Heliyon 10 (2024) e34863

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Estimation of pathological subtypes in subsolid lung nodules using artificial intelligence

Xiaoqin Hu [a], Liu Yang [b], Tong Kang [a], Hanhua Yu [a], Tingkuan Zhao [c], Yuanyi Huang [d,**], Yuefeng Kong [a,*]

[a] Department of Radiology, The Fourth Hospital of Wuhan, Wuhan, China
[b] Department of Radiology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, China
[c] Department of Pathology, Jingzhou Central Hospital, The Second Clinical Medical College, Yangtze University, Jingzhou, China
[d] Department of Radiology, Jingzhou Central Hospital, The Second Clinical Medical College, Yangtze University, Jingzhou, China

ARTICLE INFO

Keywords:
Adenocarcinomas
Artificial intelligence (AI)
Lung neoplasms
Pulmonary nodule
Subsolid nodules

ABSTRACT

Objective: This study aimed to investigate the value of artificial intelligence (AI) for distinguishing pathological subtypes of invasive pulmonary adenocarcinomas in patients with subsolid nodules (SSNs).

Materials and methods: This retrospective study included 110 consecutive patients with 120 SSNs. The qualitative and quantitative imaging characteristics of SSNs were extracted automatically using an artificially intelligent assessment system. Then, radiologists had to verify these characteristics again. We split all cases into two groups: non-IA including 11 Atypical adenomatous hyperplasia (AAH) and 25 adenocarcinoma in situ (AIS) or IA including 7 minimally invasive adenocarcinoma (MIA) and 77 invasive adenocarcinoma (IAC). Variables that exhibited statistically significant differences between the non-IA and IA in the univariate analysis were included in the multivariate logistic regression analysis. Receiver operating characteristic (ROC) analyses were conducted to determine the cut-off values and their diagnostic performances.

Results: Multivariate logistic regression analysis showed that the major diameter (odds ratio [OR] = 1.38; 95 % confidence interval [CI], 1.02–1.87; P = 0.036) and entropy of three-dimensional (3D) CT value (OR = 3.73, 95 % CI, 1.13–2.33, P = 0.031) were independent risk factors for adenocarcinomas. The cut-off values of the major diameter and the entropy of 3D CT value for the diagnosis of invasive adenocarcinoma were 15.5 mm and 5.17, respectively. To improve the classification performance, we fused the major diameter and the entropy of 3D CT value as a combined model, and the (AUC) of the model was 0.868 (sensitivity = 0.845, specificity = 0.806).

Conclusion: The major diameter and entropy of 3D CT value can distinguish non-IA from IA. AI can improve performance in distinguishing pathological subtypes of invasive pulmonary adenocarcinomas in patients with SSNs.

* Corresponding author.
** Corresponding author.
E-mail addresses: yyhuangjz@163.com (Y. Huang), kongyuefeng1990@163.com (Y. Kong).

https://doi.org/10.1016/j.heliyon.2024.e34863
Received 3 February 2024; Received in revised form 17 July 2024; Accepted 17 July 2024
Available online 19 July 2024

1. Introduction

Lung cancer remained the leading cause of cancer mortality in 2020, accounting for approximately one in five (18.0 %) deaths, with an estimated 1.8 million fatalities [1]. A relatively high proportion of subsolid nodules (SSNs), including pure ground glass nodules (pGGNs) and part-solid nodules (PSNs), was detected in lung cancer CT screening [2,3]. Although most SSNs are transient, persistent SSNs, especially PSNs, can be a specific type of lung adenocarcinoma or its precursors [4–6]. Atypical adenomatous hyperplasia (AAH) and adenocarcinoma in situ (AIS) were included in the revised 2021 categorization of thoracic tumors by the WHO as precursor glandular lesions, whereas minimally invasive adenocarcinoma (MIA) and invasive adenocarcinoma (IAC) were categorized as adenocarcinomas [7]. The probability of patients with AIS or MIA being recurrence-free for 5 years postoperatively is 100 %, and the risk of recurrence is also relatively low >5 years after resection of AIS or MIA [8]. The high long-term survival of MIA suggests that patients with AAH or AIS might not benefit from surgical procedures. Therefore, patients with AAH or AIS only need to receive conservative treatment and require surgical resection rarely.

AI has been applied to detect various features of lung nodules images [9]. Meanwhile, AI has also performed well in differentiating benign and malignant lesions [10,11]. However, the density contrast between the SSNs and the adjacent lung tissue is relatively small, making it difficult to evaluate the benign and malignant lesions clinically. Therefore, early identification of adenocarcinomas in patients with SSNs using non-invasive CT imaging characteristics based on AI would help physicians in clinical decision-making.

In this study, we used the AI evaluation system constructed by Hangzhou YITU Healthcare Technology Co., Ltd as the CT image analysis tool, and analyzed the value of AI in distinguishing pathological subtypes of invasive lung adenocarcinomas presenting as SSNs. This will help better patient management by identifying lung adenocarcinomas that require surgical removal and may require follow-up.

