ARTICLE IN PRESS

Engineering xxx (xxxx) xxx

Contents lists available at ScienceDirect

Engineering

journal homepage: www.elsevier.com/locate/eng

Research
Artificial Intelligence—Article

From Signal to Knowledge: The Diagnostic Value of Raw Data in the Artificial Intelligence Prediction of Human Data for the First Time

Bingxi He [a,b,c,†], Yu Guo [d,†], Yongbei Zhu [a,b,c,†], Lixia Tong [e], Boyu Kong [d], Kun Wang [c], Caixia Sun [a,b,c], Hailin Li [a,b,c], Feng Huang [e], Liwei Wu [f], Meng Wang [d], Fanyang Meng [d], Le Dou [d], Kai Sun [c], Tong Tong [c], Zhenyu Liu [c], Ziqi Wei [c], Wei Mu [a,b,c], Shuo Wang [a,b,c], Zhenchao Tang [a,b,c], Shuaitong Zhang [a,b,c], Jingwei Wei [c], Lizhi Shao [c], Mengjie Fang [a,b,c], Juntao Li [g,h], Shouping Zhu [g,h], Lili Zhou [d], Shuo Wang [d], Di Dong [c,i,*], Huimao Zhang [d,*], Jie Tian [a,b,c,*]

[a] Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, School of Engineering Medicine, Beihang University, Beijing 100191, China
[b] Key Laboratory of Big Data-Based Precision Medicine, Beihang University, Ministry of Industry and Information Technology, Beijing 100191, China
[c] CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[d] Department of Radiology, The First Hospital of Jilin University, Changchun 130021, China
[e] Neusoft Medical Systems Co. Ltd., Shenyang 110167, China
[f] Neusoft Research of Intelligent Healthcare Technology Co. Ltd., Shenyang 110167, China
[g] School of Life Science and Technology, Xidian University, Xi'an 710126, China
[h] Engineering Research Center of Molecular and Neuro Imaging, Ministry of Education, Xi'an 710126, China
[i] School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Computed tomography
Diagnosis
Deep learning
Lung cancer
Raw data

ABSTRACT

Encouraging and astonishing developments have recently been achieved in image-based diagnostic technology. Modern medical care and imaging technology are becoming increasingly inseparable. However, the current diagnosis pattern of signal to image to knowledge inevitably leads to information distortion and noise introduction in the procedure of image reconstruction (from signal to image). Artificial intelligence (AI) technologies that can mine knowledge from vast amounts of data offer opportunities to disrupt established workflows. In this prospective study, for the first time, we develop an AI-based signal-to-knowledge diagnostic scheme for lung nodule classification directly from the computed tomography (CT) raw data (the signal). We find that the raw data achieves almost comparable performance with CT, indicating that it is possible to diagnose diseases without reconstructing images. Moreover, the incorporation of raw data through three common convolutional network structures greatly improves the performance of the CT models in all cohorts (with a gain ranging from 0.01 to 0.12), demonstrating that raw data contains diagnostic information that CT does not possess. Our results break new ground and demonstrate the potential for direct signal-to-knowledge domain analysis.

© 2023 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

The discovery of X-rays in 1895 ushered in a new era in the use of imaging for medical diagnostic purposes. Since then, noninvasive medical imaging technology has subverted the traditional palpation and cut-and-see approaches [1]. Astounding technological advances have been made in medical imaging over the past 120 years, and modern medical care is increasingly inseparable from imaging technology. Medical imaging is essential for clinicians to observe the patient's body from images and thereby diagnose diseases. This process can be defined as a path from image to knowledge. However, it has recently been found that clinicians' experience has become a bottleneck in this path, hindering the accurate diagnosis and treatment of diseases [2,3].

* Corresponding authors at: CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (D. Dong). Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, School of Engineering Medicine, Beihang University, Beijing 100191, China (J. Tian).

E-mail addresses: di.dong@ia.ac.cn (D. Dong), huimao@jlu.edu.cn (H. Zhang), jie.tian@ia.ac.cn (J. Tian).

† These authors contributed equally to this work.

https://doi.org/10.1016/j.eng.2023.02.013
2095-8099/© 2023 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Please cite this article as: B. He, Y. Guo, Y. Zhu et al., From Signal to Knowledge: The Diagnostic Value of Raw Data in the Artificial Intelligence Prediction of Human Data for the First Time, Engineering, https://doi.org/10.1016/j.eng.2023.02.013

The emergence of artificial intelligence (AI) technology has partially solved the problem of humans' limited ability in the diagnosis process [4–7]. AI can be used to automatically mine the radiographic patterns in imaging data related to the occurrence and progression of diseases, and it has been shown to match and even surpass human abilities in many clinical applications [8–12]. The essential reason why AI can surpass humans may be that it treats images as data rather than as visual images and extracts huge quantities of features for analysis [13,14]. In comparison, a medical image is compressed or its data is filtered to fit the human eye, which may be insufficient or imperfect for diagnosis. Taking computed tomography (CT) as an example, the CT system first collects raw data (i.e., the signal) from the patient; then, a reconstruction method converts the raw data to images (i.e., signal to image) [15]. Therefore, both AI-based and human-based diagnosis are signal-to-image-to-knowledge processes. Medical images suffer from information distortion in both the acquisition and reconstruction processes. The current high sampling frequency greatly compresses the influence of factors such as motion artifacts in the acquisition process, so the main reason for the loss of resolution lies in operations such as interpolation and suboptimal statistical weighting in the reconstruction process [16]. In fact, the unprocessed data size of raw data is about 10–20 times larger than the size of the final CT images (e.g., 2 GB compared with 180 MB). In general, we can imagine such processes as following a chain from signals to images to knowledge (as shown in Fig. 1). In the signal domain, complete diagnostic knowledge is obtained. However, during the process from the signal domain to the image domain, the introduction of reconstruction errors leads to a loss of diagnostic knowledge, which is difficult to recover. Although AI-based models have been designed to continuously reduce their impact, losses are inevitable due to the nature of solving inverse problems [17]. In this way, no additional information is added in this process, but some information is lost. Hence, the huge amount of information present in the raw data is not optimally mined in current signal-to-image-to-knowledge processes, and there is great scientific interest in the question of how to analyze raw data.

Skipping the image process and going directly from signal to knowledge will hopefully lead to new breakthroughs in disease diagnosis [17]. This idea can be traced back to 2016, when Wang et al. [18] proposed a direct pathway from data acquisition to therapeutic actions. Inspired by this idea, several studies have discussed the potential value of analyzing raw data [19–23] and thereby going directly from signal to knowledge. De Man et al. [21] conducted a simulation experiment to detect and estimate the vessel centerline from raw data in the signal domain. They achieved encouraging initial results that demonstrated the feasibility of raw data analysis for clinical CT analysis tasks. Through the simulation of raw data with CT images of pulmonary nodules, Gao et al. [23] found that pulmonary nodules could be detected in the signal domain in reality and obtained exciting results showing that this method could effectively solve the different-nodule-size problem encountered in common image-based diagnosis. Wu et al. [22] also demonstrated that an end-to-end network prediction scheme from the signal domain was superior in terms of both sensitivity and accuracy to an abnormality detection model developed with reconstructed images. We have reported our simulation results on lung cancer at the American Association for Cancer Research (AACR) conference [24]. However, most current studies are based on simulation; thus, there is no published research on signal-to-knowledge analysis in real clinical tasks based on patient data, as far as we know.

In this prospective study, for the first time, we developed an AI-based signal-to-knowledge diagnostic scheme for lung nodule classification directly from CT raw data (a flowchart of this process is shown in Fig. 1). The value of raw data alone and its added value to CT are studied for 276 patients. We found that the raw data achieved almost comparable performance with CT, indicating that it is possible to diagnose diseases without reconstructed images. Moreover, the introduction of raw data greatly promoted the performance of CT, demonstrating that raw data contains diagnostic information that CT does not possess. This research breaks the routinely used circle of image-based diagnosis and may open up a new signal-to-knowledge pathway for disease diagnosis.

