Translational Oncology 42 (2024) 101894

Contents lists available at ScienceDirect

Translational Oncology

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

Original Research

Development and validation of a deep learning signature for predicting lymphovascular invasion and survival outcomes in clinical stage IA lung adenocarcinoma: A multicenter retrospective cohort study

Kunfeng Liu [a,1], Xiaofeng Lin [a,1], Xiaojuan Chen [b,1], Biyun Chen [a], Sheng Li [a], Kunwei Li [c,1,*], Huai Chen [d,1,*], Li Li [a,*]

[a] Department of Medical Imaging, State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, Guangzhou 510060, PR China
[b] Department of Radiology, First Affiliated Hospital of Guangzhou Medical University, Guangzhou 510120, PR China
[c] Department of Radiology, The Fifth Affiliated Hospital of Sun Yat-sen University, Zhuhai 519000, PR China
[d] Department of Radiology, The Second Affiliated Hospital of Guangzhou Medical University, Guangzhou 510260, PR China

ARTICLE INFO

Keywords:
Lung cancer
Adenocarcinoma
Lymphovascular invasion
Survival
Deep learning

ABSTRACT

Purpose: The presence of lymphovascular invasion (LVI) influences the management and outcomes of patients with clinical stage IA lung adenocarcinoma. The objective was the development of a deep learning (DL) signature for the prediction of LVI and stratification of prognosis.

Methods: A total of 2077 patients from three centers were retrospectively enrolled and divided into a training set (n = 1515), an internal validation set (n = 381), and an external set (n = 181). A -three-dimensional residual neural network was used to extract the DL signature and three models, namely, the clinical, DL, and combined models, were developed. Diagnostic efficiency was assessed by ROC curves and AUC values. Kaplan-Meier curves and Cox proportional hazards regression analyses were conducted to evaluate links between various factors and disease-free survival.

Results: The DL model could effectively predict LVI, shown by AUC values of 0.72 (95 %CI: 0.68–0.76) and 0.63 (0.54–0.73) in the internal and external validation sets, respectively. The incorporation of DL signature and clinical-radiological factors increased the AUC to 0.74 (0.71–0.78) and 0.77 (0.70–0.84) in comparison with the DL and clinical models (AUC of 0.71 [0.68–0.75], 0.71 [0.61–0.81]) in the internal and external validation sets, respectively. Pathologic LVI, LVI predicted by both DL and combined models were associated with unfavorable prognosis (all $p < 0.05$).

Conclusion: The effectiveness of the DL signature in the diagnosis of LVI and prognosis prediction in patients with clinical stage IA lung adenocarcinoma was demonstrated. These findings suggest the potential of the model in clinical decision-making.

Abbreviations

AUC	area under the curve
CT	computed tomography
CTR	consolidation tumor ratio
DL	deep learning
DFS	disease-free survival
HR	hazard ratio
ICC	intraclass correlation coefficient
LLL	left lower lobe
LUL	left upper lobe
LVI	lymphovascular invasion
NSCLC	non-small cell lung cancer
OR	odds ratio
RLL	right lower lobe
RML	right middle lobe

* Corresponding author at: Department of Medical Imaging, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University Cancer Center, No. 651 Dongfeng East Road, Yuexiu District, Guangzhou 510060, Guangdong, China.
E-mail address: lil@sysucc.org.cn (L. Li).
[1] Kunfeng Liu, Xiaofeng Lin, Xiaojuan Chen, Kunwei Li and Huai Chen contributed equally to this work as first authors.

https://doi.org/10.1016/j.tranon.2024.101894
Received 28 September 2023; Received in revised form 30 December 2023; Accepted 24 January 2024
Available online 6 February 2024

ResNet	residual neural network
ROC	receiver operating characteristic
RUL	right upper lobe
VPI	visceral pleural invasion.

Introduction

Lung cancer is a deadly malignancy with high global mortality. Lung adenocarcinoma is the most prevalent histological subtype of non-small cell lung cancers (NSCLCs) [1]. The proportion of clinical stage IA lung cancer has increased by 58 % over the past decade [2]. They are usually treated with lobectomy together with dissection of the mediastinal lymph nodes [3]. Screening using low-dose computed tomography (LDCT), the) has improved the diagnosis of early-stage disease, and thus the sublobar resection may be a reasonable alternative to lobectomy for stage IA NSCLC patients. Several studies have shown that in properly selected patients, lobectomy and sublobar resection lead to comparable survival outcomes [4,5]. However, it has been reported that clinical stage IA patients may have recurrence and metastasis approaching 35 % [6]. Inappropriate sublobar resections may lead to undesirable surgical outcomes, even in patients with high-risk features. In addition, in terms of the improved survival rate of patients with NSCLC, recent clinical studies have shown that immunotherapy, either as a neoadjuvant or adjuvant therapy, is more effective than chemoradiotherapy alone for early-stage patients [7–10]. Since immunotherapy has made great progress, neoadjuvant and adjuvant therapy plays an important role in improving the survival rate of lung cancer. Thus, whether it is to guide the choice of surgical scope or adjuvant treatment in the clinical stage IA lung adenocarcinoma, the investigation of factors predicting metastasis and recurrence following resection is urgently needed for the stratification and management of these patients.

Lymphovascular invasion (LVI) represents the invasion of malignant cells into arteries, veins, and lymph vessels and is known to be predictive of disease recurrence [11–13]. Neoadjuvant chemotherapy is usually recommended for the preoperative treatment of LVI-positive NSCLC, as it can reduce tumor staging and increase the likelihood of better outcomes [14,15]. It has also been found that lobectomy together with more extensive lymph node dissection results in better outcomes compared with sub-lobectomy in patients with LVI [16]. Thus, the accurate detection of LVI would assist in determining patients likely to benefit from individualized clinical decision-making. While previous studies have indicated that tumor size or the consolidation tumor ratio (CTR) may serve as indicators for assessing LVI, the tumor size cutoff (4.5 cm) is not suitable for clinical stage IA lung cancer [17], and the CTR is dubious (0.25 or 0.5) for those tumor size $\leq$ 3 cm [18]. Therefore, there has been limited investigation of the use of imaging for the prediction of LVI in lung cancer.

The development of deep learning (DL) has altered the landscape of imaging investigations, shifting from human interpretation to self-taught machine analysis [19]. DL is effective in many fields, including the detection and malignancy prediction of lung nodules on CT imaging [20,21]. However, there is limited information on its use for predicting LVI in lung cancer. Here, we investigated the construction of a DL model for the prediction of clinical stage IA lung adenocarcinoma from CT images to develop an effective, simpler, and machine-based means of identifying patients who would benefit from clinical intervention.

Materials and methods

The ethical review board of Sun Yat-sen University Cancer Center and the First Affiliated Hospital of Guangzhou Medical University (B2022-293-01), the Fifth Affiliated Hospital of Sun Yat-sen University (K107-1) approved this study, and waived the requirement for written informed consent.

Patients

The patients included in the study were selected from three hospitals in China, specifically, center 1 (Sun Yat-sen University Cancer Center) between 2010 and 2021, center 2 (The First Affiliated Hospital of Guangzhou Medical University) between 2015 and 2019, and center 3 (The Fifth Affiliated Hospital of Sun Yat-sen University) between 2016 and 2018. The inclusion and exclusion criteria are presented in Fig. 1. The final number of patients enrolled was 2077, including 1070, 826 and 181 patients from three centers, respectively. The participants from centers 1 and 2 were assigned to the training (n = 1515) and internal validation (n = 381) cohorts (ratio 8: 2) using simple randomization. The cohort of 181 patients from center 3 was used as the external validation set. A previous report has described 603 patients from center 1 [22], focusing on the preoperative prediction of malignancy in solitary lung nodules using a DL model.

Clinical and pathological data

The clinical and pathological features of the participants were obtained, including age, sex, symptoms, smoking history (smoker or non-smoker), and family history of cancer. Tumor staging was defined according to TNM from the 8th edition of the American Joint Committee of Cancer (AJCC) [23] and histopathological classifications were according to the International Association for the Study of Lung Cancer/American Thoracic Society/ European Respiratory Society classification system [24]. Data on LVI, visceral pleural invasion (VPI), lymph node metastasis, and distant metastasis status were acquired from pathological reports. LVI was defined as the presence of tumor cells in lymphatic, arterial or venous vessels in the surrounding pulmonary tissue that was visible on microscopy. VPI was classified based on the hematoxylin-eosin–stained slice: PL0 (lack of pleural invasion beyond the elastic layer), PL1 (invasion beyond the elastic layer), PL2 (invasion into the surface of the visceral pleura), and PL3 (parietal pleura involvement). When hematoxylin-eosin staining indicated that the lesion was adjacent to the pleura, and it was uncertain whether the visceral pleura was involved, an elastic stain can be used to determine whether VPI was present. The elastic stain was not used at our institution to reassess VPI during the study inclusion period. PL1 and PL2 could not be recorded separately in this study.

CT image acquisition

CT scans were conducted from the apex to the base of the lung. Supplementary Table E1 shows the details of the scanning and reconstruction parameters.