2. Materials and methods

2.1. Patients

Our institutional review board approved this retrospective study. All patients provided written informed permission prior to undergoing the pathological test.

Patients assessed for SSNs at our institution between January 2017 and June 2021 were eligible for this study. The inclusion criteria were as follows: (1) No preoperative chemotherapy, radiation, or other treatments were given in any of the cases. (2) Patients received CT examination before surgery with a diagnosis of SSNs, and we were able to evaluate the CT imaging data using AI software. (3) The major diameter of SSNs less than or equal to 3 cm. (4) All cases were identified through postoperative pathology. The another exclusion criteria were as follows: (1) The patient's electronic medical records were incomplete. (2) The poor image quality made it difficult for AI software to delineate the tumor boundary on CT scans. (3) Postoperative pathology results were not associated with lung adenocarcinoma or precursor glandular lesions.

A total of 110 patients with 120 SSNs, including 11 AAHs, 25 AISs, 7 MIAs, and 77 IACs, were studied (Table 1). We split all cases into two groups: non-IA (including 11 AAH and 25AIS) or IA (including 7 MIA and 77 IAC).The average age of patients was 52.95 ± 9.10 (27 ~ 72) year of age. Of the cases, 49 (44.5 %) were males and 56 (55.5 %) were females. A total of 22 (20.0 %) cases had a history of smoking and 7 (6.1 %) cases had a family history of lung cancer.

2.2. CT scanning

All patients underwent an unenhanced chest thin-section CT examination. The CT images were obtained on one of the following two scanners: the 64-layer Philip CT scanner (Holland) or the 16-layer Siemens CT scanner (Germany) with 120 kVp, 160 mA, pitch 1.2:1, collimation 40 mm, and FOV 35–36 cm. The conventional scanning layer has a thickness of 5 mm, layer spacing of 5 mm, and an

Table 1
Relationships between the clinical features and pathological diagnosis of 110 patients with GGNs.

Clinical features	Total ($N = 110$)	Non-IA group ($N = 33$)	IA group ($N = 77$)	t/χ^2	P value
Age (years)	52.95 ± 9.10 (27 ~ 72)	49.78 ± 10.56 (27 ~ 69)	54.76 ± 7.65 (36 ~ 72)	−2.617	0.011[a]
Gender				0.506	0.477[b]
Male	49(44.5 %)	13(39.4 %)	36(46.8 %)		
Female	61(55.5 %)	20(60.6 %)	41(53.2 %)		
Smoking history				0.693	0.405[b]
Present	22(20.0 %)	5(15.2 %)	17(22.1 %)		
Absent	88(80.0 %)	28(84.8 %)	60(77.9 %)		
Family history of lung cancer				0.588	0.426[c]
Present	7(6.1 %)	3(9.1 %)	4(5.2 %)		
Absent	103(93.6 %)	30(90.9 %)	73(94.8 %)		

Note. -Values are presented as mean ± SD (rang) or no. (%).
[a] : Student's t-test, b: χ2 tests, c: Fisher's exact test.

image matrix of 512 × 512. Images were reconstructed using the standard and lung algorithm reconstruction algorithms with a thickness of 1 mm and spacing of 1 mm. Images were observed in the pulmonary window (window width: 1500 HU; window: −500 HU).

Both scanning devices used the same pre-programs and reconstruction processes as for chest CT scans. Patients were trained in the fundamentals of inhaling and exhaling while lying down with their arms elevated. The patient underwent a whole-lung scan while holding his breath after taking a deep breath, and the scan was conducted from the thoracic inlet to the base of the lung.

2.3. Imaging analysis

The AI assessment system developed by Hangzhou YITU Healthcare Technology Co., Ltd. was used as the CT image-analysis tool. The system combined thresholding techniques with convolutional neural networks to separate the left and right lungs and detect pulmonary nodules and compute the qualitative and quantitative parameters automatically. Subsequently, radiologists reviewed the qualitative characteristics by applying multi-plane technology, window technology, and volume reconstruction. The qualitative variables included the type of nodules, spiculation, lobulation, vacuole sign, air bronchogram, and pleural indentation. The quantitative variables included the longest diameter, shortest diameter, lung nodules volume, and the minimum, maximum, mean, median, standard deviation, skewness, kurtosis, and entropy of the three-dimensional (3D) CT values (Figs. 1 and 2).

2.4. Pathological evaluation

All pathological specimens enrolled in the study were assessed individually by two senior pathologists, and the final pathological conclusions were reached by consensus. The pathological diagnosis of lung adenocarcinomas was classified as AAH, AIS, MIA, and IAC according to the 2015 WHO classification [12]. Notably, according to the 2021 WHO classification, MIA and IAC were categorized as adenocarcinomas and AAH and AIS as precursor glandular lesions [7]. All cases were divided into the non-IA group (including AAH and AIS) or the IA group (including MIA and IAC).

