC values

TABLE III: MEAN AUC VALUES FOR DL-BASED CNN DIAGNOSIS
MODEL

Data Input	Dataset 1	Dataset 2	Dataset 3
Conventional CT (120kVp)	74.20	52.42	59.54
VMIs with TER	**80.71**	**69.33**	**71.97**

than the conventional 120 kVp data, which verifies the effectiveness of the contrast enhancement brought from the VMIs. The AUC values of each dataset can be improved 6.51%, 16.91%&12.55% for lesion characterization, respectively. This is powerful proof that our proposed CADx model could benefit from the energy spectral information in VMIs.

4. DISCUSSION

This study aims at exploring the energy spectral information from a conventional CT image by using the TER to generate VMIs for lesion diagnosis. The proposed CADx framework with the VMIs achieved improved diagnosis performance than the traditional CADx pipeline with the conventional CT data. As we have demonstrated in this work, the VMIs at different energy bins show significant effectiveness for lesion characterization. How to choose the energy range and energy number of the VMIs to capture the meaningful lesion features is still a promising direction [13]. With the analysis from the above, this work will also have a great potential in guiding the energy selection in PCCT imaging for diagnosis. Meanwhile, to compare the PCCT images from the practical measurements with the VMIs by the use of TER from the conventional CT would be a very interesting research topic in the future. And last but not least, clinical evaluations with more disease data sets are needed to test the robustness of the proposed method.

5. CONCLUSIONS

In conclusion, we proposed to use tissue-energy response to generate the virtual monoenergetic images from the conventional CT for CADx of lesions. In this framework, each tissue contrast distribution along the energy axis is fully enhanced, which brings richer information to ML-based CADx. Experimental results demonstrated that the VMIs generated with the use of TER from the conventional CT applied to the ML-based CADx obtain better performance than the traditional CADx pipeline in lesion classification.

REFERENCES

[1] C. H. McCollough, S. Leng, L. Yu, and J. G. Fletcher, "Dual-and multi-energy CT: principles, technical approaches, and clinical applications," Radiology, vol. 276, no. 3, pp. 637-653, 2015.

[2] Y. Gao, Y. Shi, W. Cao, S. Zhang, and Z. Liang, "Energy enhanced tissue texture in spectral computed tomography for lesion classification," Visual Computing for Industry, Biomedicine, and Art, vol. 2, no. 1, pp. 1-12, 2019.

[3] S. Leng et al., "Photon-counting detector CT: system design and clinical applications of an emerging technology," Radiographics, vol. 39, no. 3, pp. 729-743, 2019.

[4] H. Machida et al., "Dual-energy spectral CT: various clinical vascular applications," Radiographics, vol. 36, no. 4, pp. 1215-1232, 2016.

[5] R. Forghani, B. De Man, and R. Gupta, "Dual-energy computed tomography: physical principles, approaches to scanning, usage, and implementation: part 1," Neuroimaging Clinics, vol. 27, no. 3, pp. 371-384, 2017.

[6] S. Chang et al., "Spectrum estimation-guided iterative reconstruction algorithm for dual energy ct," IEEE transactions on medical imaging, vol. 39, no. 1, pp. 246-258, 2019.

[7] A. A. A. Setio et al., "Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks," IEEE transactions on medical imaging, vol. 35, no. 5, pp. 1160-1169, 2016.

[8] H. Jiang, H. Ma, W. Qian, M. Gao, and Y. Li, "An automatic detection system of lung nodule based on multigroup patch-based deep learning network," IEEE journal of biomedical and health informatics, vol. 22, no. 4, pp. 1227-1237, 2017.

[9] H. Wang et al., "A hybrid CNN feature model for pulmonary nodule malignancy risk differentiation," Journal of X-ray Science and Technology, vol. 26, no. 2, pp. 171-187, 2018.

[10] J. Tan et al., "3D-GLCM CNN: A 3-dimensional gray-level Co-occurrence matrix-based CNN model for polyp classification via CT colonography," IEEE transactions on medical imaging, vol. 39, no. 6, pp. 2013-2024, 2019.

[11] S. Zhang et al., "An investigation of CNN models for differentiating malignant from benign lesions using small pathologically proven datasets," Computerized Medical Imaging and Graphics, vol. 77, p. 101645, 2019.

[12] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.

[13] S. Chang et al., "A feasibility study of computer-aided diagnosis with DECT Bayesian reconstruction for polyp classification." Medical Imaging 2022: Computer-Aided Diagnosis. Vol. 12033. SPIE, 2022.