Evaluation of radiological features

The radiological features were evaluated by three radiologists with 5, 10, and 15 years of experience, respectively, in thoracic imaging and diagnosis. They were blinded to the diagnosis and data of the patients, and interpretation was done by consensus. The specific characteristics evaluated were the maximum diameter of the tumor, the consolidation maximum diameter, CTR, the tumor location (right upper, middle, or lower lung, or left upper, middle, or lower lung), the tumor density (pure-solid or subsolid), shape (regular or irregular), boundary (clear or vague), presence or absence of lobulation (defined as irregular undulations in the margin of the nodule), spiculation (the extension of strands, 2 mm or greater, from the tumor margin to the surrounding parenchyma but not into the pleura), vacuole sign (air attenuation, either ovoid or round, $<$ 5 mm within the tumor), cavity (air attenuation, either ovoid or round, $\geq$ 5 mm within the tumor), air bronchogram (presence of air-filled bronchi within the tumor), vessel convergence (convergence of multiple blood vessels directed to the tumor), and pleural indentation (presence of linear or triangular strands from the

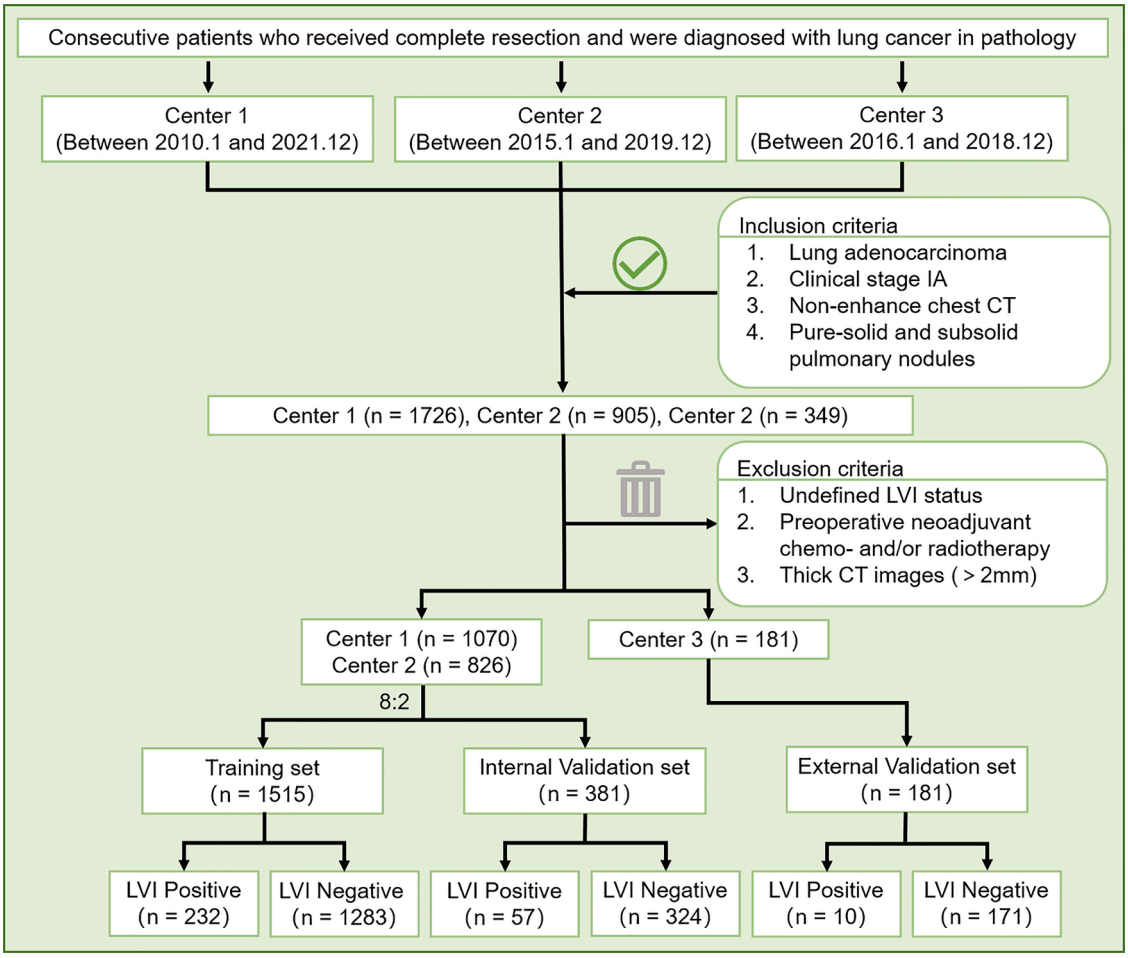

Fig. 1. Flowchart of patient selection. LVI = lymphovascular invasion.

tumor to the pleura). CT images were viewed in lung windows (level, –600 Hounsfield units [HU]; width, 1500 HU).

Data splitting and pretreatment

Learning and selection of the model were conducted in the training cohort while model performance was evaluated in the internal and external validation cohorts. Five-fold stratified cross-validation was used for both cohorts; details are shown in the first part of Fig. 2. For developing the DL system, the CT data samples were preprocessed. The details about the CT image preprocessing are shown in the Appendix E1 (Supplemental Material).

3D-ResNet-9 model architecture

The Residual Neural Network algorithm was used to classify the images into nine layers (3D-ResNet-9), allowing the extraction of specific features for prediction [25]. The 3D-ResNet-9 model is one of the methods used for a residual neural network that functions in a similar manner to other residual neural networks but differs in the number of layers. The main structure of 3D-ResNet-9 consisted of the following parts:

(a) Initial convolution layer: The 3D-ResNet-9 structure began with a 3 × 3 convolution layer with step size 1, followed by a maximum pool (maxpool) layer. This part of the model was responsible for the extraction of preliminary features.
(b) Residual blocks: Four residual blocks formed the main components of the 3D-ResNet-9. Each residual block contained two 3D convolutional layers that operated by batch normalization (BN) and ReLU activation. BN refers to the normalization of batch data, in which a feature map of a data batch function with a distribution rule has an average value of 0 and variance of 1. To reduce resource consumption and training time, the number of channels was reduced to 16, 32, 64, and 96, respectively. The input was then connected directly to the output of these convolutional layers via a skip join (also called an identity join). This design allowed the network to learn the residual mapping between input and output, thus avoiding the problem of disappearing or exploding gradients.
(c) Average global pool (avgpool) layer: Following the residual blocks, 3D-ResNet-9 used an avgpool layer in place of the fully connected layer. This significantly reduced the number of parameters required by the model, thus avoiding overfitting. A full connection layer was used after the avgpool layer for the output of the results of the final classification.

The second part of Fig. 2 illustrates the 3D-ResNet-9 architecture. The images were processed into various layers allowing the extraction of features for prediction. A significant advantage of the 3D-ResNet-9 algorithm is that it avoids the problem of vanishing gradients, rendering it more efficient in comparison with other algorithms.

Model construction and validation

LVI status was used as the predicted outcome, and a binary classification model based on DL was designed, namely, the fully connected neural network [26]. The classifier used a one-dimensional array as input, with the output being a binary score, which was distinguished

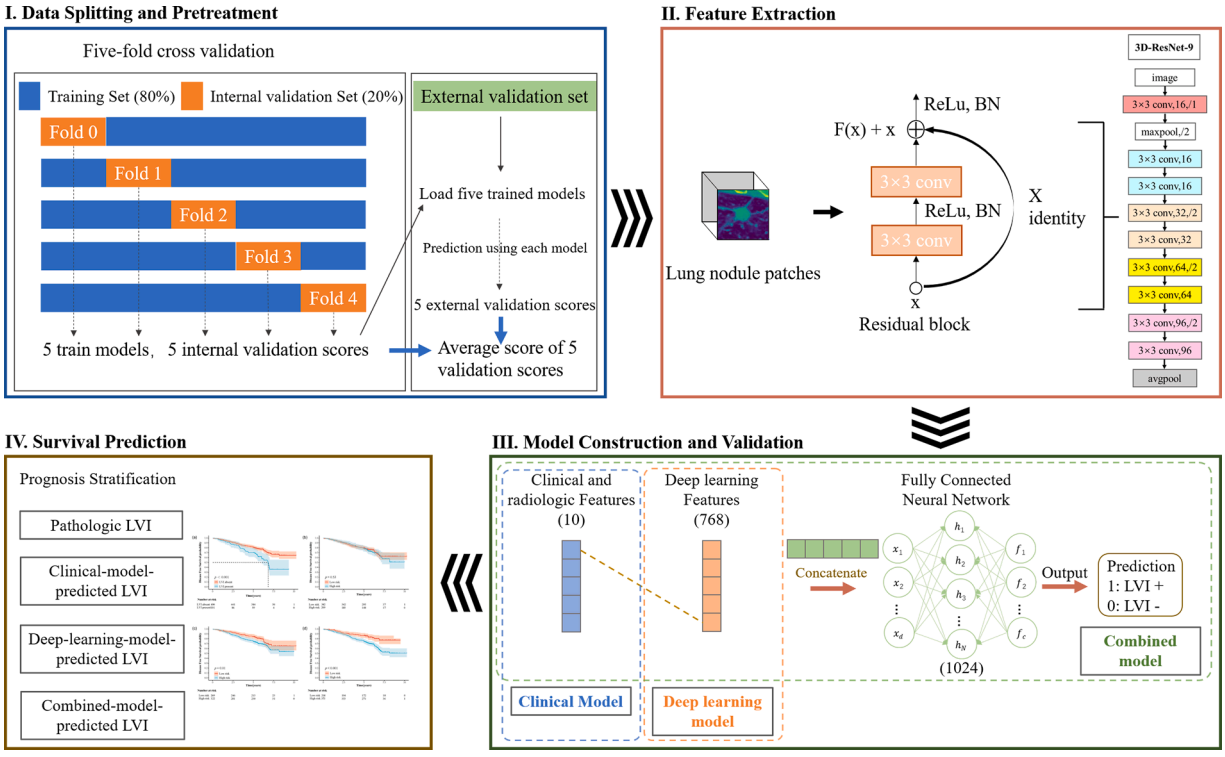

Fig. 2. Flowchart of the study design. BN = batch normalization, LVI = lymphovascular invasion, ResNet = residual neural network.

according to the appropriate threshold value. The category was defined as 1 when the score was greater than the threshold value and if not, the category was described as 0. The third part of Fig. 2 shows the complete architecture of the fully connected neural network. An Adam optimizer was used for training. The learning rate was 0.0005, and the batch size was set to 16. The loss function adopted a weighted L1 norm, and the weight was set according to the positive proportion. Tne model was then constructed to reflect the improvement in classification by the integration of multidimensional features. The model reached convergence within an average of five training cycles. The model was then validated in both the internal and external validation sets. To improve classification performance via the integration of multidimensional features, a clinical model, a DL model, and a combined model were established. Ten features were found to be correlated with LVI status in the univariate regression analyses (Table 2) and placed into model for automatic analysis. The DL model was established using 768 high-quality DL features extracted from 3D-ResNet-9. The combined model was constructed by the insertion of the clinical and radiological factors into the DL model through the fully connected neural network which was designed as a single hidden layer, with the neuron number in the layer set to 1024. The neuron number in the input layer was set according to the number of input features while the number in the output layer was set to 1, indicative of positive output probability. The third part of Fig. 2 shows the model construction.