2.5. Statistical analysis

Statistical analysis was performed using IBM SPSS statistics, version 20 (IBM Corp., Armonk, NY). The univariate analysis was conducted on the qualitative variables and quantitative variables. Quantitative variables were compared by *t*-test or the Mann-Whitney *U* test according to whether or not the data conformed to a normal distribution. Differences in qualitative variables were analyzed using the χ^2 tests or Fisher's exact test. Variables that exhibited significant differences in the univariate analysis were included in the multivariate logistic regression analysis. Using the obtained factors with differences as predictors, a receiver operating curve (ROC) was plotted to obtain the cut-off value, sensitivity, and specificity. A *P* value of <0.05 is considered statistically significant.

3. Results

3.1. Relationships between clinical features and pathological diagnosis

The age difference of patients between the non-IA and IA groups was statistically significant ($P = 0.011$), and the average age of patients (54.76 ± 7.65) in the IA group was older than that in the non-IA group (49.78 ± 10.56). The gender ($P = 0.477$), smoking history ($P = 0.405$), and family history of lung cancer ($P = 0.426$) were not significantly different between the two groups (Table 1).

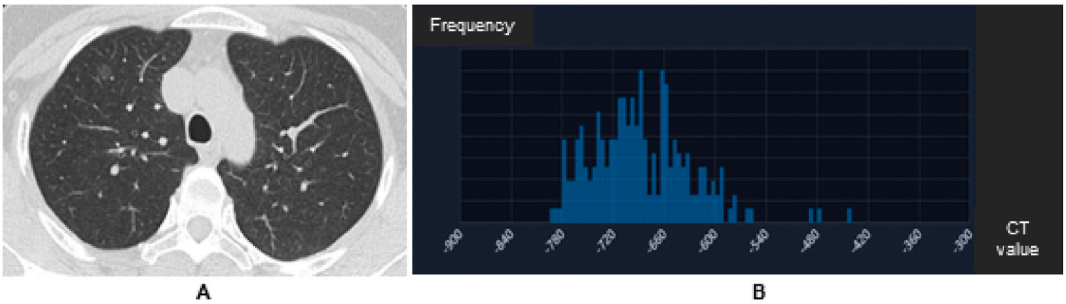

Fig. 1. A 51-year-old woman with atypical adenomatous hyperplasia (AAH) in the right upper lobe. (a) A transverse thin-section CT (1 mm thickness) scan shows a pure ground glass nodules (pGGN) (arrow) with a major diameter of 8 mm, a well-defined, smooth interface and a mean CT value of −697 HU. (b) The three-dimensional (3D) CT values histogram of the pGGN was calculated using artificial intelligence (AI) automated image analysis tool.

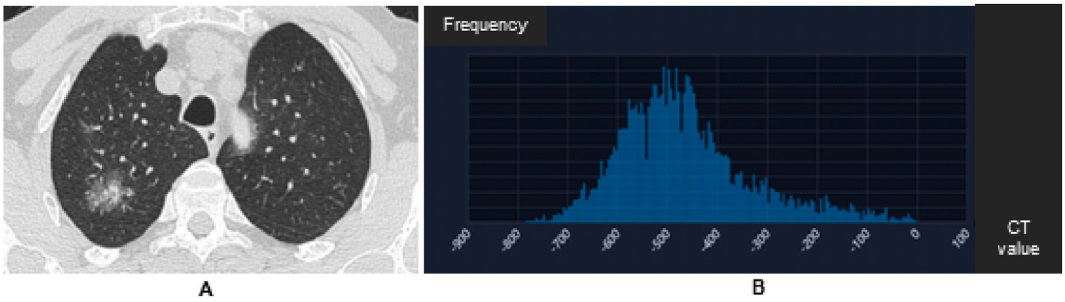

Fig. 2. A 67-year-old man with invasive adenocarcinoma (IAC) in the right upper lobe. (a) A transverse thin-section CT (1 mm thickness) scan shows part-solid nodules (PSN) (arrow) with a major diameter of 29 mm, spiculation, lobulation, vacuole sign, air bronchogram and a mean CT value of −463 HU. (b) The three-dimensional (3D) CT values histogram of the PSN was calculated using artificial intelligence (AI) automated image analysis tool.