2. Material and methods

2.1. Patients

In this prospective study, 626 patients who had a chest CT scan in the First Hospital of Jilin University from November 2019 to May 2021 were recruited. Eligible patients were included according to the following inclusion criteria: ① patients who had a pulmonary lesion larger than 2 cm with a contrast-enhanced chest CT scan, ② raw data obtained from a CT machine after the imaging examination, and ③ a pathological diagnosis of pulmonary lesion within a two-week interval from the CT scan. Patients were excluded based on the following: ① previous systemic antineoplastic treatments or ② CT images with poor image quality or an unreadable scan. After exclusion, a total of 276 patients were included in the modeling experiments.

The experiments were performed in accordance with the Standards for Reporting Diagnostic Accuracy Studies (STARD) and approved by the Ethics Committee of the First Hospital of Jilin University (AF-IRB-032-05).

2.2. Collection of CT images and raw data

Both the CT images and the raw data were collected from the First Hospital of Jilin University and were acquired using a NeuViz Prime CT system (Neusoft Medical Systems Co., Ltd., China). The system parameters of the CT scanner included a source-to-isocenter distance of 570 mm, a source-to-detector distance of 1040 mm, and a scanning field of view (FOV) of 500 mm. The imaging protocol included a contrast-enhanced CT of the chest with variable imaging parameters. Contrast-enhanced CT scans were performed in spiral scan mode using a 324 mA tube current, 100 kVp tube voltage, 0.5 s ration time, and 0.9 spiral pitch. CT raw data were reconstructed using a kernel F20 at a slice thickness of 1.0 mm with an image pixel range from 0.59 mm to 0.98 mm and an image matrix of 512 by 512. In addition, we acquired the initial height and initial view angle of the CT detector each time the patient underwent a scan. Finally, CT images and raw data from each scanner were randomly stratified into one of three cohorts in a 6:2:2 ratio: a training cohort, a validation cohort, and a test cohort. Table S1 in Appendix A provides the CT scanner information, system parameters, and imaging parameters.

2.3. Lesion segmentations in CT images

The primary lesion segmentations were manually delineated across all the sections in the axial view using the annotation tool in IntelliSpace Discovery (ISD, Philips, Germany). The regions of interest were annotated and reviewed by four radiologists with 8–25 years' experience with chest CT. All radiologists were blinded to any clinical or histopathologic information. The annotation was labeled into five common categories according to the lesions' pulmonary lobe.

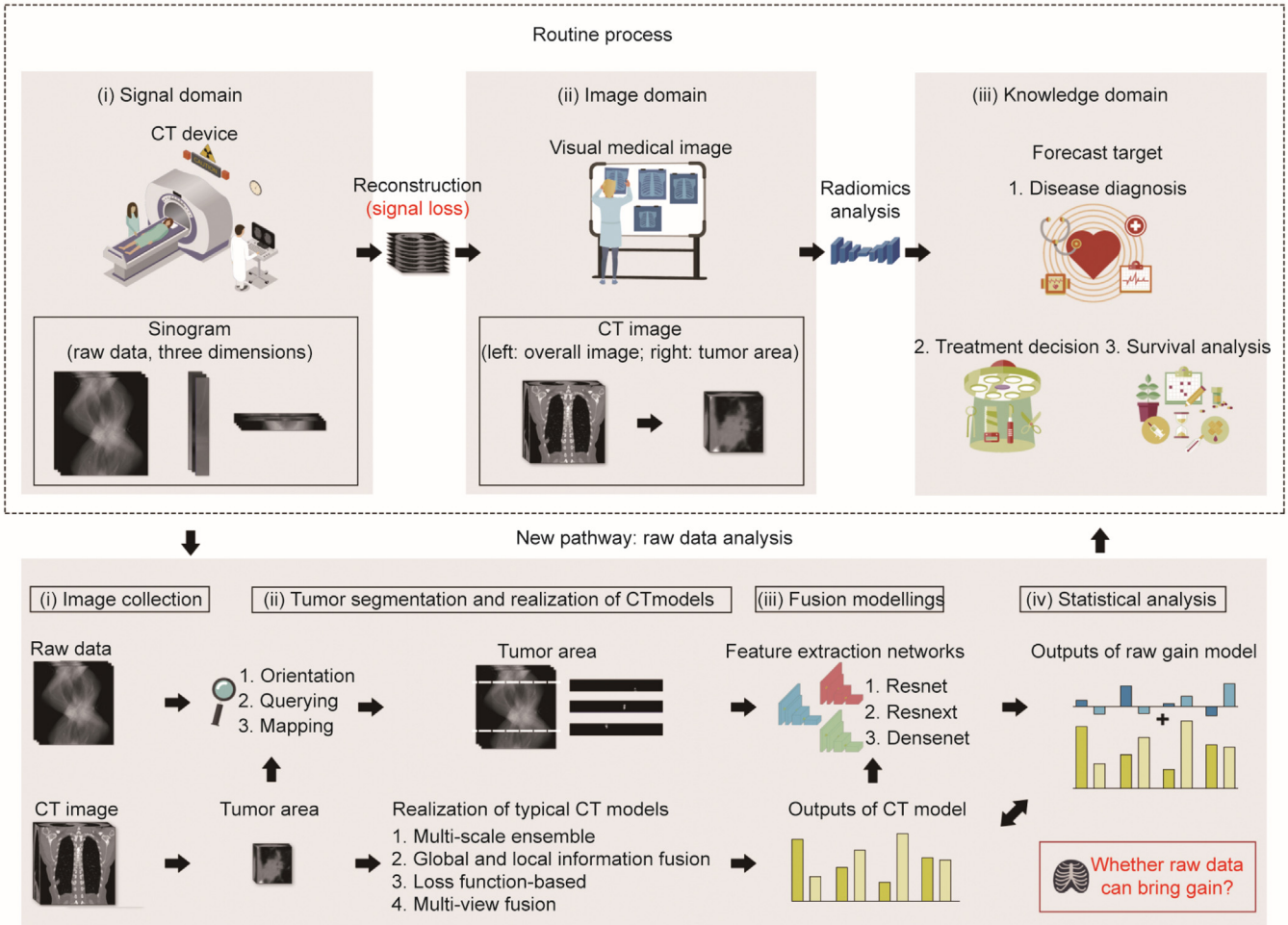

Fig. 1. Flow chart of a raw data gain experiment.

2.4. Realization of typical CT models

Many studies have been conducted on benign–malignant lung nodule classification in chest CT. We selected four typical papers from the major journals *IEEE Transactions on Medical Imaging*, *Medical Image Analysis*, and *Nature Medicine* to construct a CT model (CTM) using the multi-scale ensemble method (CTM 1) [25], the global and local information fusion method (CTM 2) [26], the loss function-based method (CTM 3) [27], and the multi-view fusion (CTM 4) method, respectively [28]. We further performed experiments on these four typical models with our dataset. All realization details are described in Appendix A.

2.5. Extraction of the lesion region from raw data

After acquiring the four CTMs, we proceeded to perform raw data gain experiments. The first step of the experiment was to select a projection surface containing the lesions in the raw data. The raw data of CT scans has three dimensions, which consists of 1D scan index and 2D projection data. Specifically, the scan index represents the acquisition order, and the projection data represents the detector receives the X-ray attenuation, for which the channel and row directions are defined as x and y, respectively. All lesion segmentation regions of the raw data were derived from the binarized segmentation of the CT image after being represented in a unified coordinate system. The complete derivation can be condensed into three steps: orientation, querying, and mapping (Fig. S1 in Appendix A).