Gradient-weighted class activation mapping (Grad-CAM) was used to determine the region of the CT image that contributed most to prediction [27] and the results were visualized in heatmaps. The Grad-CAM details are shown in the Appendix E2 (Supplemental Material).

Follow-up and survival

Following surgery, the patients were evaluated by CT every six months for the initial two years and every 12 months thereafter. Overall, 591 patients were followed up for a minimum of five years. The study endpoint was defined as disease-free survival (DFS), defined as the interval between the surgery and recurrence, metastasis, death, or final follow-up. Significant variables from the univariate analysis were then incorporated into the multivariate analysis for the determination of independent DFS predictors. Kaplan-Meier curves were plotted and differences between the high- and low-risk groups were examined by log-rank tests.

Statistical analysis

Data were analyzed using Python version 3.7.0, PyTorch version 1.6.0, CUDA version 9.2, and R version 3.5.3. Clinicopathological and radiological characteristics are expressed as means ± standard deviations for continuous variables and numbers (percentages) for categorical variables. Inter-group comparisons were evaluated by unpaired t-tests, chi-squared tests, or Mann-Whitney U tests. Inter-observer agreement (readers 1, 2, and 3) was evaluated by Cohen's kappa test and intraclass correlation coefficients. Receiver operating characteristic (ROC) curves, the areas under the curves (AUCs) and the DeLong test were used to compare the predictive performance of the models. The evaluation matrix included accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and F1 scores. Calibration curves and box plots were used to determine the degree of fit of the models. Survival differences were analyzed by Kaplan-Meier curves and log-rank tests. All significance levels were two-tailed with $p < 0.05$ considered statistically significant.

Results

Patient characteristics

Details of the participants, including clinical, radiological, and pathological features, are provided in Table 1. There was good agreement between the three chest radiologists in the evaluation of the radiological characteristics ($p < 0.05$) (Supplementary Tables E2 and E3). No significant differences were observed in the patient characteristics between the training and internal validation cohorts ($p > 0.05$). It was found that LVI-positive and -negative participants in three cohorts

Table 1
Clinical, Pathological, and Radiological Characteristics

	Training Set (n = 1515)		*p1 Value*	*Internal validation Set (n = 381)*		*p1 Value*	*Internal validation Set (n = 181)*		*p1 Value*	p2 Value
	LVI Positive (n = 232)	LVI Negative (n = 1283)		LVI Positive (n = 57)	LVI Negative (n = 324)		LVI Positive (n = 10)	LVI Negative (n = 171)		
A. Clinical characteristics										
Age*	58.8±10.6	58.4±9.79	0.660	60.5±11.2	57.6±9.86	0.072	57.8±10.2	59.2±10.4	0.690	0.452
Sex			0.149			0.417			0.927	0.562
Male	119 (51.3 %)	589 (45.9 %)		31 (54.4 %)	154 (47.5 %)		4 (40.0 %)	80 (46.8 %)		
Female	113 (48.7 %)	694 (54.1 %)		26 (45.6 %)	170 (52.5 %)		6 (60.0 %)	91 (53.2 %)		
Symptom			0.145			0.856			0.823	0.333
Yes	134 (57.8 %)	809 (63.1 %)		36 (63.2 %)	212 (65.4 %)		2 (20.0 %)	21 (12.3 %)		
No	98 (42.2 %)	474 (36.9 %)		21 (36.8 %)	112 (34.6 %)		8 (80.0 %)	150 (87.7 %)		
Smoking history			0.277			0.985			0.113	0.714
Ever smoker	68 (29.3 %)	329 (25.6 %)		15 (26.3 %)	89 (27.5 %)		0 (0 %)	48 (28.1 %)		
Never smoker	164 (70.7 %)	954 (74.4 %)		42 (73.7 %)	235 (72.5 %)		10 (100 %)	123 (71.9 %)		
Family history of cancer			0.242			0.170			0.997	0.965
Yes	25 (10.8 %)	178 (13.9 %)		4 (7.0 %)	48 (14.8 %)		1 (10.0 %)	8 (4.7 %)		
No	207 (89.2 %)	1105 (86.1 %)		53 (93.0 %)	276 (85.2 %)		9 (90.0 %)	163 (95.3 %)		
Clinical T stage			<0.001			<0.001			0.005	0.999
cT1a	5 (2.2 %)	362 (28.2 %)		2 (3.5 %)	90 (27.8 %)		1 (10.0 %)	56 (32.7 %)		
cT1b	97 (41.8 %)	593 (46.2 %)		20 (35.1 %)	154 (47.5 %)		2 (20.0 %)	75 (43.9 %)		
cT1c	130 (56.0 %)	328 (25.6 %)		35 (61.4 %)	80 (24.7 %)		7 (70.0 %)	40 (23.4 %)		
B. Pathological characteristics										
Histologic subtype			<0.001			<0.001			0.533	0.563
Lepidic predominant	5 (2.2 %)	163 (12.7 %)		0 (0 %)	48 (14.8 %)		0 (0 %)	6 (3.5 %)		
Acinar predominant	134 (57.8 %)	777 (60.6 %)		32 (56.1 %)	180 (55.6 %)		4 (40.0 %)	84 (49.1 %)		
Papillary predominant	39 (16.8 %)	160 (12.5 %)		7 (12.3 %)	45 (13.9 %)		4 (40.0 %)	47 (27.5 %)		
Micropapillary predominant	15 (6.5 %)	20 (1.6 %)		4 (7.0 %)	12 (3.7 %)		0 (0 %)	3 (1.8 %)		
Solid predominant	23 (9.9 %)	61 (4.8 %)		10 (17.5 %)	13 (4.0 %)		1 (10.0 %)	10 (5.8 %)		
Mucous predominant	7 (3.0 %)	37 (2.9 %)		2 (3.5 %)	13 (4.0 %)		0 (0 %)	5 (2.9 %)		
Special type	1 (0.4 %)	5 (0.4 %)		0 (0 %)	1 (0.3 %)		1 (10.0 %)	2 (1.2 %)		
MIA	0 (0 %)	10 (0.8 %)		0 (0 %)	2 (0.6 %)		0 (0 %)	1 (0.6 %)		
AIS	0 (0 %)	1 (0.1 %)		0 (0 %)	0 (0 %)		0 (0 %)	0 (0 %)		
Non-classified	8 (3.4 %)	49 (3.8 %)		2 (3.5 %)	10 (3.1 %)		0 (0 %)	13 (7.6 %)		
Pathologic stage			<0.001			<0.001			<0.001	0.838
IA	93 (40.1 %)	1032 (80.4 %)		23 (40.4 %)	263 (81.2 %)		6 (60.0 %)	139 (81.3 %)		
IB	50 (21.6 %)	164 (12.8 %)		14 (24.6 %)	43 (13.3 %)		1 (10.0 %)	16 (9.4 %)		
IIB	32 (13.8 %)	38 (3.0 %)		9 (15.8 %)	8 (2.5 %)		0 (0 %)	5 (2.9 %)		
IIIA	56 (24.1 %)	45 (3.5 %)		11 (19.3 %)	10 (3.1 %)		2 (20.0 %)	6 (3.5 %)		
IIIB	0 (0 %)	1 (0.1 %)		0 (0 %)	0 (0 %)		1 (10.0 %)	0 (0 %)		
IVA	1 (0.4 %)	3 (0.2 %)		0 (0 %)	0 (0 %)		0 (0 %)	5 (2.9 %)		
VPI			<0.001			<0.001			0.823	0.949
Yes	91 (39.2 %)	196 (15.3 %)		21 (36.8 %)	50 (15.4 %)		2 (20.0 %)	21 (12.3 %)		
No	141 (60.8 %)	1087 (84.7 %)		36 (63.2 %)	274 (84.6 %)		8 (80.0 %)	150 (87.7 %)		
Lymph node metastasis			<0.001			<0.001			0.036	0.477
Yes	89 (38.4 %)	84 (6.5 %)		20 (35.1 %)	18 (5.6 %)		3 (30.0 %)	11 (6.4 %)		
No	143 (61.6 %)	1199 (93.5 %)		37 (64.9 %)	306 (94.4 %)		7 (70.0 %)	160 (93.6 %)		
Distant metastasis			1.000			0.325			1.000	1.000
Yes	1 (0.4 %)	3 (0.2 %)		1 (1.8 %)	0 (0 %)		0 (0 %)	5 (2.9 %)		
No	231 (99.6 %)	1280 (99.8 %)		56 (98.2 %)	324 (100 %)		10 (100 %)	166 (97.1 %)		
C. Radiological characteristics										
Whole maximum diameter*	21.5±5.22	18.6±5.87	<0.001	21.3±4.72	18.4±5.89	<0.001	23.3±7.48	19.9±7.49	0.194	0.492
Consolidation maximum diameter*	20.9±5.51	15.4±7.06	<0.001	20.9±5.49	15.4±6.96	<0.001	22.2±6.08	14.9±7.67	0.004	0.921
CTR*	0.97±0.11	0.82±0.25	<0.001	0.98±0.12	0.83±0.25	<0.001	0.97±0.10	0.77±0.30	<0.001	0.608

(*continued on next page*)

Table 1 (*continued*)