3.2. Univariate analysis of qualitative and quantitative characteristics

We then investigated the qualitative and quantitative imaging characteristics between the two groups based on AI. Fig. 1 shows a transverse thin-section CT (1 mm thickness) scan (Fig. 1a) and the three-dimensional (3D) CT values histogram of the PSN (Fig. 1b) of a 51-year-old woman with atypical adenomatous hyperplasia (AAH) in the right upper lobe. Fig. 2 shows a transverse thin-section CT (1 mm thickness) scan (Fig. 2a) and the three-dimensional (3D) CT values histogram of the PSN (Fig. 2b) of a 67-year-old man with invasive adenocarcinoma (IAC) in the right upper lobe. Table 2 shows the difference in qualitative data between the non-IA and IA groups. The results show that the nodule type ($P < 0.001$), spiculation ($P < 0.001$), lobulation ($P = 0.005$), air bronchogram ($P = 0.002$), and pleural indentation ($P < 0.001$) between the non-IA and IA groups were all statistically significant. The SSNs in the IA group had an increased probability of solid component, spiculation, lobulation, air bronchogram, and pleural indentation characteristics than those in the non-IA group (Table 2). The vacuole sign ($P = 0.062$) and pulmonary nodule location ($P = 0.353$) were not significantly different between the non-IA and IA groups. The vacuole sign was present in 14 (11.7 %) SSNs of the non-IA group and IA groups, with 1 (2.8 %) case in the non-IA group and 13 (15.5 %) cases in the IA group. However, 106 (88.3 %) SSNs of the two groups lacked the vacuole sign. There were 81 (67.5 %) SSNs in the upper lobes of both lungs, including 54 (45.0 %) cases in the right upper lobe and 27 (22.5 %) cases in the left upper lobe. In the right upper lobe of the lung, there were 15 (41.7 %) and 39 (46.4 %) SSNs in the non-IA group and IA group, respectively. The SSNs of the non-IA group and IA groups were located mainly in the upper lobes of both

Table 2
The qualitative imaging characteristics of 120 SSNs and the results of univariate analysis.

Characteristics	Total ($N = 120$)	Non-I group ($N = 36$)	IA group ($N = 84$)	χ^2	P values
Type of nodules				14.292	0.000[a]
pGGN	43(35.8 %)	22(61.1 %)	21(25.0 %)		
PSN	77(64.2 %)	14(38.9 %)	63(75.0 %)		
Spiculation				19.755	0.000[a]
Present	50(41.7 %)	4(11.1 %)	46(54.8 %)		
Absent	70(58.3 %)	32(88.9 %)	38(45.2 %)		
Lobulation				7.902	0.005[a]
Present	76(63.3 %)	16(44.4 %)	60(71.4 %)		
Absent	44(36.7 %)	20(55.6 %)	24(28.6 %)		
Vacuole sign				3.943	0.062[b]
Present	14(11.7 %)	1(2.8 %)	13(15.5 %)		
Absent	106(88.3 %)	35(97.2 %)	71(84.5 %)		
Air bronchogram				3.178	0.002[a]
Present	62(51.7 %)	11(30.6 %)	51(60.7 %)		
Absent	58(48.3 %)	25(69.4 %)	33(39.3 %)		
Pleural indentation				14.939	0.000[a]
Present	59(49.2 %)	8(22.2 %)	51(60.7 %)		
Absent	61(50.8 %)	28(77.2 %)	33(39.3 %)		
Location of pulmonary nodule in lung				4.416	0.353[b]
Right upper	54(45.0 %)	15(41.7 %)	39(46.4 %)		
Right middle	8(6.7 %)	2(5.6 %)	6(7.1 %)		
Right lower	22(18.3 %)	4(11.1 %)	18(21.4 %)		
Left upper	27(22.5 %)	12(33.3 %)	15(17.9 %)		
Left lower	9(7.5 %)	3(8.3 %)	6(7.1 %)		

Note. -Values are presented as no. (%).
Non-IA group = AAH + AIS, IA group = MIA + IAC.
[a] : χ^2 tests.
[b] : Fisher's exact test.

lungs, particularly in the upper lobe of the right lung (Table 2).

Further, quantitative imaging characteristics between the non-IA and the IA groups are presented in Table 3. The major diameter, vertical short diameter, volume, maximum, mean, median, standard deviation, and entropy of 3D CT values between the non-IA and IA groups statistically significantly differed (for all, $P < 0.001$). Compared with the non-IA group, the IA group had greater major diameter, vertical short diameter, volume, maximum, mean, median, standard deviation, and entropy of 3D CT values. However, the minimum ($P = 0.286$), skewness ($P = 0.069$), and kurtosis ($P = 0.258$) of 3D CT values were not significantly different between the two groups (Table 3).

3.3. Multivariate logistic regression analysis and ROC analysis

Multivariate logistic regression analysis showed that the major diameter (odds ratio [OR] = 1.38; 95 % confidence interval [CI], 1.02–1.87; $P = 0.036$) and entropy of 3D CT value (OR = 3.73, 95 % CI, 1.13–2.33, $P = 0.031$) were independent risk factors for adenocarcinomas (Table 4). ROC curve analysis showed that, for differentiating IA from the non-IA group, the area under the curve (AUC) was 0.854 (sensitivity = 0.690, specificity = 0.889) with a major diameter value of 15.5 mm and 0.831 (sensitivity = 0.774, specificity = 0.833) with entropy of 3D values of 5.17. To improve the classification performance, we fused the major diameter and the entropy of 3D CT value as a combined model, and the AUC of the model was 0.868 (sensitivity = 0.845, specificity = 0.806) (Fig. 3). These results suggested that the major diameter of 15.5 mm and the entropy of 3D CT value of 5.17 had good predictive significance for distinguishing the non-IA from the IA groups.

