Table S1 Predictive value for lung cancer development using 10–90% thresholds in the validation set

Models	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)
Threshold >0.1					
3D-CNN	100.0	0.0	4.0	0.0	4.0
MobileNet v2	100.0	0.0	4.0	0.0	4.0
EfficientNet-B0	100.0	4.9	4.2	100.0	8.7
SEResNet18	98.1	12.6	4.4	99.4	16.0
Threshold >0.2					
3D-CNN	96.2	13.7	4.4	98.9	17.0
MobileNet v2	100.0	0.0	4.0	0.0	4.0
EfficientNet-B0	94.2	19.1	4.6	98.8	22.1
SEResNet18	94.2	37.8	5.9	99.4	40.0
Threshold >0.3					
3D-CNN	86.5	50.6	6.8	98.9	52
MobileNet v2	100.0	0.0	0.0	0.0	4.0
EfficientNet-B0	90.4	31.6	5.2	98.8	33.9
SEResNet18	88.5	56.1	7.7	99.2	57.4
Threshold >0.4					
3D-CNN	71.2	70.8	9.2	98.3	70.8
MobileNet v2	98.1	21.8	4.9	99.6	24.8
EfficientNet-B0	84.6	44.4	5.9	98.6	46.0
SEResNet18	84.6	69.3	10.3	99.1	69.9
Threshold >0.5					
3D-CNN	53.8	83.3	11.8	97.8	82.1
MobileNet v2	38.5	88.3	12.0	97.2	86.3
EfficientNet-B0	76.9	51.0	6.1	98.2	52.0
SEResNet18	69.2	78.0	11.5	98.4	77.6
Threshold >0.6					
3D-CNN	36.5	92.1	16.1	97.2	89.9
MobileNet v2	26.9	95.2	18.9	96.9	92.5
EfficientNet-B0	75.0	55.7	6.6	98.2	56.4
SEResNet18	69.2	81.3	13.3	98.5	80.9
Threshold >0.7					
3D-CNN	17.3	97.3	20.5	96.6	94
MobileNet v2	15.4	98.1	25.0	96.5	94.8
EfficientNet-B0	65.4	64.0	7.0	97.8	64.0
SEResNet18	65.4	86.8	17.1	98.4	86.0
Threshold >0.8					
3D-CNN	9.6	100.0	100.0	96.4	96.4
MobileNet v2	9.6	99.8	71.4	96.4	96.2
EfficientNet-B0	46.2	70.5	6.1	96.9	69.5
SEResNet18	42.3	91.2	16.7	97.4	89.3
Threshold >0.9					
3D-CNN	1.9	100.0	100.0	96.1	96.1
MobileNet v2	3.8	100.0	100.0	96.2	96.2
EfficientNet-B0	32.7	82.0	7.0	96.7	80.0
SEResNet18	25.0	96.6	23.2	96.9	93.7

PPV, positive predictive value; NPV, negative predictive value; 3D, three-dimensional; CNN, convolutional neural network.

Table S2 Predictive value for lung cancer development using 10–90% thresholds in the test set