2.5.1. Orientation

For the derivation of orientation, we took the segmented regions in the CT image as the research object. The orientation calculation mainly includes cross-sectional orientation and height orientation. For cross-sectional positioning, we set the center point between the CT source and detector as the coordinate origin (which is also the rotation center of the CT gantry), parallel to the cross-section of the CT image. Next, the motion trajectory was characterized by the scan index t, the rotation radius r, and the angle θ. In order to obtain these parameters, we first read the origin coordinates (x_{origin}, y_{origin}, and z_{origin}) and calculated the offset values (x_{offset} and y_{offset}) through voxel spacing and image size (x_{size} and y_{size}).

$$x_{\text{offset}} = x_{\text{origin}} + x_{\text{spacing}} \times \tfrac{1}{2} x_{\text{size}} \tag{1}$$

$$y_{\text{offset}} = y_{\text{origin}} + y_{\text{spacing}} \times \tfrac{1}{2} y_{\text{size}} \tag{2}$$

Next, with the help of the offset values and the voxel coordinates x_{CT} and y_{CT} in the CT image, the distance from the coordinate origin (x_{length} and y_{length}) was calculated as follows:

$$x_{\text{length}} = x_{\text{spacing}} \times \left(x_{\text{CT}} - \tfrac{1}{2} x_{\text{size}}\right) - x_{\text{offset}} \tag{3}$$

$$y_{\text{length}} = y_{\text{spacing}} \times \left(y_{\text{CT}} - \tfrac{1}{2} y_{\text{size}}\right) - y_{\text{offset}} \tag{4}$$

In the case of obtaining the above variable, the rotation radius r, and the starting angle θ_0 were obtained.

$$r = \sqrt{x_{\text{length}}{}^2 + y_{\text{length}}{}^2} \tag{5}$$

$$\theta_0 = \arctan \frac{y_{\text{length}}}{x_{\text{length}}} \tag{6}$$

By introducing the scanning period of the machine, we obtained the angle change $\Delta\theta$ with the following relationship:

$$y_{\text{length}} = r \times \cos\left(\Delta\theta \times t + \theta_0\right) \tag{7}$$

$$x_{\text{length}} = r \times \sin\left(\Delta\theta \times t + \theta_0\right) \tag{8}$$

For height positioning, we directly obtained the initial height H_{start} of the voxel through the coordinate z, the slice thickness, and the origin of the voxel point in the CT images.

2.5.2. Querying

Our purpose in this step was to determine the interval of the scan index t in which the tumor voxels appeared in the raw data. Since there is a cone beam in the projection, we first calculated the change function H_{area} of the voxels.

$$d = \tfrac{1}{2} l + x_{\text{length}} \tag{9}$$

where d is the distance from the voxel to the X-ray focal spot on the x-axis, and l is the distance from the focal spot to the detector.

$$H_{\text{area}} = \tfrac{1}{l}\left(\tfrac{1}{2} n_y \times \Delta n_y \times d\right) \tag{10}$$

where n_y is the number of detector rows and Δn_y is the channel spacing along the y-axis.

Next, we determined the scan index t range of the voxels in the raw data by means of the following inequality.

$$H_0 - \Delta H \times t \le H_{\text{start}} + H_{\text{area}} \tag{11}$$

where H_0 is the initial height at which the detector starts to scan and ΔH is the height change in a scan.

$$H_0 - \Delta H \times t \ge H_{\text{start}} - H_{\text{area}} \tag{12}$$

To reduce the computational complexity, we first extracted the highest and lowest masks in the segmentation images and calculated the start and end indices of two voxels. Then, we initially located the range of index dimensions. Within this interval, we computed the mapping result of the voxels within the layer.

2.5.3. Mapping

Through the above calculation, we obtained the scan index interval corresponding to the voxels; then, the voxel appearing in the scan index was obtained by calculating the projection data of the scan index layer by layer. The coordinates of each voxel on the projection surface were defined as x_{raw} and y_{raw}, respectively. Since y_{raw} is related to the height H_t, H_{area} at the scan index t, and the number of detector rows n_y in the detector, we determined its height difference relative to the detector by means of the following formula, and then calculated its coordinates in the projection.

$$y_{\text{raw}} = \frac{H_t}{H_{\text{area}}} \times \frac{\left(n_y - 1\right)}{2} \tag{13}$$

Since the x-axis of the projection plane is equiangularly sampled, x_{raw} can be acquired through the angle at the t scan index θ_t, the view angle θ_d of the detector, and the number of channels in the detector n_x.

$$x_{\text{raw}} = \frac{2\theta_t}{\theta_d} \times \frac{(n_x - 1)}{2} \tag{14}$$

After obtaining the segmentation files of the lesions in the raw data, we saved the raw data segment through the initial scan index interval and used this as the training data for this gain experiment. It should be noted that there are different directions in the actual retrieval of raw data (i.e., from head to foot or from foot to head). We used the same spatial relationship to modify the inequality for different directions and then located the lesion.

2.6. Construction of a raw data gain model based on CT images

To explore whether the raw data contained unique information, we built residual fusion models through the raw data and fused it with the CTMs' output to determine whether the raw data could improve the output. First, we built three feature extraction networks using the raw data. Based on the memory needed for the calculation, we sampled the index dimension of the raw data fragments containing lesions to one-eighth, and resampled the same size based on the average value by means of equal interval sampling. For the channel, we directly removed the data outside the reconstruction area from both sides and resampled with the row dimension at half the size. In the model building, we did not modify Densenet121 (DN) [29], Resnet18 (RE) [30], or Resnext18 (RX) [31] in three dimensions in order to direct the direct gain of the raw data as much as possible. The training settings and parameters are detailed in Appendix A.

The primary aim of the residual fusion model was to achieve a correction of the CTM output; the origin of this idea was that the learning residual is easier, which is mentioned in Resnet. The probabilities of the CTMs predicting the patients as being positive were fused with the predicted probabilities of the raw data models (RDM), which was performed during the training process. The name raw gain model (RGM) was given to the fusion model. All models have two output nodes, positive and negative, which were calculated separately. More specifically, the probability of predicting one patient as being positive was calculated as follows:

$$\text{RGM}_{\text{positive}} = \text{RDM}_{\text{positive}} + \text{CTM}_{\text{positive}} \tag{15}$$

The probability of predicting one patient as being negative was calculated as follows:

$$\text{RGM}_{\text{negative}} = \text{RDM}_{\text{negative}} + \left(1 - \text{CTM}_{\text{positive}}\right) \tag{16}$$

After the output fusion of the CTM and the raw data mode, the Softmax function was used to keep the above two output sums equal to 1. In addition, the loss function was used to calculate the loss and optimize the model. The three feature extraction networks built with raw data were fused with each of the four representative CTMs described above to obtain four RGMs, for a total of 12 RGMs: RGB-DN (RGM-DN 1/2/3/4), RGM-RE (RGM-RE 1/2/3/4), and RGB-RX (RGM-RX 1/2/3/4). Next, the RGMs were compared with the CTMs to evaluate the benefits of the raw data.

2.7. Calculation of the average attention score

To calculate the average attention score of each voxel, we first used the segmentation data of the lesion in the raw data to obtain the non-lesion area by means of the unary complement. Next, we dotted and summed the segmentation data of the lesion and non-lesion areas with the attention matrix. Finally, the average attention score was obtained by dividing the total amount of attention in these two areas by the number of voxels in the segmented regions, respectively. It should be noted that we also normalized the average attention score of the lesion area and the non-lesion area in each piece of raw data to obtain a more intuitive comparison result.

3. Result

3.1. Clinical characteristics

The clinical characteristics are summarized in Table S2 in Appendix A. A total of 276 patients were included in this study, with 166 patients in the training cohort, 55 in the validation cohort, and 55 in the test cohort. Of the patients, 50% (n = 138) were female, and the mean age in the entire dataset was 58.48 years. Among the included patients, there were 21 (8%) cases of small cell carcinoma, 35 (13%) squamous cell carcinomas, and 149 (54%) adenocarcinomas. Among all the patients, the lesion location in most patients was identified as the right upper lobe (n = 89, 32%), followed by the left lower lobe (n = 67, 24%), and the left upper lobe (n = 64, 23%). For lung cancer diagnosis, most patients (n = 225, 82%) were evaluated as showing malignant cancer.