4. Discussion

SSNs have a higher probability of being malignant than solid nodules, regardless of size [13]. Lung adenocarcinoma and its precursors, when manifested as SSNs, have an indolent course with slow growth and low metastasis potential. This is in particular the case for precursor glandular lesions, such as AAH and AIS, which have long-term survival without intervention, and a more aggressive strategy might even result in pulmonary function loss [2,14]. Because of the lower attenuation of SSNs and inexperienced radiologists, identifying lung adenocarcinoma and its precursors reliably before surgery remains a major challenge [15].

AI uses machine deep-learning algorithms or artificial neural networks (ANNs) to mine imaging characteristics and the heterogeneity of lesions thoroughly [16]. Chamberlin et al. reported that the identification sensitivity and performance of deep-learning systems were superior to the average sensitivity and performance of radiologists [17]. Wang et al. suggested that the deep-learning and radiomics networks based on radiomics performed well in classifying lung adenocarcinoma subtypes [18]. AI was good at recognizing SSNs and measuring their quantitative features, but it needed some help to identify the morphological characteristics of the SSNs accurately, in part because SSNs exhibit less attenuation and contrast than the nearby lung tissue.

In our study, to ensure the correctness of the six morphological characteristics derived from AI, radiologists extracted them all after twice analyzing the relationship between the lesion and surrounding tissues by applying multi-plane technology, window technology, and volume reconstruction. Consistent with previous studies [19,20], we discovered that adenocarcinomas were more prevalent in older individuals and exhibited spiculation, lobulation, air bronchogram, and pleural indentation more frequently than precursor glandular lesions. PSNs also favored an adenocarcinoma diagnosis. Nevertheless, all of these morphological features were not independent predictors for adenocarcinomas.

The major diameter of ≥15.5 mm (OR = 1.38, 95%CI, 1.02–1.87, $P = 0.036$) was identified as an independent risk factor for lung adenocarcinomas, with an AUC of 0.868, and a diagnostic sensitivity and specificity of 69.0 % and 88.9 %, respectively. Similar to the findings of this study, Wang et al. found that the major diameter of PSNs of ≥14.50 mm (OR = 0.171, 95%CI = 0.063–0.467, $P = 0.001$) was highly likely to be pathological as IAC [19]. Xiao et al. showed that the risk of the nodules being invasive adenocarcinomas increases when SSNs have a diameter of ≥14 mm [21]. The British Thoracic Society guidelines for the investigation and management of pulmonary nodules indicate that the maximum diameter of SSNs increased by 2 mm, which was a reliable indicator of malignancy

Table 3
The quantitative imaging characteristics of 120 SSNs and the results of univariate analysis.

Characteristics	Total ($N = 120$)	Non-IA group ($N = 36$)	IA group ($N = 84$)	P values
Major diameter (mm)	16.89(16.89, 22.00)	11.00(9.00, 13.50)	19.00(15.00, 25.50)	<0.001
Vertical short diameter (mm)	12.00(9.00, 16.25)	9.00(7.00, 11.00)	14.00(11.00, 17.00)	<0.001
Volume (ml)	882.88(433.75,2332.36)	333.00(200.36, 690.13)	1275.00(754.156, 3535.91)	<0.001
Maximum (HU)	76.50(-98.00, 215.75)	−198.00(-370.50, 75.50)	180.00(1.50, 272.00)	<0.001
Minimum (HU)	−793.00(-799.00,709.25)	−773.00(-789.00, −682.00)	−797.00(-800.00, −721.50)	0.286
Mean (HU)	−420.54(-549.63, −323.53)	−537.63(-621.77, −418.00)	−375.41(-465.15, −303.46)	<0.001
Median (HU)	−439.00(-567.50, −326.75)	−550.00(-637.50, −425.50)	−410.00(-517.25, −294.50)	<0.001
Standard deviation	158.22(109.98, 211.37)	117.68(70.75, 164.02)	178.42(119.70, 222.45)	<0.001
Skewness	−0.89(-1.34, −0.28)	−0.38(-1.00, −0.05)	−0.20(-0.74,0.01)	0.069
Kurtosis	2.25(2.00, 2.78)	3.00(2.43, 4.15)	2.70(2.20, 3.32)	0.258
Entropy	4.87(3.54, 5.58)	4.80(3.48, 5.00)	5.31(5.93, 6.21)	<0.001

Note. -Values are presented as median (P_{25}, P_{75}).
HU = Hounsfield unit.
Non-IA group = AAH + AIS, IA group = MIA + IAC.