	Training Set (n = 1515)		*p1 Value*	*Internal validation Set (n = 381)*		*p1 Value*	*Internal validation Set (n = 181)*		*p1 Value*	p2 Value
	LVI Positive (n = 232)	LVI Negative (n = 1283)		LVI Positive (n = 57)	LVI Negative (n = 324)		LVI Positive (n = 10)	LVI Negative (n = 171)		
Location			0.277			0.635			0.382	0.258
RUL	76 (32.8 %)	444 (34.6 %)		14 (24.6 %)	108 (33.3 %)		6 (60.0 %)	67 (39.2 %)		
RML	18 (7.8 %)	98 (7.6 %)		5 (8.8 %)	29 (9.0 %)		0 (0 %)	13 (7.6 %)		
RLL	41 (17.7 %)	240 (18.7 %)		15 (26.3 %)	73 (22.5 %)		0 (0 %)	28 (16.4 %)		
LUL	69 (29.7 %)	301 (23.5 %)		16 (28.1 %)	69 (21.3 %)		2 (20.0 %)	44 (25.7 %)		
LLL	28 (12.1 %)	200 (15.6 %)		7 (12.3 %)	45 (13.9 %)		2 (20.0 %)	19 (11.1 %)		
Density			<0.001			<0.001			0.196	0.392
Subsolid	25 (10.8 %)	514 (40.1 %)		3 (5.3 %)	123 (38.0 %)		8 (80.0 %)	92 (53.8 %)		
Pure-solid	207 (89.2 %)	769 (59.9 %)		54 (94.7 %)	201 (62.0 %)		2 (20.0 %)	79 (46.2 %)		
Shape			0.342			0.287			0.806	0.933
Regular	4 (1.7 %)	40 (3.1 %)		0 (0 %)	12 (3.7 %)		3 (30.0 %)	67 (39.2 %)		
Irregular	228 (98.3 %)	1243 (96.9 %)		57 (100 %)	312 (96.3 %)		7 (70.0 %)	104 (60.8 %)		
Boundary			0.208			0.100			0.418	0.570
Clear	48 (20.7 %)	318 (24.8 %)		23 (40.4 %)	75 (23.1 %)		2 (20.0 %)	65 (38.0 %)		
Vague	184 (79.3 %)	965 (75.2 %)		34 (59.6 %)	249 (76.9 %)		8 (80.0 %)	106 (62.0 %)		
Vacuole sign			0.775			0.977			1.000	0.608
Yes	44 (19.0 %)	230 (17.9 %)		9 (15.8 %)	55 (17.0 %)		2 (20.0 %)	30 (17.5 %)		
No	188 (81.0 %)	1053 (82.1 %)		48 (84.2 %)	269 (83.0 %)		8 (80.0 %)	141 (82.5 %)		
Cavity			0.369			0.160			1.000	0.307
Yes	9 (3.9 %)	33 (2.6 %)		6 (10.5 %)	9 (2.8 %)		0 (0 %)	2 (1.2 %)		
No	223 (96.1 %)	1250 (97.4 %)		51 (89.5 %)	315 (97.2 %)		10 (100 %)	169 (98.8 %)		
Spiculated sign			<0.001			<0.001			0.093	0.455
Yes	185 (79.7 %)	791 (61.7 %)		41 (71.9 %)	196 (60.5 %)		8 (80.0 %)	81 (47.4 %)		
No	47 (20.3 %)	492 (38.3 %)		16 (28.1 %)	128 (39.5 %)		2 (20.0 %)	90 (52.6 %)		
Lobulated sign			<0.001			<0.001			0.120	0.738
Yes	226 (97.4 %)	1098 (85.6 %)		54 (94.7 %)	282 (87.0 %)		10 (100 %)	124 (72.5 %)		
No	6 (2.6 %)	185 (14.4 %)		3 (5.3 %)	42 (13.0 %)		0 (0 %)	47 (27.5 %)		
Air bronchogram			0.102			0.044			1.000	0.413
Yes	40 (17.2 %)	286 (22.3 %)		7 (12.3 %)	83 (25.6 %)		4 (40.0 %)	75 (43.9 %)		
No	192 (82.8 %)	997 (77.7 %)		50 (87.7 %)	241 (74.4 %)		6 (60.0 %)	96 (56.1 %)		
Vessel convergence			<0.001			<0.001			0.082	0.649
Yes	219 (94.4 %)	1004 (78.3 %)		56 (98.2 %)	247 (76.2 %)		6 (60.0 %)	49 (28.7 %)		
No	13 (5.6 %)	279 (21.7 %)		1 (1.8 %)	77 (23.8 %)		4 (40.0 %)	122 (71.3 %)		
Pleural indentation			<0.001			<0.001			0.951	0.668
Yes	193 (83.2 %)	892 (69.5 %)		46 (80.7 %)	222 (68.5 %)		3 (30.0 %)	62 (36.3 %)		
No	39 (16.8 %)	391 (30.5 %)		11 (19.3 %)	102 (31.5 %)		7 (70.0 %)	109 (63.7 %)		

Note. — Unless otherwise noted, values are numbers of patients, with percentages in parentheses. AIS = adenocarcinoma in situ, CTR = consolidation tumor ratio, LLL = left lower lobe, LUL = left upper lobe, LVI = lymphovascular invasion, MIA = minimally invasive adenocarcinoma, RLL = right lower lobe, RML = right middle lobe, RUL = right upper lobe, VPI = visceral pleural invasion.

* Data are means ± standard deviations. The P1 value was derived from the univariate association analyses between LVI positive and LVI negative. P2 value was derived from univariate association analyses between the training and internal validation sets.

differed significantly in terms of clinical T stage, pathological stage, lymph node metastasis, consolidation maximum diameter of the tumor, and CTR ($p < 0.05$). After the univariate analysis, multivariate analysis found that the presence of CTR was the independent predictor of LVI status ($p < 0.05$). Table 2 lists the results of the univariate and multivariate analyses for training cohort.

Model efficiency evaluation

As shown in Table 3, the combined model had AUCs of 0.76 (95 % confidence interval [CI]: 0.74–0.77), 0.74 (95 %CI: 0.71–0.78), and 0.77 (95 %CI: 0.70–0.84) for the training, internal, and external validation sets, respectively, which were greater than those for the clinical (0.74 [95 %CI: 0.73–0.76], 0.71 [95 %CI: 0.68–0.75], and 0.71 [95 %CI: 0.61–0.81], respectively) and DL models (0.76 [95 %CI: 0.75–0.78], 0.72 [95 %CI: 0.68–0.76], and 0.63 [95 %CI: 0.54–0.73], respectively).

Table 2
Univariate and Multivariate Logistic Regression Analyses of Factors in the Training Set

Factors	Univariate logistic regression		Multivariate logistic regression	
	OR (95 %CI)	*p*-value	OR (95 %CI)	*p*-value
Age	1.010 (0.996–1.024)	0.176		
Sex	0.843 (0.637–1.117)	0.234		
Symptom	0.819 (0.616–1.090)	0.169		
Smoking history	1.152 (0.840–1.564)	0.373		
Family history of cancer	0.697 (0.434–1.074)	0.117		
Clinical T stage				
cT1b	4.459 (1.970–12.80)	0.001	1.257 (0.498–3.865)	0.656
cT1c	11.11 (4.957–31.76)	<0.001	1.936 (0.702–6.351)	0.231
Whole maximum diameter	1.077 (1.051–1.104)	<0.001	1.231 (0.990–1.570)	0.074
Consolidation maximum diameter	1.109 (1.085–1.134)	<0.001	8.218 (0.641–1.029)	0.100
Density	6.062 (3.939–9.798)	<0.001	2.296 (0.878–6.483)	0.101
CTR	77.72 (24.92–294.2)	<0.001	805.6 (5.024–3.611)	0.018
Shape	2.557 (0.921–10.62)	0.119		
Boundary	1.065 (0.767–1.500)	0.710		
Vacuole sign	1.104 (0.766–1.563)	0.585		
Cavity	1.780 (0.820–3.547)	0.119		
Calcification	1.118 (0.171–4.278)	0.886		
Spiculated sign	2.592 (1.853–3.697)	<0.001	1.131 (0.757–1.712)	0.552
Lobulated sign	6.239 (2.981–16.00)	<0.001	1.557 (0.648–4.391)	0.357
Air bronchogram	0.664 (0.452–0.953)	0.031	0.832 (0.538–1.265)	0.399
Vessel convergence	5.154 (2.965–9.868)	<0.001	1.507 (0.763–3.204)	0.259
Pleural indentation	2.265 (1.580–3.328)	<0.001	9.667 (0.633–1.500)	0.877

Note. — CI = confidence interval, CTR = consolidation tumor ratio, OR = odds ratio.

Fig. 3 illustrates the ROC curves of the three models for three sets for LVI prediction. However, irrespective of the cohort, DeLong's test indicated that the models had relatively similar AUCs ($p > 0.05$) (Supplementary Table E4). Fig. 4 shows that the output from all three models was significantly associated with LVI (all $p < 0.05$), supporting the reliability of the results in predicting LVI. We then comprehensively evaluated the accuracy, sensitivity, specificity, PPV, NPV, and F1 scores of the three models. The details are shown in Table 3. These results indicate that DL features and clinical characteristics complement each other. The calibration curves are provided in supplementary Figure E1.

Fig. 5 presents the clinicopathological information (A), CT images (B), and heatmaps (C) of both LVI-positive and -negative patients with comparable pathological staging. However, LVI-positive patients with bone metastases had significantly shorter DFS than LVI-negative patients without recurrence or metastasis. The heatmaps indicated differences between tumors with LVI-positivity and LVI-negativity. The areas that were most associated with LVI prediction were the interface between the tumor and its adjacent parenchyma, and the tumor region contacting the adjacent pleura. The DL features extracted from the images were thus helpful in identifying the LVI status, assisting further in the risk stratification of the patients.

Survival prediction

In total, 591 participants were assessed in the survival comparisons. The median DFS duration of the participants was 66.6 months, with 163 of 591 (27.6 %) patients showing poor outcomes after surgery. Participants were followed up for a median of 68.7 months. The univariate analysis indicated that DFS was associated with age, density, CTR, clinical T stage, lymph node metastasis, VPI, pathologic LVI, and LVI predicted by both the DL and combined models. After multivariate analysis, age, CTR, clinical T1c stage, pathologic LVI, LVI predicted by both DL model and combined model were independent predictors of DFS. The results of both univariate and multivariate analyses of three cohorts are provided in Tables 4 and 5 while the survival curves are illustrated in Fig. 6. The Kaplan–Meier analysis found that the pathological LVI status, LVI predicted by the DL model, and the combined model output and were effective predictors of patient mortality ($p < 0.05$ for all). However, LVI predicted by the clinical model could not be used as a predictor of DFS.