3.2. Performance of CT model and raw data gain model

This experiment explores the performance improvement shown by the residual fusion model based on both raw data and CT images, in comparison with the model based on CT images only. To further explore the repeatability and stability of this gain, we tested four different CT models (abbreviated as CTM 1–CTM 4) and adopted three backbone network architectures for raw data feature extraction, as follows: DN [29], RE [30], and RX [31]. For each CTM, three raw data gain models based on different backbone feature extraction networks were constructed. The performance of each RGM was compared with that of the original CTM. The

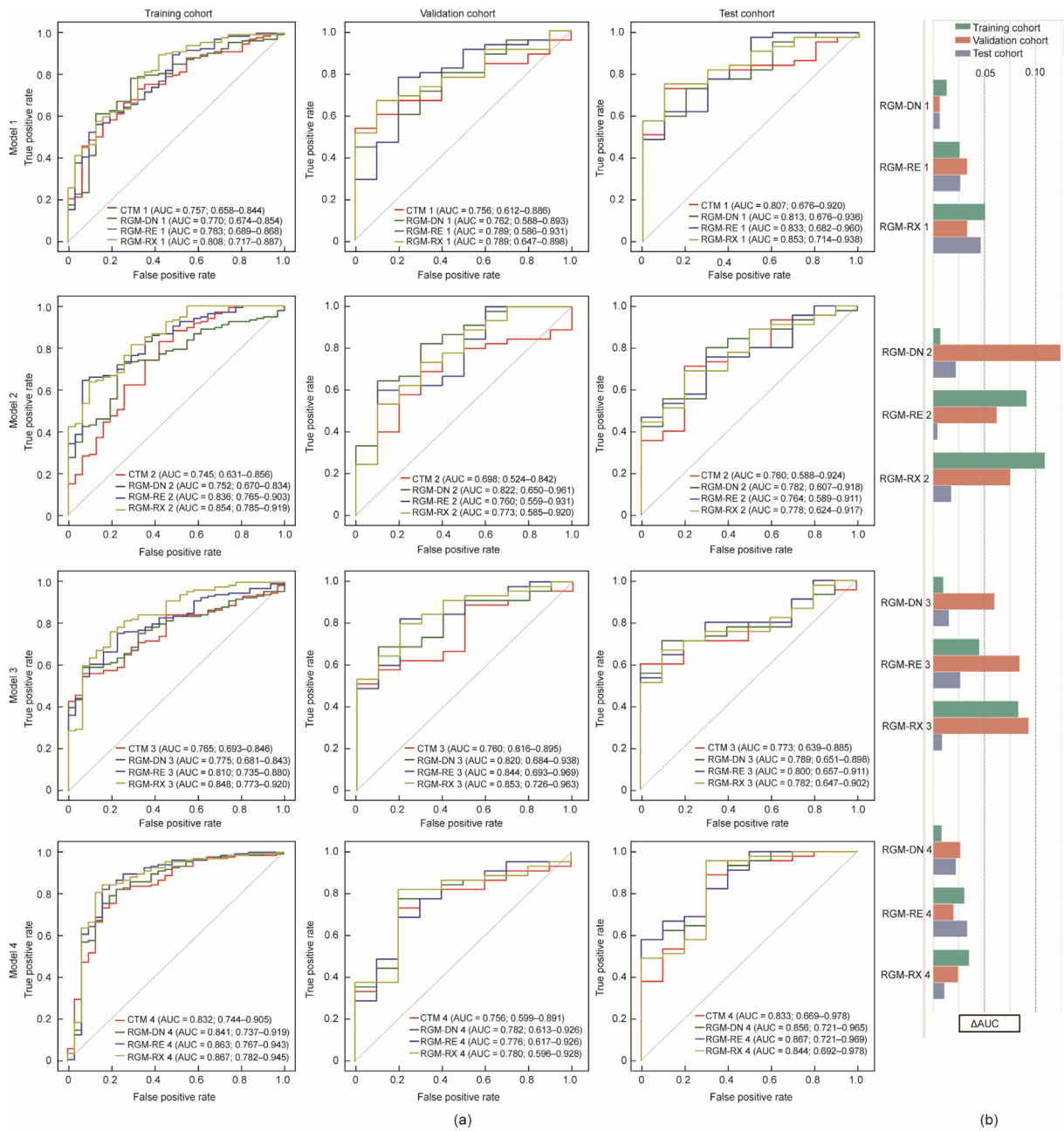

Fig. 2. CT model and raw data gain results. (a) ROC curves of each CTM and residual fusion model based on different backbone feature extraction networks; (b) bar graph of reaction AUC gain.

Table 1
Detailed result statistics of raw data for CTMs.

Model	True positive proportion	False negative proportion	False positive proportion	True negative proportion	True positive rate	True negative rate
CTM 1	0.616	0.199	0.069	0.116	0.756	0.627
RGM-DN 1	0.699	0.116	0.091	0.094	0.858	0.508
RGM-RE 1	0.543	0.272	0.043	0.141	0.666	0.766
RGM-RX 1	0.761	0.054	0.120	0.065	0.934	0.351
CTM 2	0.678	0.138	0.101	0.083	0.831	0.451
RGM-DN 2	0.656	0.159	0.076	0.109	0.805	0.589
RGM-RE 2	0.594	0.221	0.051	0.134	0.729	0.724
RGM-RX 2	0.605	0.210	0.058	0.127	0.742	0.686
CTM 3	0.460	0.355	0.022	0.163	0.564	0.881
RGM-DN 3	0.529	0.286	0.025	0.159	0.649	0.864
RGM-RE 3	0.656	0.159	0.072	0.112	0.805	0.609
RGM-RX 3	0.685	0.130	0.080	0.105	0.840	0.568
CTM 4	0.634	0.181	0.043	0.141	0.778	0.766
RGM-DN 4	0.652	0.163	0.047	0.138	0.800	0.746
RGM-RE 4	0.681	0.134	0.051	0.134	0.836	0.724
RGM-RX 4	0.696	0.120	0.043	0.141	0.853	0.766

receiver operating characteristic (ROC) curves and the area under the curve (AUC) of the four CTMs and the corresponding RGMs based on different backbone networks are shown in Fig. 2.

For each CTM, the residual fusion models based on different backbone networks obtained a better classification performance on the training, validation, and test cohorts. For the CTM 1 model, the fusion model that produced the maximum performance improvement for the training cohort was RGM-RX 1, with an AUC improvement of 0.051 (from 0.757 to 0.808). The fusion models that produced the maximum performance improvement for the validation cohort were RGM-RE 1 and RGM-RX 1, with an AUC improvement of 0.033 (from 0.756 to 0.789). The fusion model that produced the maximum performance improvement for the test cohort was RGM-RE 1, with an AUC improvement of 0.046 (from 0.807 to 0.853).

For the CTM 2 model, the fusion model that produced the maximum performance improvement for the training cohort was RGM-RX 2, with an AUC improvement as high as 0.109 (from 0.745 to 0.854). The fusion model that produced the maximum performance improvement for the validation cohort was RGM-DM 2, with an AUC improvement as high as 0.124 (from 0.698 to 0.822). The fusion model that produced the maximum performance improvement for the test cohort was RGM-DM 2, with an AUC improvement as high as 0.022 (from 0.760 to 0.782).

For the CTM 3 model, the fusion model that produced the maximum performance improvement for the training cohort was RGM-RX 3, with an AUC improvement as high as 0.083 (from 0.765 to 0.848). The fusion model that produced the maximum performance improvement for the validation cohort was RGM-RX 3, with an AUC improvement as high as 0.093 (from 0.760 to 0.853). The fusion model that produced the maximum performance improvement for the test cohort was RGM-RE 3, with an AUC improvement as high as 0.027 (from 0.773 to 0.800).

For the CTM 4 model, the fusion model that produced the maximum performance improvement for the training cohort was RGM-RX 4, with an AUC improvement as high as 0.035 (from 0.832 to 0.867). The fusion model that produced the maximum performance improvement for the validation cohort was RGM-DM 4, with an AUC improvement as high as 0.026 (from 0.756 to 0.782). The fusion model that produced the maximum performance improvement for the test cohort was RGM-RE 4, with an AUC improvement as high as 0.034 (from 0.833 to 0.867). Overall, using RX as the backbone network for the raw data feature extraction resulted in the maximum average performance improvement on the three cohorts.