Table 4
The results of multivariate logistic regression analysis.

Characteristics	OR	95 % CI	*P* values
Age	1.014	0.994–1.089	0.707
Type of nodules	0.253	0.039–1.652	0.151
Spiculation	0.241	0.028–2.058	0.194
Lobulation	1.303	0.281–6.044	0.735
Air bronchogram	4.155	0.753–22.914	0.102
Pleural indentation	0.704	0.150–3.318	0.658
Major diameter	1.383	1.021–1.873	**0.036**
Vertical short diameter	0.830	0.565–1.220	0.344
Volume	1.000	0.999–1.001	0.973
Maximum	1.001	0.995–1.008	0.632
Mean	0.993	0.974–1.011	0.429
Median	1.013	0.996–1.031	0.144
Standard deviation	0.996	0.976–1.017	0.714
Entropy	3.731	1.1291–2.331	**0.031**

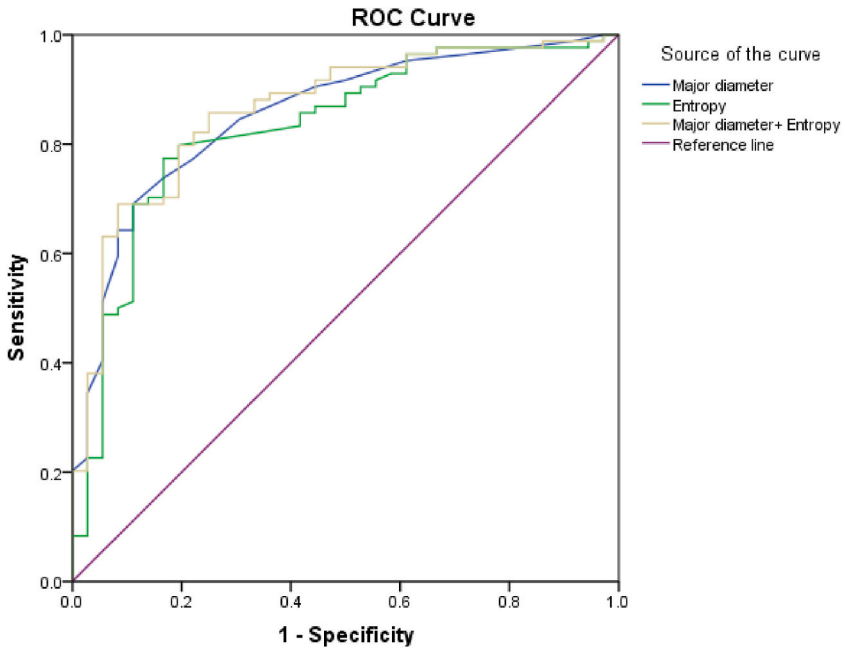

Fig. 3. The figure shows ROC curves of major diameter, entropy, major diameter entropy predicted non-IA and IA value, and the areas under the curve (AUC) for the identification of IA were 0.854, 0.831, and 0.868, respectively.

[22]. As a consequence, the lesion size enlarged accordingly with the increased invasiveness of lung adenocarcinoma presented as GGNs.

An entropy of 3D CT value of ≥5.17 (OR = 3.73, 95%CI, 1.13–2.33, $P = 0.031$) was cited as another independent risk factor for lung adenocarcinomas, with an AUC of 0.831, and a diagnostic sensitivity and specificity of 77.4 % and 83.3 %, respectively. Entropy had been used in previous research to assess the invasiveness of subsolid nodules [23,24], but the entropy of 3D CT values is used rarely, which can reflect the heterogeneity within the lesion more comprehensively. Zhang et al. found that nodules had a higher chance of being invasive adenocarcinomas when pGGNs were 3.29 or more in entropy [25]. In our study, 5.17 was the best cut-off value to make a distinction between the non-IA and IA groups, higher than prior studies had shown. The most probable explanation is that the object of our study was SSNs, including pGGNs and PSNs. The heterogeneity of lesions was increased by the solid component of PSNs and the related samples had greater overall entropy. Entropy is a measure of randomness, reflecting the inhomogeneity or complexity of the density of the lesion. The entropy increases as the heterogeneity of the lesion increases. The entropy of 3D CT values would grow in accordance with the degree of lung adenocarcinoma invasion, the difference in density within it, and the degree of the uneven density distribution of lesions.

To improve the classification performance, we fused the major diameter and the entropy of 3D CT value as a combined model, and the AUC of the model was 0.868 (sensitivity = 0.845, specificity = 0.806). These results suggested that the major diameter of 15.5 mm and the entropy of 3D CT value of 5.17 were of good predictive significance for distinguishing the non-IA from the IA group. However,

these findings need to be confirmed using a larger sample size.