The median DFS of patients with and without pathological LVI was 77.4 and 99.1 months, respectively ($p < 0.001$) (Fig. 6a), with corresponding predicted 5-year survival rates of 70.2 % and 81.2 %. Patients were additionally assigned to low- and high-risk groups using the clinical, DL, and combined models. The median DFS of patients predicted to be high- and low-risk using the clinical model was 92.4 and 95.8 months, respectively ($p = 0.53$) (Fig. 6b), with corresponding predicted 5-year survival rates of 78.2 % and 80.0 %. As such, LVI status as predicted using the clinical model was not independently predictive of survival in clinical stage IA lung adenocarcinoma patients. The median DFS of patients predicted to be high- and low-risk using the DL model was 87.7 and 103.2 months, respectively ($p = 0.01$) (Fig. 6c), with corresponding 5-year survival rates of 74.0 % and 84.4 %. The median DFS of patients predicted to be high- and low-risk using the combined model, the median DFS was 82.4 and 94.5 months, respectively ($p < 0.001$) (Fig. 6d), with corresponding 5-year survival rates of 75.2 % and 86.7 %. Kaplan-

Table 3
Prediction Performance of the Three Models

Models	Dataset	AUC (95 % CI)	ACC	SEN	SPE	PPV	NPV	F1 score
Clinical model	Training set	0.74 (0.73–0.76)	0.62	0.79	0.59	0.26	0.94	0.39
	Internal validation set	0.71 (0.68–0.75)	0.63	0.76	0.61	0.28	0.94	0.40
	External validation set	0.71(0.61–0.81)	0.62	0.88	0.60	0.13	0.99	0.22
DL model	Training set	0.76 (0.75–0.78)	0.64	0.79	0.63	0.28	0.94	0.42
	Internal validation set	0.72 (0.68–0.76)	0.63	0.79	0.60	0.27	0.94	0.40
	External validation set	0.63(0.54–0.73)	0.50	0.86	0.48	0.09	0.98	0.16
Combined model	Training set	0.76 (0.74–0.77)	0.63	0.81	0.60	0.27	0.95	0.40
	Internal validation set	0.74 (0.71–0.78)	0.68	0.77	0.60	0.31	0.94	0.43
	External validation set	0.77(0.70–0.84)	0.72	0.84	0.70	0.15	0.98	0.26

Note. — ACC = accuracy, AUC = area under the receiver operating characteristic curve, CI = confidence interval, DL = deep learning, NPV = negative predictive value, PPV = positive predictive value, SEN = sensitivity, SPE = specificity.

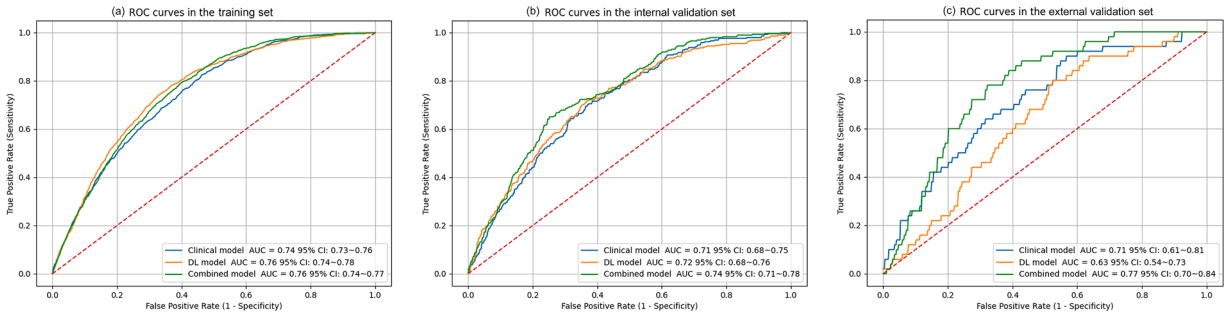

Fig. 3. Predictive performance of the DL signature in the identification of LVI. ROC curves of the three models for the prediction of lymphovascular invasion in the training set (a), internal validation set (b), and external validation set (c). AUC = area under the curve, CI = confidence interval, DL = deep learning, ROC = receiver operating characteristic.

Meier analyses indicated that both the combined and DL models were capable of readily identifying patients facing a higher risk of mortality.

Discussion

Here, the effectiveness of a DL signature derived from 3D-ResNet-9 of CT was investigated for predicting the outcomes of patients with clinical stage IA lung adenocarcinoma involving LVI. It was found that the combined model integrating DL with clinical-radiological characteristics achieved the best performance (AUCs = 0.76, 0.74 and 0.77 for the training, internal, and external validation sets, respectively). Despite a lack of significant advantage of the combined model over the DL and clinical models, the DL signature showed comparable predictive efficacy with clinical-radiological characteristics. Therefore, even the use of the DL model alone provided a simple and more accurate prediction of LVI from CT. High-risk status as predicted with the DL model (HR for DFS, 2.396 [$p = 0.009$]) or the combined model (HR for DFS, 2.439 [$p = 0.006$]) and the presence of pathologic LVI (HR for DFS, 1.949 [$p < 0.001$]) were all found to be related to poor prognostic outcomes. Furthermore, LVI status as predicted by both the DL and combined models was independently associated with patient DFS in multivariate analyses, i.e., shortened DFS [11–13], which was consistent with our findings.

Primary tumors are connected to lymph nodes through peritumoral lymphatic vessels and arteriovenous connections, which facilitate metastasis [28,29]. Thus, early-stage lung cancer accompanied by LVI-positivity requires consideration in terms of appropriate surgical scope [4,5,28,29]. Although the benefits of adjuvant chemotherapy are now recognized for stage II or III NSCLC, there is no consensus on its use for stage IA [30]. The present study aimed to predict the LVI status preoperatively, which may provide evidence of appropriate therapeutic regimens. This study appears to be the first investigation to incorporate ResNet features with preoperative CT to determine the LVI status. Notably, this DL model provides a simpler means of identifying patients with a high risk of tumor recurrence even in stage IA disease, thus providing information on the biology of the primary tumor and further information for accurate prediction of prognosis. Beck et al developed a 3D convolutional neural network using a transfer learning algorithm to predict LVI or nodal involvement (AUCs of 0.63–0.72) [31], which differs from the DL features used in this study that only focused on the LVI status. The present results allow the excavation of hidden features that are not available to clinical examinations and imaging, thus providing a more accurate interpretation of the relationship between CT imaging and LVI status.

Recent studies have shown that the lung parenchyma surrounding the primary tumor may be involved in tumor invasion and metastasis [32,33]. Zuo et al [34] and Yang et al [35] concentrated on CT intratumor features associated with LVI and outcomes. On the basis of the previous two studies, Chen et al extracted CT-based radiomics features from 145 patients, and found that the gross tumor volume incorporating peritumoral regions 9 mm from the tumor can predict LVI in NSCLC (AUCs of 0.67–0.82) [36]. However, these findings required further verification with large samples. Fig. 4 suggests the possibility that additional factors not visible to the naked eye may also affect tumor invasiveness at the tumor-parenchyma and tumor-pleura interfaces. This investigation used a large sample of patients from two centers, and the model was found to be both robust and generalizable. We believe that such an approach integrating DL models and available clinical-radiological information can be applied to the development of similar models in medicine.

The multivariate analysis demonstrated that preoperative CTR could independently predict LVI, as observed by a previous study [37,38]. CTR has also been observed to predict prognosis following surgery or radiation treatment for NSCLC, as well as having the ability to predict lymph node metastasis and VPI. The findings of the long-term JCOG 0201 trial indicated that limited surgery could result in satisfactory outcomes in patients with predominantly GGO lung tumors with CTR values of 0.5 or less and sizes between 2 and 3 cm [39]. Thus, higher CTR values are indicative of increased proliferation and invasiveness of the cancer cells, together with an increased risk of LVI development.

There has been intense emphasis on the part played by LVI in NSCLC in the TNM staging manual. Several studies have demonstrated that LVI adversely impacts survival in NSCLC [40,41]. Here, LVI was observed to be an independent risk factor for DFS in univariate regression analysis. LVI was then used to stratify the patients and the predictive ability of three models for DFS was investigated. The LVI predictions made by DL and combined models were also found to be independently associated with DFS in multivariate analyses. Also, the Kaplan-Meier analysis found that the output of DL and combined models clearly separated the lung adenocarcinoma patients into high- and low-risk mortality groups. The results of prognostic analysis in this study were consistent with previous study [30,32]. However, LVI predicted by the clinical model is not a prognostic predictor. This indicates that the DL signature, in comparison with clinical and radiological features will assist in the assessment of the survival risk in individual patients, as well as providing a reference for appropriate treatment.