3.3. Image feature distribution of the RGMs and gain stability analysis

We performed a *t*-distributed stochastic neighbor embedding (*t*-SNE) dimensionality reduction on all deep learning features obtained by the different feature extraction networks and counted the true positives, false positives, true negatives, and false negatives for each patient (Table 1). In addition, we assigned different colors and markers to visualize these metrics in the same coordinate system (Fig. 3). As shown in Fig. 3, the results of the various RGMs within each CTM are relatively similar, even though they come from different feature extraction networks. Table 1 shows the same situation. The gain of the RGMs inside each CTM is approaching the same trend, such as improving the malignant or benign detectable rates. Moreover, RGM-RE 1, RGM-DN 2, RGM-RE 2, RGM-DN 3, RGM-DN 4, RGM-RE 4, and RGM-RX 4 achieve a significant increase in the detectable rate of one category at the expense of a small number of the other category detectable rates.

Therefore, we calculated the optimization rate and error rate of each RGM for the CTM; we also calculated the proportion of at least two model optimizations to all optimization samples, which can reflect the stability of the raw data's gain. All results are summarized in Table S3 in Appendix A. The results show that the analysis method incorporating the raw data has a high optimization rate for CTMs 1–3 that is greater than the error rate, which is also reflected in the improvement of the AUC. In addition, although different feature extraction networks were used to analyze the raw data, the proportion of at least two networks that can be optimized in each CTM is about 80%. Finally, we found that seven samples were mispredicted within four CTMs. Of these seven samples, the input of raw data corrected the prediction results of six CTMs, and a corrected model existed in each CTM. In summary, the gain of the raw data for the CTM is very stable.

3.4. Visual statistics and analysis of the RGMs

To better explain the prediction process of the RGMs, we visualized the region of greatest interest in the RGM using gradient-weighted class activation mapping (Grad-CAM). The predictive results of the RGMs were most dependent on the information of the RGM-discovered suspicious areas. Fig. 4 illustrates the lesion masks and corresponding attention maps from different views of the raw data. From Fig. 4, it can be seen that the RGMs can always focus on the lesion areas for prediction, although the input data includes some non-lesion areas. We also calculated the average attention score of each voxel in the lesion and non-lesion areas

in the raw data; the result showed that the attention score of the lesion area was 1–2 times greater than that of the non-lesion area.

3.5. Stratified analysis of different malignant subgroups

The results of the subgroup analysis for age, sex, and lesion size are shown in Table 2 and in Table S4 in Appendix A.

In the subgroup of patients aged 60 or lower, RGM-RX 4 and CTM 4 achieved a similar highest model performance, with AUCs of 0.837 (0.746–0.924) and 0.831 (0.749–0.904), respectively. In the over-60 subgroup, RGM-RX 4 achieved the highest model performance, with an AUC of 0.845 (0.713–0.949), outperforming the best CTM (CTM 4 with an AUC of 0.790). In the male subgroup, CTM 4 and RGM-RX 4 performed best, with similar AUCs of 0.804

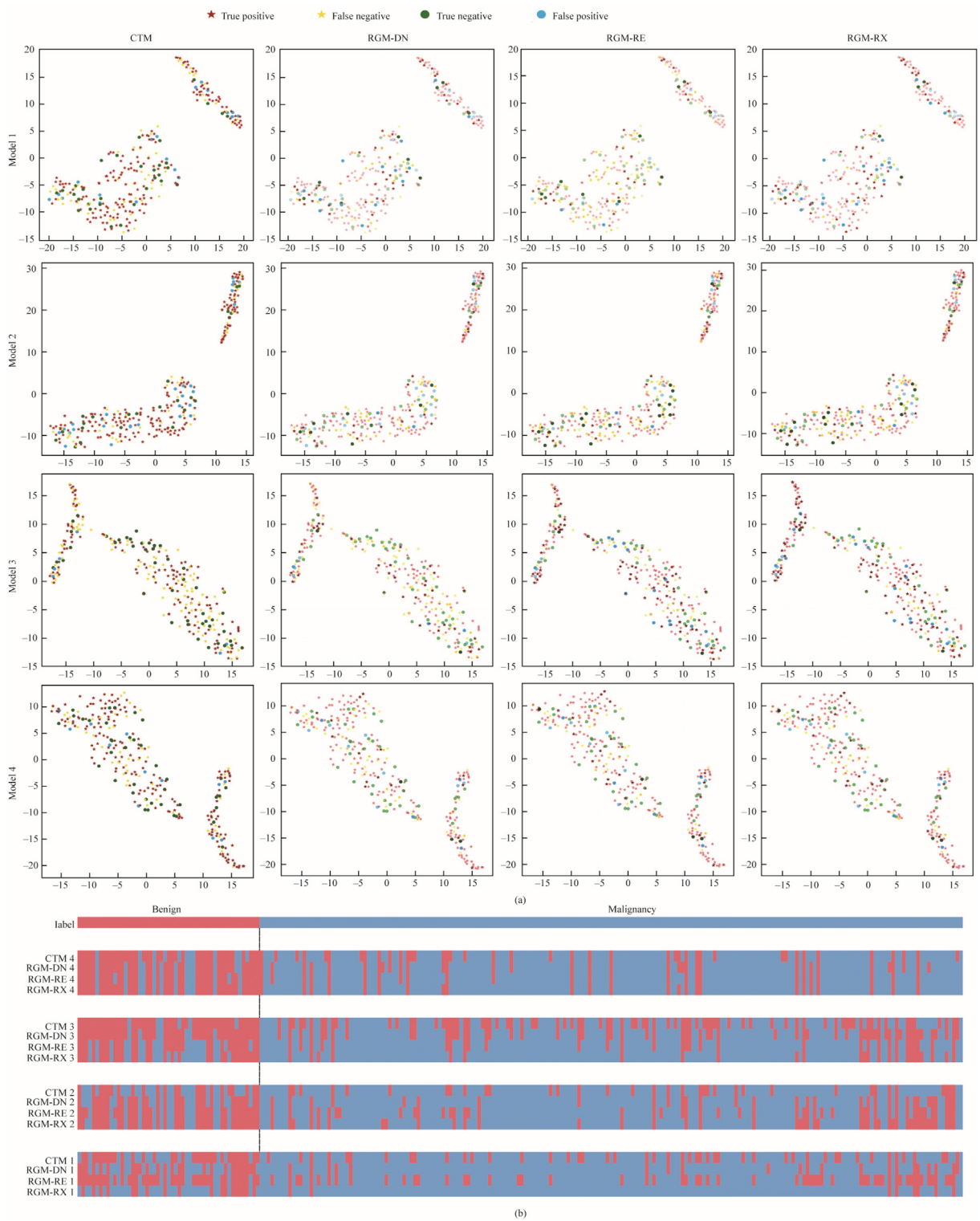

Fig. 3. Feature distribution of the RGMs and a heatmap of prediction results. (a) Feature distribution in the RGM, where different marks and colors reflect different authenticity and prediction categories; (b) prediction results of all real categories, CTMs, and RGMs. In this figure, red and blue represent benign and malignant, respectively. The row named "Label" represents the true category of patients.