Models	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)
Threshold >0.1					
3D-CNN	98.4	12.3	38.5	93.3	43.1
MobileNet v2	98.4	12.3	38.5	93.3	43.1
EfficientNet-B0	95.3	29.8	43.1	91.9	53.2
SEResNet18	97.6	27.6	42.9	95.5	52.7
Threshold >0.2					
3D-CNN	92.9	38.2	45.6	90.6	57.7
MobileNet v2	95.3	27.2	42.2	91.2	51.5
EfficientNet-B0	87.4	43.4	46.2	86.1	59.2
SEResNet18	87.4	53.5	51.2	88.4	65.6
Threshold >0.3					
3D-CNN	71.1	74.0	58.8	83.1	72.1
MobileNet v2	88.2	42.5	46.1	86.6	58.9
EfficientNet-B0	80.3	54.4	49.5	83.2	63.7
SEResNet18	82.7	68.9	59.7	87.7	73.8
Threshold >0.4					
3D-CNN	61.4	84.2	68.4	79.7	76.1
MobileNet v2	79.5	62.7	54.3	84.6	68.7
EfficientNet-B0	73.2	62.7	52.2	80.8	66.5
SEResNet18	71.7	77.2	63.6	83.0	75.2
Threshold >0.5					
3D-CNN	42.5	93.0	77.1	74.4	74.9
MobileNet v2	67.7	76.8	61.9	81.0	73.5
EfficientNet-B0	66.9	69.3	54.8	79.0	68.5
SEResNet18	60.6	84.6	68.8	79.4	76.1
Threshold >0.6					
3D-CNN	29.9	95.6	79.2	71.0	72.1
MobileNet v2	54.3	87.7	71.1	77.5	75.8
EfficientNet-B0	58.3	75.0	56.5	76.3	69.0
SEResNet18	52.0	87.3	69.5	76.5	74.6
Threshold >0.7					
3D-CNN	17.3	98.7	88.0	68.2	69.6
MobileNet v2	37.0	96.5	85.5	73.3	75.2
EfficientNet-B0	49.6	80.7	58.9	74.2	69.6
SEResNet18	44.9	94.7	82.6	75.5	76.9
Threshold >0.8					
3D-CNN	8.7	99.6	91.7	66.2	67.0
MobileNet v2	17.3	99.1	91.7	68.3	69.9
EfficientNet-B0	43.3	86.8	64.7	73.3	71.3
SEResNet18	37.0	96.9	87.0	73.4	75.5
Threshold >0.9					
3D-CNN	4.7	100.0	100.0	65.3	65.9
MobileNet v2	2.4	100.0	100.0	64.8	65.1
EfficientNet-B0	30.7	94.3	75.0	71.0	71.5
SEResNet18	21.3	99.6	96.4	69.4	71.5