In recent years, the application of DL in medical imaging has made significant advances in the field of tumor evaluation, including tumor diagnosis, clinical grading and staging, genetic analyses, efficacy evaluation, and prognosis prediction [19-21,42-44]. This study concentrated on the prediction of LVI status in clinical stage IA lung adenocarcinoma through the deep mining of high-throughput, noninvasive, and comprehensive features in CT images, which reflect the spatial heterogeneity of the lesions and have important prognostic value for disease treatment. Currently, the main challenge faced by DL models is that of biological interpretability. Traditional manual radiomics features have complete formulas and definitions that are closely related to the semantic features of lesions described in medical imaging diagnosis. Therefore, these can be used as an approximate explanation of the potential biological significance of radiomics features. However, DL

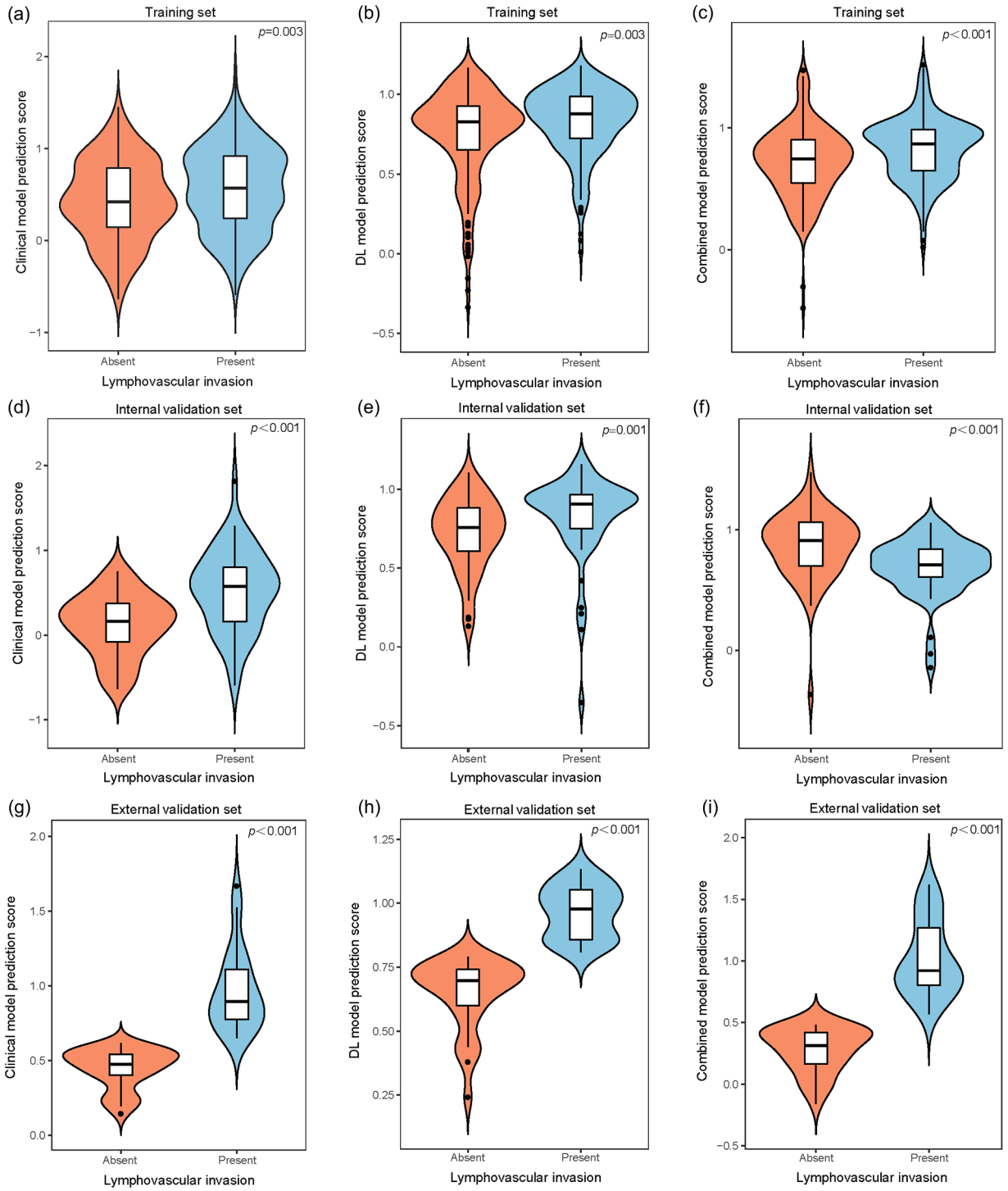

Fig. 4. Difference in the DL signature between Clinical Stage IA lung adenocarcinoma patients with LVI and without LVI. Boxplots showing the lymphovascular invasion outputs of the clinical model (a, d, g), DL model (b, e, h), and combined model (c, f, i) in the training set (a–c), internal validation set (d–f), and external validation set (g–i). The outputs of all three models were significantly associated with lymphovascular invasion (all $p < 0.05$). DL = deep learning, LVI = lymphovascular invasion.

features represent a "black box" [45] as they lack accurate and complete formulas and definitions, making them difficult to interpret. A common method used is their reverse interpretation using DL features in heatmaps which can provide information on the clinical significance of specific regions of the model. The model output of the current study is shown in the heatmap in Fig. 5, this presentation solves the problem of the biological interpretability of the model to a certain extent. However, the heatmap cannot truly meet the requirements of biological interpretability due to the large amount of error. In addition, some imaging features with high predictive ability may be association with increased expression of specific genes or proteins. The biological interpretability of DL models could thus be further improved by exploring the relationships between these genes, proteins and clinical endpoint events [46]. In the future, more research on the application of DL in medical imaging can address and resolve the problem of biological interpretability. We believe that an efficient and generalized artificial

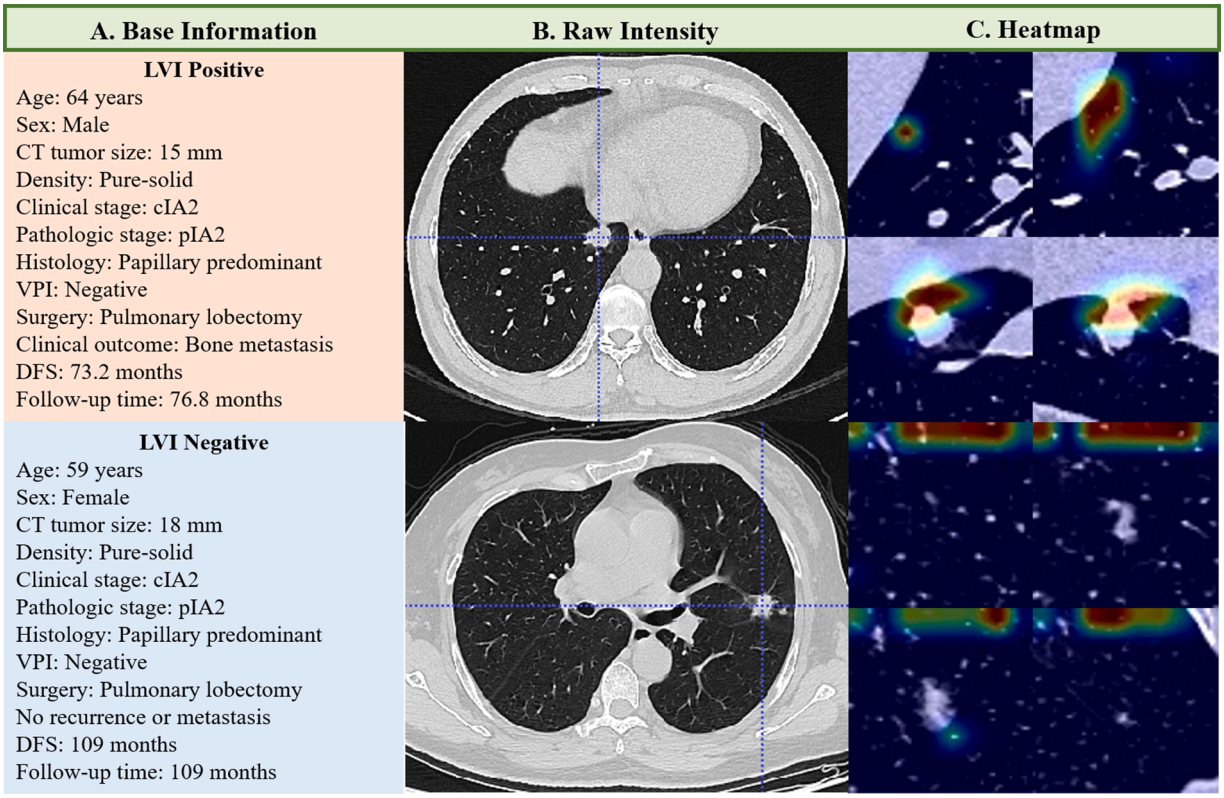

Fig. 5. Gradient-weighted class activation heatmaps. Clinicopathological information (A), CT images (B), and heatmaps (C) of two typical LVI-positive and LVI-negative cases. Although, apart from LVI status, the clinicopathological information was similar, LVI-positive patients with bone metastasis had significantly shorter DFS than LVI-negative patients without recurrence or metastasis. The heatmaps indicate that the areas most associated with LVI prediction were the tumor-parenchyma and tumor-pleura interfaces.

Table 4
Univariate Cox Regression Analyses of Disease-Free Survival in the Training Set Comprising 591 Patients with Clinical Stage IA Lung Adenocarcinoma

Variable	Univariate analysis	
	HR (95 %CI)	*p* value
Age	1.017(1.000,1.034)	0.044
Sex (female)	0.801(0.588,1.090)	0.158
Smoking history (ever smoker)	1.245(0.905,1.714)	0.178
Family history of lung cancer	0.818(0.530,1.261)	0.363
Density (Pure-solid)	2.096(1.283,3.423)	0.003
CTR	21.88(4.824,99.25)	<0.001
Clinical T stage		<0.001
cT1b	3.839(1.664,8.857)	0.002
cT1c	6.545(2.862,14.97)	<0.001
Lymph node metastasis	2.916(2.072,4.103)	<0.001
VPI	1.618(1.181,2.216)	0.003
Pathologic LVI	1.949(1.375,2.763)	<0.001
Clinical-model-predicted LVI	1.225(0.876,1.713)	0.236
DL-model-predicted LVI	2.396(1.247,4.603)	0.009
Combined-model-predicted LVI	2.439(1.860,4.704)	0.006

Note. — CI = confidence interval, DL = deep learning, HR = hazard ratio, LVI = lymphovascular invasion, VPI = visceral pleural invasion.

intelligence system for disease diagnosis and treatment, integrating new technologies and theories, will result in the accurate diagnosis and evaluation of diseases.