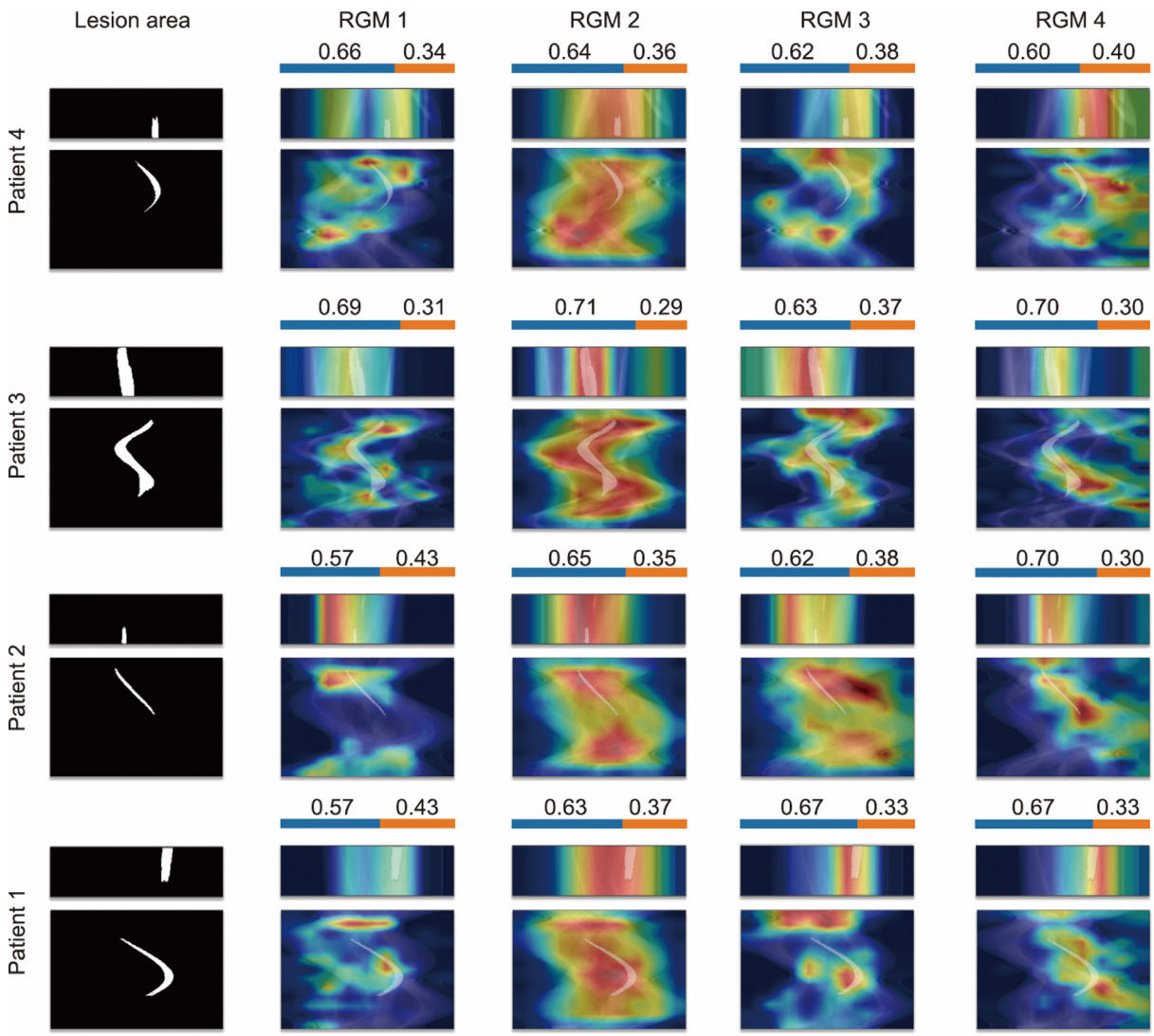

Fig. 4. Lesion trajectory in raw data and the Grad-CAM graphs of the RGMs. Lesion area reflected the segmentation of lesions in the raw data. The blue and red in the class activation mapping represent low- and high-attention areas of the model, respectively, and all attention values were normalized to 0 to 1 in the same data.

(0.706–0.882) and 0.810 (0.707–0.897), respectively. In the female subgroup, RGM-RX 3 achieved the highest performance, with an AUC of 0.885 (0.818–0.945), far exceeding the best CTM (CTM 4 with an AUC of 0.823 (0.720–0.920)). In the subgroup with a lesion size less than or equal to 23 mm, RGM-RX 3 achieved the highest model performance, with an AUC of 0.847 (0.781–0.916), far exceeding the best CTM (CTM 4 with an AUC of 0.806). In the subgroup with a lesion size greater than 23 mm, CTM 4 and RGM-RE 4 showed a similar highest model performance, with AUCs of 0.819 (0.719–0.906) and 0.833 (0.703–0.925), respectively. For the lesion location subgroups, CTM 4 and RGM-RX 4 showed a similar performance in the subgroup of superior lobe of left lung, while the RGMs outperformed the CTMs, with an AUC of 0.840 versus 0.812 in the subgroup of inferior lobe of left lung, 0.849 versus 0.807 in the subgroup of superior lobe of right lung, and 0.872 versus 0.843 in the subgroup of inferior lobe of right lung.

4. Discussion

In this prospective study, we validated the potential value of raw data in real clinical practice for the first time. Interestingly, the raw data analysis showed a performance comparable with CT images, which indicates that leveraging non-image information holds promise as an alternative to image-based methods. Moreover, this study confirmed the value of adding raw data to CT images, indicating that the combination of non-image and image data will further promote the advance of disease diagnosis. This study proposed and validated a feasible method for diagnosis without image reconstruction, which has the potential to change existing imaging-based diagnosis and treatment strategies.

The classification of benign and malignant pulmonary nodules is a matter of great clinical concern [25,32,33]. This study explored the feasibility of raw data analysis in classifying indeterminate lung nodules greater than 2 cm in size. The results indicate that raw data can well discriminate malignant nodules from benign nodules. The AUCs of the raw data in the training cohort, validation cohort, and test cohort were 0.768 (95%CI: 0.681–0.851), 0.760 (95%CI: 0.558–0.922), and 0.782 (95%CI: 0.592–0.924), respectively (Fig. S2 in Appendix A), and there was no statistical difference between the performance of the raw data and that of CT. These results indicate that the classification of lung nodules may not need image reconstruction and clinician participation. However, it is still uncertain whether a convolutional neural network is the

Table 2
The performance of 12RGMs in the subgroup analysis.