Admittedly, our study presents several limitations. First, this study was conducted in a single institution with a limited sample size. Second, all SSNs with lower attenuation were automatic extracts by AI, which might not be as accurate at drawing the borders of SSNs. Thus, small blood vessels that were not eliminated might have had an impact on the textural characteristics. Third, the CT images were acquired with two distinct CT scanners, which might have influenced the heterogeneity of textural characteristics. Fourth, no internal or external validation of the predictive model in this study was carried out. These need to be refined in future studies.

5. Conclusion

The study result demonstrates using AI to improve the performance of distinguishing early lung adenocarcinomas in patients with SSNs, which can assist the design of subsequent exact treatment strategies and increase patient survival rates.

Funding

Not.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval

This study was conducted in accordance with the ethical regulations of the Declaration of Helsinki. The experiments were admitted to the Ethics Committee of the Fourth Hospital of Wuhan. The number of the Ethics Committee's acceptance is: 2021-005-01.

Informed consent

All patients signed the informed consent form.

CRediT authorship contribution statement

Xiaoqin Hu: Writing – original draft, Conceptualization. **Liu Yang:** Software, Formal analysis. **Tong Kang:** Formal analysis, Data curation. **Hanhua Yu:** Validation. **Tingkuan Zhao:** Formal analysis. **Yuanyi Huang:** Writing – review & editing, Conceptualization. **Yuefeng Kong:** Writing – review & editing, Conceptualization.

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.

Acknowledgments

We thank the members of the Fourth Hospital of Wuhan We apologize to the scientists whose work could not be cited due to space limitations.

References

[1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA A Cancer J. Clin. 71 (3) (2021) 209–249.
[2] S. Ricciardi, R. Booton, R.H. Petersen, M. Infante, M. Scarci, G. Veronesi, G. Cardillo, Managing of screening-detected sub-solid nodules-a European perspective, Transl. Lung Cancer Res. 10 (5) (2021) 2368–2377.
[3] Y.I. Cheng, M.P.A. Davies, D. Liu, W. Li, J.K. Field, Implementation planning for lung cancer screening in China, Precision clinical medicine 2 (1) (2019) 13–44.
[4] L. Azour, J.P. Ko, D.P. Naidich, W.H. Moore, Shades of gray: subsolid nodule considerations and management, Chest 159 (5) (2021) 2072–2089.
[5] T. Miyoshi, K. Aokage, S. Katsumata, K. Tane, G. Ishii, M. Tsuboi, Ground-glass opacity is a strong prognosticator for pathologic stage IA lung adenocarcinoma, Ann. Thorac. Surg. 108 (1) (2019) 249–255.
[6] A. Hattori, S. Hirayama, T. Matsunaga, T. Hayashi, K. Takamochi, S. Oh, K. Suzuki, Distinct clinicopathologic characteristics and prognosis based on the presence of ground glass opacity component in clinical stage IA lung adenocarcinoma, J. Thorac. Oncol. 14 (2) (2019) 265–275.
[7] A.G. Nicholson, M.S. Tsao, M.B. Beasley, A.C. Borczuk, E. Brambilla, W.A. Cooper, S. Dacic, D. Jain, K.M. Kerr, S. Lantuejoul, M. Noguchi, M. Papotti, N. Rekhtman, G. Scagliotti, P. van Schil, L. Sholl, Y. Yatabe, A. Yoshida, W.D. Travis, The 2021 WHO classification of lung tumors: impact of advances since 2015, J. Thorac. Oncol. 17 (3) (2022) 362–387.
[8] M. Yotsukura, H. Asamura, N. Motoi, J. Kashima, Y. Yoshida, K. Nakagawa, K. Shiraishi, T. Kohno, Y. Yatabe, S.I. Watanabe, Long-term prognosis of patients with resected adenocarcinoma in situ and minimally invasive adenocarcinoma of the lung, J. Thorac. Oncol. 16 (8) (2021) 1312–1320.
[9] A. Hosny, C. Parmar, J. Quackenbush, L.H. Schwartz, H. Aerts, Artificial intelligence in radiology, Nat. Rev. Cancer 18 (8) (2018) 500–510.
[10] Y. Sim, M.J. Chung, E. Kotter, S. Yune, M. Kim, S. Do, K. Han, H. Kim, S. Yang, D.J. Lee, B.W. Choi, Deep convolutional neural network-based software improves radiologist detection of malignant lung nodules on chest radiographs, Radiology 294 (1) (2020) 199–209.