The study has several limitations. First, the study was retrospective which may have led to possible selection bias. It is possible that the positive rate of LVI positivity among the participants may not reflect the true situation. The inclusion and exclusion criteria were applied strictly to ensure the definite inclusion of LVI in the study. Future prospective studies should be conducted to further validate the model. Second, the combined model was not found to be superior to either the clinical or DL models. This may be because the combined model faces the problem of multi-modal data fusion. Therefore, the combined model did not show superiority over the clinical and imaging data of a single modality. Future studies should address the issue of multimodal data fusion with further model adjustments. Third, the study used CT images acquired from different machines, with different manufacturers, and using different protocols. Standardization of the imaging data, as far as possible, would allow greater generalization of the model. However, data standardization is a difficult problem that will require future solution. Nonetheless, we consider that this heterogeneity in CT images is an accurate reflection of real-world clinical data, and thus a trained DL system based on such data would be more applicable to real-world clinical practice. Fourth, the data used for survival analysis were derived from a relatively small number of cases as the follow-up time was limited to five years. Further verification is required from large-sample studies in the future.

Conclusion

In conclusion, this DL algorithm using 3D-ResNet-9 based on CT images offers convenient and potentially useful assistance for the preoperative prediction of LVI in patients with clinical stage IA lung adenocarcinoma. The integrated model combining multimodal imaging data with clinical and radiological data may thus be useful for the preoperative prediction of LVI. This model may also assist with preoperative clinical decision-making.

CRediT authorship contribution statement

Kunfeng Liu: Conceptualization, Data curation, Methodology, Software, Visualization, Writing – original draft. **Xiaofeng Lin:**

Table 5
Multivariate Cox Regression Analyses of Disease-Free Survival in Patients with Clinical Stage IA Lung Adenocarcinoma

	Pathologic LVI		Clinical-model-predicted LVI		DL-model-predicted LVI		Combined-model-predicted LVI	
Variable	HR (95 %CI)	*p* value	HR (95 %CI)	*p* value	HR (95 %CI)	*p* value	HR (95 %CI)	*p* value
Age	1.015(0.998,1.032)	0.078	1.014(0.997,1.031)	0.097	1.014(0.997,1.031)	0.098	1.014(0.998,1.031)	0.095
Density (Pure-solid)	1.427(0.857,2.376)	0.171	1.470(0.880,2.457)	0.141	1.463(0.881,2.430)	0.142	1.526(0.902,2.582)	0.115
CTR	4.865(0.892,26.53)	0.068	5.308(0.970,29.04)	0.054	5.265(0.961,28.86)	0.056	5.236(0.957,28.65)	0.056
Clinical T stage		0.006		0.002		0.003		0.001
cT1b	2.270(0.938,5.493)	0.069	2.328(0.963,5.632)	0.061	2.304(0.953,5.571)	0.064	2.350(0.971,5.684)	0.058
cT1c	3.359(1.369,8.240)	0.008	3.677(1.502,9.000)	0.004	3.580(1.450,8.841)	0.006	3.772(1.538,9.248)	0.004
LVI	1.454(1.014,2.083)	0.042	0.972(0.685,1.381)	0.875	4.081(1.353,8.115)	0.019	4.837(1.476,8.472)	0.038

Note. — CI = confidence interval, DL = deep learning, HR = hazard ratio, LVI = lymphovascular invasion

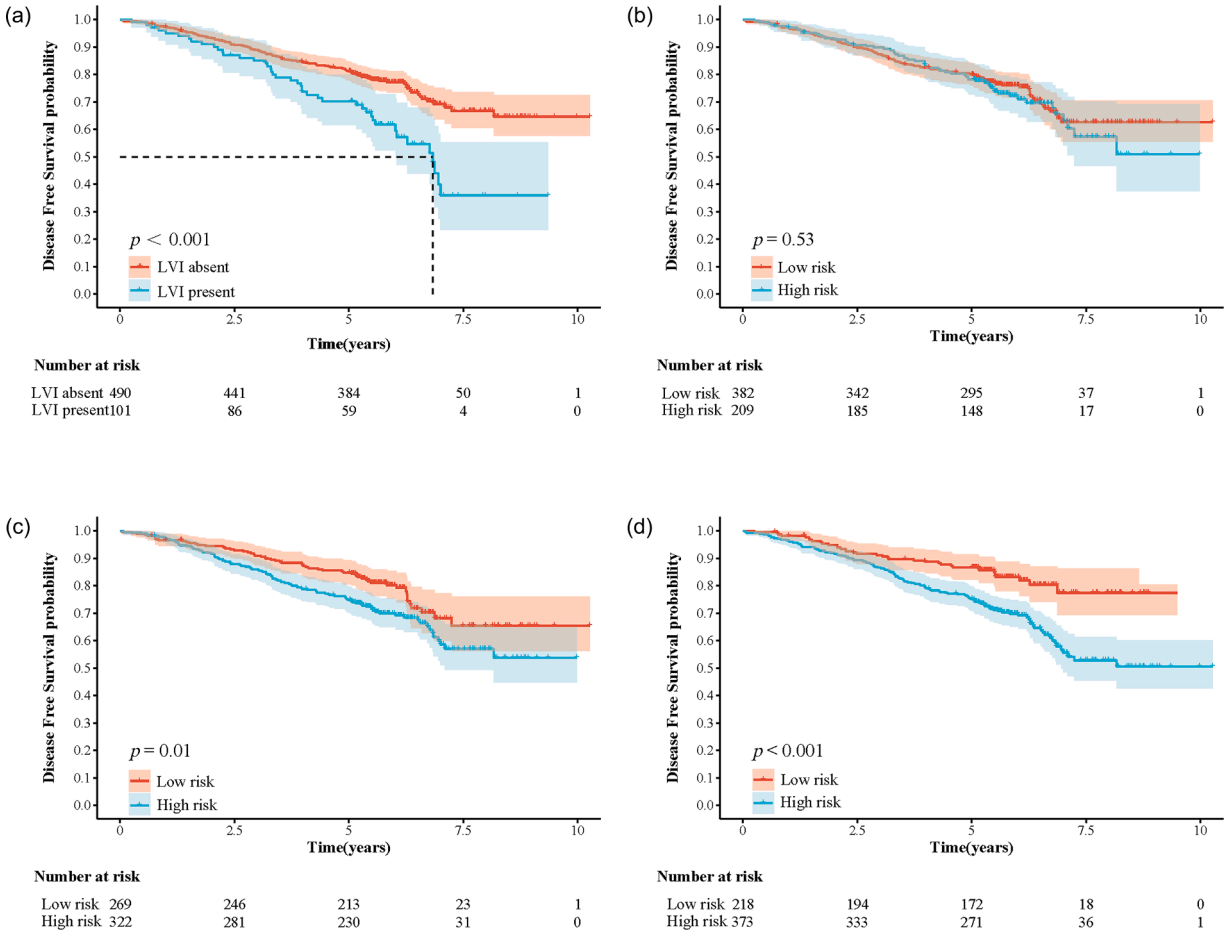

Fig. 6. Prognostic value of the DL signature in Clinical Stage IA lung adenocarcinoma. Kaplan-Meier survival curves according to pathological LVI status (a), or their LVI status as predicted by the clinical model (b), DL model (c), and combined model (d). DFS = disease-free survival, DL = deep learning, LVI = lymphovascular invasion.

Conceptualization, Data curation, Methodology, Writing – original draft. **Xiaojuan Chen:** Conceptualization, Data curation, Methodology. **Biyun Chen:** Data curation. **Sheng Li:** Investigation, Supervision. **Kunwei Li:** Visualization, Writing – review & editing. **Huai Chen:** Validation, Writing – review & editing. **Li Li:** Investigation, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We sincerely thank Han Jiao from the School of Electronics and Information, Sun Yat-sen University for providing technical support.

Funding

This study has received funding from Guangdong-Hongkong-Macao Applied Mathematics Center project of Guangdong Basic and Applied Basic Research Foundation (Grant N0. 2021B1515310002) and ZHONGNANSHAN Medical Foundation of Guangdong Province (Grant No. ZNSXS-20230001).

Availability of data and materials

All data generated during this study are included in this published article. Further inquiry data can be obtained by the corresponding author.

Ethics approval and consent to participate

This study received ethical approval from the Sun Yat-sen University Cancer Center and the First Affiliated Hospital of Guangzhou Medical University (Ethics approval No. B2022-293-01), and the Fifth Affiliated Hospital of Sun Yat-sen University (Ethics approval No. K107-1). And the written informed consent requirement was waived.

Consent for publication

Not applicable.

Cover letter

Lymphovascular invasion (LVI) in patients with early-stage lung cancer is often associated with poor prognosis, as well as having implications for patient treatment and clinical outcomes. While most studies have focused on the assessment of assessing LVI-related morphological characteristics in CT images, such assessments are not always accurate as the identification of specific features is strongly dependent on the expertise of the radiologist. Deep learning (DL) algorithms allow automated characterization of data, such that imaging data associated with particular regions of interest can be transformed into high-resolution spatial data from which features can be extracted and analyzed. To date, few studies have sought to apply DL as a means of predicting LVI status. This study was designed with the goal of establishing and validating a CT-based DL model capable of predicting LVI status and stratifying patients based on their prognostic outcomes. The analysis showed that the model outputs were independently predictive of both LVI status and disease-free survival in patients with clinical stage IA lung adenocarcinoma, and allowed the identification of latent characteristics underlying clinical assessments and semantic features.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.tranon.2024.101894.