Clinical characteristics	The AUCs (95%CI) of all RGMs											
	RGM-DN 1	RGM-RE 1	RGM-RX 1	RGM-DN 2	RGM-RE 2	RGM-RX 2	RGM-DN 3	RGM-RE 3	RGM-RX 3	RGM-DN 4	RGM-RE 4	RGM-RX 4
Age												
≤60	0.731 (0.647–0.825)	0.760 (0.667–0.844)	0.756 (0.667–0.837)	0.764 (0.668–0.845)	0.790 (0.709–0.874)	0.786 (0.705–0.867)	0.800 (0.719–0.870)	0.825 (0.743–0.895)	0.824 (0.750–0.894)	0.829 (0.749–0.903)	0.836 (0.746–0.915)	0.837 (0.746–0.924)
>60	0.801 (0.680–0.908)	0.781 (0.664–0.894)	0.804 (0.700–0.900)	0.740 (0.610–0.859)	0.792 (0.678–0.901)	0.815 (0.699–0.932)	0.744 (0.636–0.833)	0.767 (0.648–0.863)	0.799 (0.673–0.914)	0.808 (0.661–0.926)	0.835 (0.690–0.945)	0.845 (0.713–0.949)
Sex												
Male	0.693 (0.584–0.794)	0.698 (0.591–0.793)	0.697 (0.601–0.792)	0.737 (0.626–0.826)	0.725 (0.608–0.837)	0.714 (0.600–0.824)	0.702 (0.605–0.791)	0.736 (0.634–0.834)	0.717 (0.617–0.822)	0.793 (0.695–0.877)	0.791 (0.686–0.889)	0.810 (0.707–0.897)
Female	0.808 (0.716–0.898)	0.825 (0.728–0.901)	0.835 (0.749–0.901)	0.786 (0.695–0.871)	0.842 (0.749–0.916)	0.859 (0.777–0.927)	0.852 (0.780–0.921)	0.863 (0.784–0.920)	0.885 (0.818–0.945)	0.849 (0.741–0.944)	0.873 (0.764–0.958)	0.863 (0.739–0.954)
Tumor size												
≤23 mm	0.771 (0.677–0.849)	0.786 (0.689–0.869)	0.815 (0.732–0.893)	0.775 (0.685–0.863)	0.821 (0.743–0.894)	0.828 (0.752–0.901)	0.800 (0.722–0.869)	0.823 (0.736–0.888)	0.847 (0.781–0.916)	0.821 (0.729–0.906)	0.836 (0.744–0.925)	0.839 (0.731–0.926)
>23 mm	0.737 (0.603–0.847)	0.746 (0.627–0.854)	0.709 (0.582–0.815)	0.748 (0.639–0.847)	0.753 (0.634–0.864)	0.754 (0.641–0.859)	0.759 (0.657–0.850)	0.782 (0.684–0.874)	0.775 (0.666–0.868)	0.822 (0.712–0.924)	0.833 (0.703–0.925)	0.831 (0.729–0.928)
Tumor location												
Superior left	0.679 (0.520–0.836)	0.667 (0.464–0.836)	0.743 (0.616–0.881)	0.638 (0.469–0.812)	0.683 (0.480–0.847)	0.737 (0.517–0.899)	0.824 (0.702–0.925)	0.812 (0.701–0.916)	0.810 (0.680–0.914)	0.861 (0.754–0.953)	0.885 (0.773–0.964)	0.909 (0.829–0.979)
Inferior left	0.796 (0.638–0.915)	0.789 (0.652–0.921)	0.769 (0.600–0.896)	0.741 (0.584–0.881)	0.759 (0.607–0.905)	0.771 (0.599–0.911)	0.775 (0.633–0.894)	0.793 (0.656–0.907)	0.812 (0.650–0.934)	0.813 (0.651–0.953)	0.840 (0.688–0.969)	0.833 (0.661–0.961)
Superior right	0.739 (0.618–0.855)	0.814 (0.698–0.912)	0.785 (0.673–0.876)	0.810 (0.704–0.900)	0.843 (0.740–0.930)	0.811 (0.685–0.901)	0.756 (0.633–0.864)	0.802 (0.675–0.909)	0.815 (0.675–0.899)	0.843 (0.730–0.936)	0.848 (0.741–0.934)	0.849 (0.729–0.939)
Inferior right	0.821 (0.638–0.957)	0.809 (0.655–0.957)	0.795 (0.617–0.937)	0.872 (0.744–0.970)	0.849 (0.701–0.970)	0.849 (0.700–0.970)	0.786 (0.617–0.923)	0.835 (0.705–0.961)	0.812 (0.616–0.959)	0.721 (0.430–0.942)	0.764 (0.470–0.981)	0.775 (0.524–0.962)

CI: confidence interval.

most suitable method for raw data analysis. The scanning mode and spatial structure of the raw data require a unique network structure that fits its characteristics, which is the main goal of our follow-up research. A future RDM could be applied to a wide range of grassroots hospitals that have mainstream CT systems but lack technical personnel and clinicians.

Our study shows that the introduction of raw data to CT resulted in an overall improvement over different CTMs, regardless of which backbone network was used. This indicates that raw data contains unique information that may be lost during reconstruction processing, which aligns with the results of a study by Gao et al. [23] showing that the fusion of raw data with image data resulted in an improvement in the model's result. Moreover, the stability of the RGMs was increased on the training cohort, validation cohort, and test cohort compared with the CTMs. The combination of both non-image and image data made the model robust, which was reflected in the fact that each convolutional network indicated optimization.

The result statistics showed that RX had a better performance improvement than the other networks, which may be related to factors such as the network structure and number of parameters. By introducing group convolution, RX enables the model to learn more multiple feature representations [31], which is similar with the multi-head attention mechanism in Transformer. Nevertheless, the structural advantages of the network for analyzing raw data will need to be evaluated on more data in the future.

Our study used Grad-CAM to visualize the model decision; the results showed that the model could self-adapt to converge to the lesion and its surrounding area during the judgment. To quantify the assessment of the visualization, we calculated the average attention score and found that the mean importance of the voxels in the lesion area was about twice that of the voxels in the non-lesion area. As a gradient-based visualization method, Grad-CAM plays a strong role in the explainable field of deep networks, although it still has some defects that can be optimized [34,35]. The question of how to associate the properties of multiple views with the network visualization algorithm is a very meaningful direction for raw data research.

We also performed intra-CT and inter-CT analyses. For intra-CTM, fused raw data prediction had a higher optimization rate than the error rate, which showed a similar gain trend, and about 80% of the optimized patients appeared in at least two feature extraction networks. For the inter-CTMs, 85% of the patients that all CTMs predicted incorrectly had optimizable RGMs within each CTM. The results demonstrated that the gain of raw data was stable across different convolutional networks and different CTM approaches. Therefore, exploring the drawbacks of post-reconstructed CT image analysis and developing models for direct diagnosis from raw data are the keys to future research. The results of the subgroup analysis showed that the RGMs performed better than the CTMs in most subgroups, especially in the subgroups of older, female, and smaller lesion size, indicating that the raw data provided valuable information that brought model gains in these subgroups, while this information may have been lost in the process of CT reconstruction.

This study has some limitations. First, it involved a small number of patients, and the proportion of positive and negative samples was unbalanced. Further study on large-scale multicenter datasets

should be performed. Second, only patients with a single nodule were included in this study, so further validation of our method on patients with multiple nodules should be performed. Third, although the raw data had a comparable performance with CT, it still had a certain gap with the best CT diagnosis. Thus, there is an urgent need to develop novel AI methods specifically for raw data.

Strong computing power is a problem that cannot be ignored when calculating raw data. It is not realistic to read the complete high-frequency scanning data directly to a computing device. Designing appropriate preprocessing algorithms and building deep networks that align with the characteristics of raw data will be potential breakthrough points in the future. Finally, the CT scan scheme is designed for image reconstruction and may be not suitable for raw data analysis. Therefore, novel scan strategies, such as scanning for specific diagnostic purposes, should be developed to maximize the gain of raw data.

5. Conclusions

In summary, for the first time, we have validated the potential value of raw data in real clinical practice. Our raw data analysis showed comparable performance with CT images, indicating that leveraging non-image information holds promise as an alternative to image-based methods. Moreover, this study confirmed the value of adding raw data to CT images, demonstrating that the combination of non-images and images will further promote the advance of disease diagnosis. This study proposed and validated a new feasible direction for diagnosis without image reconstruction, which may facilitate the development of fully automated scanning and diagnostic processes.

Acknowledgments

The authors would like to acknowledge the instrumental and technical support of the multi-modal biomedical imaging experimental platform at the Institute of Automation, Chinese Academy of Sciences. This work was supported by the National Key Research and Development Program of China (2017YFA0205200), National Natural Science Foundation of China (82022036, 91959130, 81971776, 62027901, 81930053, and 81771924), The Beijing Natural Science Foundation (Z20J00105), Strategic Priority Research Program of Chinese Academy of Sciences (XDB38040200), Chinese Academy of Sciences (GJJSTD20170004 and QYZDJ-SSW-JSC005), The Project of High-Level Talents Team Introduction in Zhuhai City (Zhuhai HLHPTP201703), the Youth Innovation Promotion Association CAS(Y2021049), and the China Postdoctoral Science Foundation (2021M700341).

Compliance with ethics guidelines

Bingxi He, Yu Guo, Yongbei Zhu, Lixia Tong, Boyu Kong, Kun Wang, Caixia Sun, Hailin Li, Feng Huang, Liwei Wu, Meng Wang, Fanyang Meng, Le Dou, Kai Sun, Tong Tong, Zhenyu Liu, Ziqi Wei, Wei Mu, Shuo Wang, Zhenchao Tang, Shuaitong Zhang, Jingwei Wei, Lizhi Shao, Mengjie Fang, Juntao Li, Shouping Zhu, Lili Zhou, Shuo Wang, Di Dong, Huimao Zhang, and Jie Tian declare that they have no conflict of interest or financial conflicts to disclose.