[11] H. Kim, J.M. Goo, K.H. Lee, Y.T. Kim, C.M. Park, Preoperative CT-based deep learning model for predicting disease-free survival in patients with lung adenocarcinomas, Radiology 296 (1) (2020) 216–224.
[12] W.D. Travis, E. Brambilla, A.G. Nicholson, Y. Yatabe, J.H.M. Austin, M.B. Beasley, L.R. Chirieac, S. Dacic, E. Duhig, D.B. Flieder, K. Geisinger, F.R. Hirsch, Y. Ishikawa, K.M. Kerr, M. Noguchi, G. Pelosi, C.A. Powell, M.S. Tsao, I. Wistuba, The 2015 world health organization classification of lung tumors, J. Thorac. Oncol. 10 (9) (2015) 1243–1260.
[13] C.I. Henschke, D.F. Yankelevitz, R. Mirtcheva, G. McGuinness, D. McCauley, O.S. Miettinen, CT screening for lung cancer: frequency and significance of part-solid and nonsolid nodules, AJR, Am. J. Roentgenol. 178 (5) (2002) 1053–1057.
[14] M. Silva, M. Prokop, C. Jacobs, G. Capretti, N. Sverzellati, F. Ciompi, B. van Ginneken, C.M. Schaefer-Prokop, C. Galeone, A. Marchianò, U. Pastorino, Long-term active surveillance of screening detected subsolid nodules is a safe strategy to reduce overtreatment, J. Thorac. Oncol. 13 (10) (2018) 1454–1463.
[15] J. Zugazagoitia, C. Guedes, S. Ponce, I. Ferrer, S. Molina-Pinelo, L. Paz-Ares, Current challenges in cancer treatment, Clin. Therapeut. 38 (7) (2016) 1551–1566.
[16] M.J. Iqbal, Z. Javed, H. Sadia, I.A. Qureshi, A. Irshad, R. Ahmed, K. Malik, S. Raza, A. Abbas, R. Pezzani, J. Sharifi-Rad, Clinical applications of artificial intelligence and machine learning in cancer diagnosis: looking into the future, Cancer Cell Int. 21 (1) (2021) 270.
[17] S. Cui, S. Ming, Y. Lin, F. Chen, Q. Shen, H. Li, G. Chen, X. Gong, H. Wang, Development and clinical application of deep learning model for lung nodules screening on CT images, Sci. Rep. 10 (1) (2020) 13657.
[18] C. Wang, J. Shao, J. Lv, Y. Cao, C. Zhu, J. Li, W. Shen, L. Shi, D. Liu, W. Li, Deep learning for predicting subtype classification and survival of lung adenocarcinoma on computed tomography, Translational oncology 14 (8) (2021) 101141.
[19] H. Wang, Q. Weng, J. Hui, S. Fang, X. Wu, W. Mao, M. Chen, L. Zheng, Z. Wang, Z. Zhao, L. Zhou, J. Tu, M. Xu, Y. Huang, J. Ji, Value of TSCT features for differentiating preinvasive and minimally invasive adenocarcinoma from invasive adenocarcinoma presenting as subsolid nodules smaller than 3 cm, Acad. Radiol. 27 (3) (2020) 395–403.
[20] X. Yue, S. Liu, S. Liu, G. Yang, Z. Li, B. Wang, Q. Zhou, HRCT morphological characteristics distinguishing minimally invasive pulmonary adenocarcinoma from invasive pulmonary adenocarcinoma appearing as subsolid nodules with a diameter of $\leq$3 cm, Clin. Radiol. 73 (4) (2018), 411.e7-411.e15.
[21] W. Xiaolu, X. Qiuzhen, C. Wenda, W. Tao, J. Fengli, L. Wenhui, L. Guangming, Establishment and analysis of prediction model for invasive subsolid pulmonary nodules based on radiomics, Natl. Med. J. China (Peking) 102 (3) (2022) 209–215.
[22] M.E. Callister, D.R. Baldwin, A.R. Akram, S. Barnard, P. Cane, J. Draffan, K. Franks, F. Gleeson, R. Graham, P. Malhotra, M. Prokop, K. Rodger, M. Subesinghe, D. Waller, I. Woolhouse, British Thoracic Society guidelines for the investigation and management of pulmonary nodules, Thorax 70 (Suppl 2) (2015) ii1–ii54.
[23] X. Chen, B. Feng, Y. Chen, Y. Hao, X. Duan, E. Cui, Z. Liu, C. Zhang, W. Long, Whole-lesion computed tomography-based entropy parameters for the differentiation of minimally invasive and invasive adenocarcinomas appearing as pulmonary subsolid nodules, J. Comput. Assist. Tomogr. 43 (5) (2019) 817–824.
[24] Q. Sun, Y. Huang, J. Wang, S. Zhao, L. Zhang, W. Tang, N. Wu, Applying CT texture analysis to determine the prognostic value of subsolid nodules detected during low-dose CT screening, Clin. Radiol. 74 (1) (2019) 59–66.
[25] T. Zhang, X.H. Pu, M. Yuan, Y. Zhong, H. Li, J.F. Wu, T.F. Yu, Histogram analysis combined with morphological characteristics to discriminate adenocarcinoma in situ or minimally invasive adenocarcinoma from invasive adenocarcinoma appearing as pure ground-glass nodule, Eur. J. Radiol. 113 (2019) 238–244.