References

[1] RL Siegel, KD Miller, HE Fuchs, A Jemal, Cancer Statistics, CA Cancer J. Clin. 71 (1) (2021) 7–33.
[2] W Yang, F Qian, J Teng, et al., Written on behalf of the AME Thoracic Surgery Collaborative Group, Community-based lung cancer screening with low-dose CT in China: results of the baseline screening, Lung Cancer 117 (2018) 20–26.
[3] A Lackey, JS Donington, Surgical management of lung cancer, Semin. Intervent. Radiol. 30 (2) (2013) 133–140.
[4] JS Donington, YT Kim, B Tong, et al., Progress in the management of early-stage non-small cell lung cancer in 2017, J. Thorac. Oncol. 13 (6) (2018) 767–778.
[5] OV Khullar, Y Liu, T Gillespie, et al., Survival after sublobar resection versus lobectomy for clinical stage IA lung cancer: an analysis from the national cancer data base, J. Thorac. Oncol. 10 (11) (2015) 1625–1633.
[6] P Goldstraw, K Chansky, J Crowley, et al., International Association for the Study of Lung Cancer Staging and Prognostic Factors Committee, Advisory Boards, and Participating Institutions, International Association for the Study of Lung Cancer Staging and Prognostic Factors Committee Advisory Boards and Participating Institutions, The IASLC Lung Cancer Staging Project: Proposals for Revision of the TNM Stage Groupings in the Forthcoming (Eighth) Edition of the TNM Classification for Lung Cancer, J. Thorac. Oncol. 11 (1) (2016) 39–51.
[7] A Rizzo, A Cusmai, F Giovannelli, et al., Impact of proton pump inhibitors and histamine-2-receptor antagonists on non-small cell lung cancer immunotherapy: a systematic review and meta-analysis, Cancers (Basel) 14 (6) (2022) 1404.
[8] M Santoni, A Rizzo, V Mollica, et al., The impact of gender on the efficacy of immune checkpoint inhibitors in cancer patients: The MOUSEION-01 study, Crit. Rev. Oncol. Hematol. 170 (2022) 103596.
[9] A. Rizzo, Identifying optimal first-line treatment for advanced non-small cell lung carcinoma with high PD-L1 expression: a matter of debate, Br. J. Cancer 127 (8) (2022) 1381–1382.
[10] V Mollica, A Rizzo, A Marchetti, et al., The impact of ECOG performance status on efficacy of immunotherapy and immune-based combinations in cancer patients: the MOUSEION-06 study, Clin. Exp. Med. (2023). Epub ahead of print.
[11] T. Kinoshita, T. Ohtsuka, M. Yotsukura, et al., Prognostic impact of preoperative tumor marker levels and lymphovascular invasion in pathological stage I adenocarcinoma and squamous cell carcinoma of the lung, J Thorac Oncol 10 (4) (2015) 619–628.
[12] L. Okiror, L. Harling, L. Toufektzian, et al., Prognostic factors including lymphovascular invasion on survival for resected non-small cell lung cancer, J. Thorac. Cardiovasc. Surg. 156 (2) (2018) 785–793.
[13] D. Noma, K. Inamura, Y. Matsuura, et al., Prognostic effect of lymphovascular invasion on TNM staging in stage I non-small-cell lung cancer, Clin. Lung Cancer 19 (1) (2018) e109–e122.
[14] S Wang, J Xu, R Wang, et al., Adjuvant chemotherapy may improve prognosis after resection of stage I lung cancer with lymphovascular invasion, J. Thorac. Cardiovasc. Surg. 156 (5) (2018), 2006–2015.e2.
[15] Y Tsutani, Y Miyata, K Kushitani, et al., Propensity score-matched analysis of adjuvant chemotherapy for stage I non-small cell lung cancer, J. Thorac. Cardiovasc. Surg. 148 (2014) 1179–1185.
[16] JK Yun, GD Lee, S Choi, et al., Comparison of prognostic impact of lymphovascular invasion in stage IA non-small cell lung cancer after lobectomy versus sublobar resection: a propensity score-matched analysis, Lung Cancer 146 (2020) 105–111.
[17] K Grbic, B Mehic, D Udovicic-Gagula, et al., Preoperative tumour size as a predictor of the presence of lymphovascular invasion in lung adenocarcinoma, Med Glas (Zenica) 17 (2) (2020) 363–368.
[18] K Suzuki, T Koike, T Asakawa, et al., A prospective radiological study of thin-section computed tomography to predict pathological noninvasiveness in peripheral clinical IA lung cancer (Japan Clinical Oncology Group 0201), J. Thorac. Oncol. 6 (4) (2011) 751–756.
[19] D Shen, G Wu, HI Suk, Deep learning in medical image analysis, Annu. Rev. BioMed. Eng. 19 (2017) 221–248.
[20] N Nasrullah, J Sang, MS Alam, M Mateen, B Cai, H Hu, Automated lung nodule detection and classification using deep learning combined with multiple strategies, Sensors (Basel) 19 (17) (2019) 3722.
[21] JL Causey, J Zhang, S Ma, et al., Highly accurate model for prediction of lung nodule malignancy with CT scans, Sci. Rep. 8 (1) (2018) 9286.
[22] X Lin, H Jiao, Z Pang, et al., Lung cancer and granuloma identification using a deep learning model to extract 3-dimensional radiomics features in CT imaging, Clin. Lung Cancer 22 (5) (2021) e756–e766.
[23] MB Amin, SB Edge, FL Greene, AJCC Cancer Staging Manual, 8th ed., Springer, New York, 2017.
[24] J Zugazagoitia, AB Enguita, JA Nuñez, L Iglesias, S Ponce, The new IASLC/ATS/ERS lung adenocarcinoma classification from a clinical perspective: current concepts and future prospects, J. Thorac. Dis. 6 (Suppl 5) (2014) S526–S536.
[25] H Kaiming, Z Xiangyu, R Shaoqing, S Jian, Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770–778.
[26] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, et al., An image is worth 16x16 words: transformers for image recognition at scale, ICLR (2021).
[27] RR Selvaraju, M Cogswell, A Das, R Vedantam, D Parikh, D Batra, GradCAM: visual explanations from deep networks via gradient-based localization, in: IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017, pp. 618–626.
[28] K Vaahtomeri, K Alitalo, Lymphatic vessels in tumor dissemination versus immunotherapy, Cancer Res. 80 (17) (2020) 3463–3465.
[29] T. Sato, Y. Shimada, T. Mimae, et al., The impact of pathological lymph node metastasis with lymphatic invasion on the survival of patients with clinically node-negative non-small cell lung cancer: a multicenter study, Lung Cancer 158 (2021) 9–14.
[30] JP Pignon, H Tribodet, GV Scagliotti, et al., Lung adjuvant cisplatin evaluation: a pooled analysis by the LACE collaborative group, J. Clin. Oncol. 26 (21) (2008) 3552–3559.
[31] K S Beck, B Gil, S J Na, et al., DeepCUBIT: predicting lymphovascular invasion or pathological lymph node involvement of clinical T1 stage non-small cell lung cancer on chest CT scan using deep cubical nodule transfer learning algorithm, Front. Oncol. 11 (2021) 661244.
[32] T Saijo, G Ishii, A Ochiai, et al., Evaluation of extratumoral lymphatic permeation in non-small cell lung cancer as a means of predicting outcome, Lung Cancer 55 (1) (2017) 61–66.
[33] X Wang, X Zhao, Q Li, et al., Can peritumoral radiomics increase the efficiency of the prediction for lymph node metastasis in clinical stage T1 lung adenocarcinoma on CT? Eur. Radiol. 29 (11) (2019) 6049–6058.
[34] Z Zuo, X Fan, Y Tang, et al., Deep learning-powered 3D segmentation derives factors associated with lymphovascular invasion and prognosis in clinical T1 stage non-small cell lung cancer, Heliyon 9 (4) (2023) e15147.
[35] G Yang, P Nie, L Zhao, et al., 2D and 3D texture analysis to predict lymphovascular invasion in lung adenocarcinoma, Eur. J. Radiol. 129 (2020) 109111.
[36] Q Chen, J Shao, T Xue, et al., Intratumoral and peritumoral radiomics nomograms for the preoperative prediction of lymphovascular invasion and overall survival in non-small cell lung cancer, Eur. Radiol. 33 (2) (2023) 947–958.
[37] M Shimomura, M Iwasaki, S Ishihara, et al., Volume-based consolidation-to-tumor ratio is a useful predictor for postoperative upstaging in stage I and II lung adenocarcinomas, Thorac. Cardiovasc. Surg. 70 (3) (2022) 265–272.

[38] K Grbic, B Mehic, D Udovicic-Gagula, et al., Preoperative tumour size as a predictor of the presence of lymphovascular invasion in lung adenocarcinoma, Med Glas (Zenica) 17 (2) (2020) 363–368.
[39] H Asamura, T Hishida, K Suzuki, et al., Radiographically determined noninvasive adenocarcinoma of the lung: Survival outcomes of japan clinical oncology group 0201, J. Thorac. Cardiovasc. Surg. 146 (1) (2013) 24–30.
[40] S Wang, B Zhang, J Qian, et al., Proposal on incorporating lymphovascular invasion as a T-descriptor for stage I lung cancer, Lung Cancer 125 (2018) 245–252.
[41] D Noma, K Inamura, Y Matsuura, et al., Prognostic effect of lymphovascular invasion on TNM staging in stage I non-small-cell lung cancer, Clin. Lung Cancer 19 (2018) e109–ee22.
[42] A Esteva, B Kuprel, RA Novoa, et al., Dermatologist-level classification of skin cancer with deep neural networks, Nature 542 (7639) (2017) 115–118.
[43] X Li, S Zhang, Q Zhang, et al., Diagnosis of thyroid cancer using deep convolutional neural network models applied to sonographic images: a retrospective, multicohort, diagnostic study, Lancet Oncol. 20 (2) (2019) 193–201.
[44] P Rajpurkar, J Irvin, RL Ball, et al., Deep learning for chest radiograph diagnosis: a retrospective comparison of the CheXNeXt algorithm to practicing radiologists, PLoS Med. 15 (11) (2018) e1002686.
[45] WL Bi, A Hosny, MB Schabath, et al., Artificial intelligence in cancer imaging: clinical challenges and applications, CA Cancer J. Clin. 69 (2019) 127–157.
[46] K Pinker, J Chin, AN Melsaether, et al., Precision medicine and radiogenomics in breast cancer: new approaches toward diagnosis and treatment, Radiology 287 (2018) 732–747.