Author contributions

Bingxi He, Yongbei Zhu, Caixia Sun, Tong Tong, Kai Sun, and Hailin Li developed the network architecture and data/modeling infrastructure, training and testing setup. Bingxi He, Yongbei Zhu, Caixia Sun, Tong Tong, Kai Sun, and Hailin Li wrote the methods. Bingxi He, Caixia Sun, and Tong Tong created the figures. Bingxi He and Mengjie Fang performed statistical analysis. Jie Tian, Di Dong, Zhenyu Liu, Kun Wang, Ziqi Wei, Wei Mu, Shuo Wang, Zhenchao Tang, Shuaitong Zhang, Jingwei Wei, and Lizhi Shao advised on the modeling techniques. Lixia Tong, Liwei Wu, Shouping Zhu, and Juntao Li provided raw data structure information. Di Dong, Bingxi He, Yongbei Zhu, Caixia Sun, Tong Tong, Kai Sun, Hailin Li, Liwei Wu, and Yu Guo wrote the manuscript. Huimao Zhang, Yu Guo, Wei Mu, Fanyang Meng, Le Dou, Lili Zhou, and Shuo Wang provided clinical expertise and guidance on the study design. Huimao Zhang, Yu Guo, Wei Mu, Fanyang Meng, and Le Dou created the clinical datasets, interpreted the data, and defined the clinical labels. Lixia Tong, Liwei Wu, and Feng Huang created the raw datasets. Jie Tian, Di Dong, Feng Huang, and Huimao Zhang initiated the project and provided guidance on the concept and design. Jie Tian, Feng Huang, and Huimao Zhang supervised the project.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eng.2023.02.013.

References

[1] Ciccarelli EC, Jacobs AJ, Berman P. Looking back on the millennium in medicine. N Engl J Med 2000;342:1365–2137.
[2] Lauwerends LJ, van Driel PBAA, Baatenburg de Jong RJ, Hardillo JAU, Koljenovic S, Puppels G, et al. Real-time fluorescence imaging in intraoperative decision making for cancer surgery. Lancet Oncol 2021;22(5):e186–95.
[3] Lehman CD, Wellman RD, Buist DSM, Kerlikowske K, Tosteson AN, Miglioretti DL, the Breast Cancer Surveillance Consortium. Diagnostic accuracy of digital screening mammography with and without computer-aided detection. JAMA. Intern Med 2015;175(11):1828–37.
[4] Bi WL, Hosny A, Schabath MB, Giger ML, Birkbak NJ, Mehrtash A, et al. Artificial intelligence in cancer imaging: clinical challenges and applications. CA Cancer J Clin 2019;69(2):127–57.
[5] Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. Nat Rev Cancer 2018;18(8):500–10.
[6] Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Med Image Anal 2017;42:60–88.
[7] Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017;14(12):749–62.
[8] Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. Lancet Digit Health 2019;1(6):e271–97.
[9] Killock D. AI outperforms radiologists in mammographic screening. Nat Rev Clin Oncol 2020;17(3):134.
[10] Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ, Ashrafian H, et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. Lancet Digit Health 2020;2(10):e549–60.
[11] Dong D, Fang MJ, Tang L, Shan XH, Gao JB, Giganti F, et al. Deep learning radiomic nomogram can predict the number of lymph node metastasis in locally advanced gastric cancer: an international multicenter study. Ann Oncol 2020;31(7):912–20.
[12] Huang YQ, Liang CH, He L, Tian J, Liang CS, Chen X, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. J Clin Oncol 2016;34(18):2157–64.
[13] Mu W, Schabath MB, Gillies RJ. Images are data: challenges and opportunities in the clinical translation of radiomics. Cancer Res 2022;82(11):2066–8.
[14] Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology 2016;278(2):563–77.
[15] Zhu B, Liu JZ, Cauley SF, Rosen BR, Rosen MS. Image reconstruction by domain-transform manifold learning. Nature 2018;555(7697):487–92.
[16] Wang G, Ye JC, De Man B. Deep learning for tomographic image reconstruction. Nat Mach Intell 2020;2(12):737–48.
[17] Chung C, Kalpathy-Cramer J, Knopp MV, Jaffray DA. In the era of deep learning, why reconstruct an image at all? J Am Coll Radiol 2021;18(1):170–3.
[18] Wang G. A perspective on deep imaging. IEEE Access 2016;4:8914–24.
[19] Kalra M, Wang G, Orton CG. Radiomics in lung cancer: its time is here. Med Phys 2018;45(3):997–1000.
[20] Wang G, Ye JC, Mueller K, Fessler JA. Image reconstruction is a new frontier of machine learning. IEEE Trans Med Imaging 2018;37(6):1289–96.
[21] De Man Q, Haneda E, Claus B, Fitzgerald P, De Man B, Qian G, et al. A two-dimensional feasibility study of deep learning-based feature detection and characterization directly from CT sinograms. Med Phys 2019;46(12):e790–800.

[22] Wu D, Kim K, Dong B, Li Q. End-to-end abnormality detection in medical imaging. In: Proceedings of the 6th International Conference on Learning Representations (ICLR 2018); 2018 Apr 30–May 3; Vancouve, BC, Canada. San Francisco: OpenReview; 2018.
[23] Gao Y, Tan J, Liang Z, Li L, Huo Y. Improved computer-aided detection of pulmonary nodules via deep learning in the sinogram domain. Vis Comput Ind Biomed Art 2019;2(1):15.
[24] Dong D, He B, Kong B, Zhang L, Tong L, Huang F, et al. Abstract CT274: diagnosis based on signal: the first time break the routinely used circle of signal-to-image-to-diagnose. Cancer Res 2020;80(Suppl 16):CT274.
[25] Xu X, Wang C, Guo J, Gan Y, Wang J, Bai H, et al. MSCS-DeepLN: evaluating lung nodule malignancy using multi-scale cost-sensitive neural networks. Med Image Anal 2020;65:101772.
[26] Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nat Med 2019;25(6):954–61.
[27] Liu L, Dou Q, Chen H, Qin J, Heng PA. Multi-task deep model with margin ranking loss for lung nodule analysis. IEEE Trans Med Imaging 2020;39(3):718–28.
[28] Xie Y, Xia Y, Zhang J, Song Y, Feng D, Fulham M, et al. Knowledge-based collaborative deep learning for benign–malignant lung nodule classification on chest CT. IEEE Trans Med Imaging 2019;38(4):991–1004.
[29] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017); 2017 Jul 21–26; Honolulu, HI, USA. New York City: IEEE; 2017. p. 4700–8.
[30] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the 2016 IEEE Conference on Computer Vision Pattern Recognition; 2016 Jun 26–Jul 1; Las Vegas, NV, USA. New York City: IEEE; 2016. p. 770–8.
[31] Xie S, Girshick R, Dollár P, Tu Z, He K. Aggregated residual transformations for deep neural networks. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017); 2017 Jul 21–26; Honolulu, HI, USA. New York City: IEEE; 2017. p. 1492–500.
[32] Shen W, Zhou M, Yang F, Yu D, Dong D, Yang C, et al. Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification. Pattern Recognit 2017;61:663–73.
[33] Mukherjee P, Zhou M, Lee E, Schicht A, Balagurunathan Y, Napel S, et al. A shallow convolutional neural network predicts prognosis of lung cancer patients in multi-institutional computed tomography image datasets. Nat Mach Intell 2020;2(5):274–82.
[34] Chattopadhay A, Sarkar A, Howlader P, Balasubramanian VN. Grad-CAM++: generalized gradient-based visual explanations for deep convolutional networks. In: Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV 2018); 2018 Mar 12–15; Lake Tahoe, NV, USA. New York City: IEEE; 2018. p. 839–47.
[35] Wang H, Wang Z, Du M, Yang F, Zhang Z, Ding S, et al. Score-CAM: score-weighted visual explanations for convolutional neural networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2020); 2020 Jun 14–19; online. New York City: IEEE; 2020. p. 24–5.