PPV, positive predictive value; NPV, negative predictive value; 3D, three-dimensional; CNN, convolutional neural network.
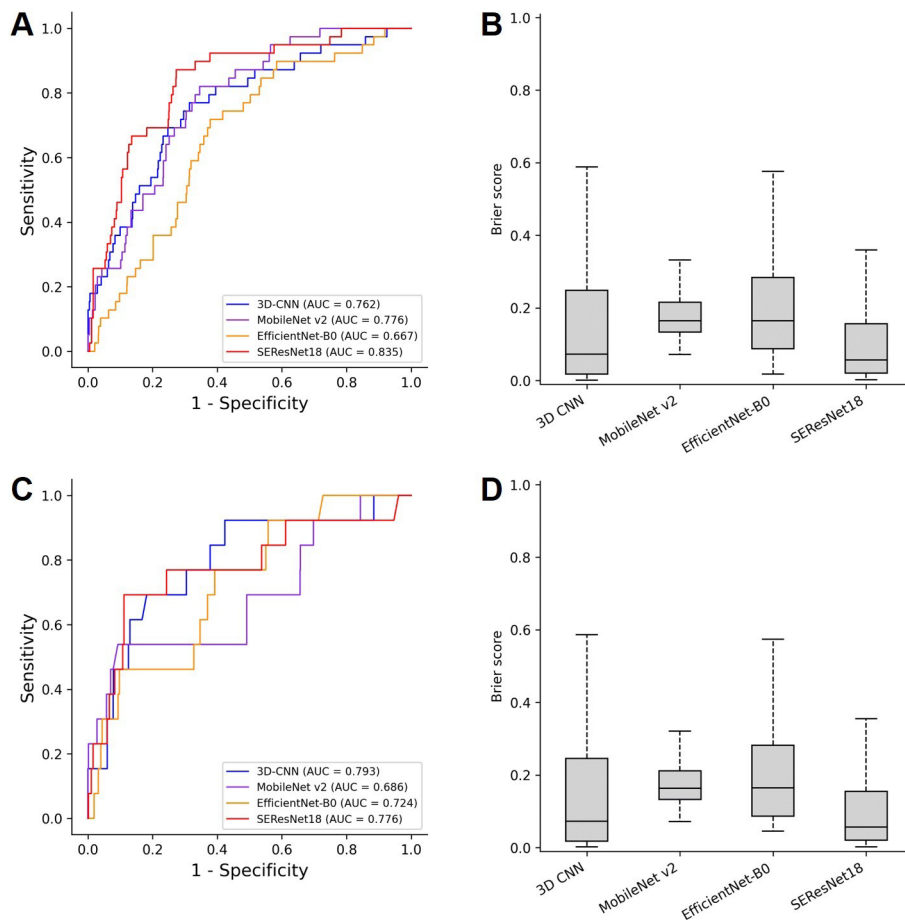
Figure S1 ROC curves, and Brier score for the 3D-CNN, MobileNet v2, SEResNet18, and EfficientNet-B0 in the test data set, categorized by histopathologic types of lung cancer. ROC curve (A), and Brier score (B) for Group A (adenocarcinoma and squamous cell carcinoma) and ROC curve (C), and Brier score (D) for Group B (SCLC, carcinoid tumors, and rare types). 3D, three-dimensional; CNN, convolutional neural network; AUC, area under the ROC curve; ROC, receiver operating characteristic; SCLC, small cell lung cancer.
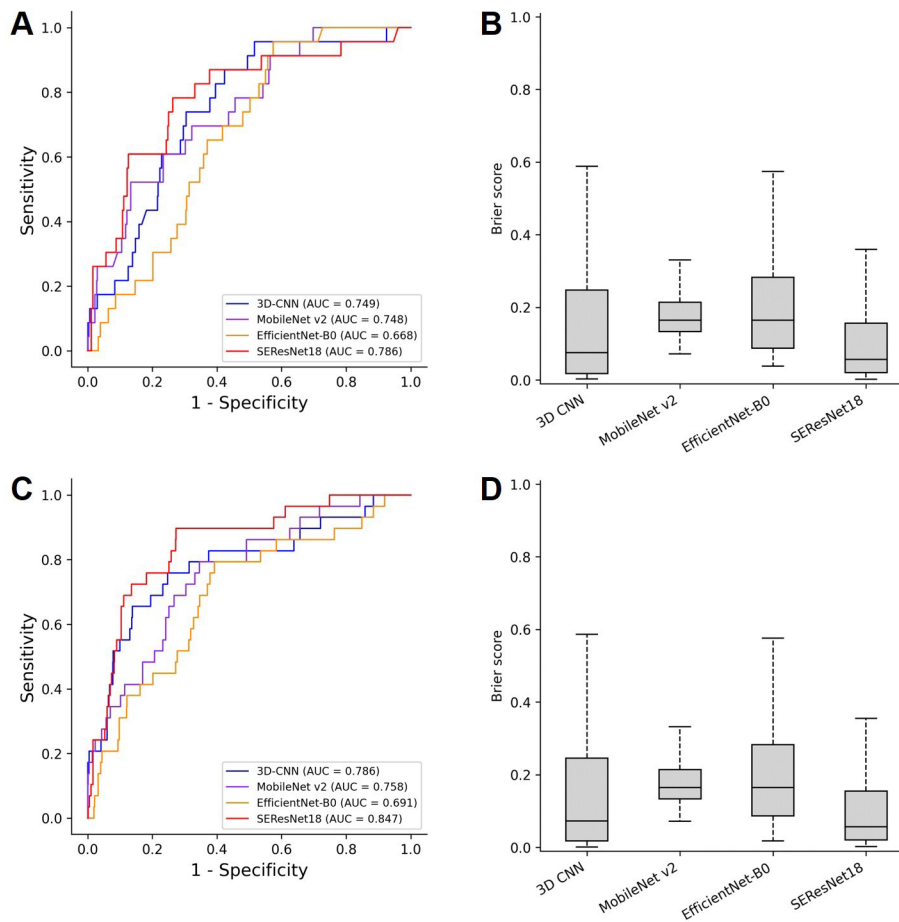
Figure S2 ROC curves, and Brier score for the 3D-CNN, MobileNet v2, SEResNet18, and EfficientNet-B0 in the test data set, categorized by presence of pulmonary nodules on LDCT images. ROC curves (A), and Brier score (B) for the non-nodule group and ROC curves (C), and Brier score (D) for the nodule group. 3D, three-dimensional; CNN, convolutional neural network; AUC, area under the ROC curve; ROC, receiver operating characteristic; LDCT, low-dose chest computed tomography.