

Table S1 Parameters used for the acquisition of images for each CT scanner

Manufacturer, model	Pitch	Tube voltage (kVp)/tube current (ma)	Rotation time (s)	Field of view (mm)	Thickness (mm)/interval (mm)	Reconstruction kernel	Reconstruction method
Phantom							
Canon Medical Systems, Aquilion ONE TSX-301C*	0.813, 0.637, 1.388	120/75, 80/25, 135/100	0.5, 0.75	350, 400	1/1, 2.0/2.0, 5.0/5.0	FC56, FC17, FC86	Standard, mild, strong, enhanced
Canon Medical Supply, Aquilion ONE TSX-301A	0.828	120/75	0.5	350	1/1	FC56	Standard
Hitachi, Scenaria	0.8281	120/75	0.5	350	1/1	66	Intelli IP (Lv. 2)
Philips Healthcare, Brilliance	0.891	120/75	0.5	350	1/1	L	Standard (enhancement =1.0)
Siemens Healthineers, SOMATOM Definition AS	0.9	120/75	0.5	350	1/1	B60f	ADMIRE (strength =3)
United Imaging Healthcare, UCT550	0.8875	120/75	0.5	350	1/1	Sharp	Adaptive filter function (enhancement =2.5)
GE Healthcare, Revolution CT	0.992	120/75	0.5	350	1.25/1.25	Lung	ASiR (Plus/SS40)
Clinical validation							
Philips Healthcare, Brilliance	0.891	120/(150–200) [†]	0.75	350	1/1	L	Standard (enhancement =1.0)
Siemens AG, SOMATOM Definition AS	0.9	120/(150–200) [†]	0.5	350	1/1	B60f	3.13%
Hitachi, Scenaria	0.8281	120/(150–200) [†]	0.5	350	1/1	66	Intelli IP (Lv.2)
Siemens, Emotion 16	0.9	120/(150–200) [†]	0.5	350	1/1	B30s	Standard
Canon Medical Supply, Aquilion ONE	0.813	120/(150–200) [†]	0.5	350	1/1	FC56	Standard

*, Aquilion ONE TSX-301C was chosen for the intra-CT protocol trial; [†], tube current was automatically adjusted within the range of 150–200 mA. CT, computed tomography; ADMIRE, advanced modeled iterative reconstruction; ASiR, adaptive statistical iterative reconstruction.

Table S2 Evaluation of feature robustness of the three models

Variable	Two-block-ResNet		ResNet18		DenseNet121	
	No. of stability features	Ratio	No. of stability features	Ratio	No. of stability features	Ratio
Test-retest	46	71.88	0	0.00	0	0.00
Inter-CT	56	87.50	33	51.56	2	3.13
Pitch	64	100.00	0	0.00	0	0.00
Rotation time	64	100.00	0	0.00	2	3.12
Tube voltage	64	100.00	0	0.00	0	0.00
Tube current	64	100.00	5	7.81	18	28.12
Field of view	64	100.00	5	7.81	6	9.38
Slice thickness and slice interval	33	51.56	0	0.00	1	1.56
Reconstruction kernel	57	89.06	0	0.00	0	0.00
Iteration level	64	100.00	58	90.62	56	87.50

Table S2 shows the outcomes from applying different DLR approaches. The feature robustness is remarkable on the shallower network structure of the two-block-ResNet. In contrast, the feature robustness is significantly lower on the deep learning network structures of ResNet18 and DenseNet121. We speculate that deep learning networks are more likely to learn image details irrelevant to the biomarker, but the effectiveness of the DLR approach for deep learning networks remains to be investigated. In the meantime, we suggest that for the selection of a network as a feature extractor, a network model with fewer layers is preferable for extracting more robust image features. CT, computed tomography; ResNet, residual net.