



Artificial intelligence neural network analysis and application of CT imaging features to predict lymph node metastasis in non-small cell lung cancer

Mingfei Geng¹, Mingsha Geng², Rong Wei³, Mingwei Chen⁴

¹Department of State-owned Assets Management, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China; ²Department of Information Management & Information Technology, The Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China; ³Department of Information Management & Information Technology, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China; ⁴Department of Respiratory and Critical Care Medicine, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China

Contributions: (I) Conception and design: Mingfei Geng; (II) Administrative support: M Chen; (III) Provision of study materials or patients: All authors; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Mingwei Chen. Department of Respiratory and Critical Care Medicine, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an 710061, China. Email: cmw_chenmingwei@163.com.

Background: Computed tomography (CT) is important in the diagnosing of lung cancer. The combination of CT features and artificial intelligence algorithm have been used in the diagnosis of various lung diseases. However, limited studies focused on the relationship between the combination of CT features and artificial intelligence algorithm and lymph node metastasis in non-small cell lung cancer (NSCLC). This study developed an algorithm for lung cancer CT image segmentation based on an artificial neural network model and investigated the role of a nomogram model based on CT images for predicting lymph node metastasis in lung cancer.

Methods: Wiener filtering and fuzzy enhancement were first used to suppress image noise and improve image contrast. Next, texture features and fractal features were extracted. In the third step, the artificial neural network model was trained and tested according to the best parameters of the network.

Results: The area under the curve (AUC) of the constructed nomogram model on the training set and the test set were 0.859 (sensitivity, 0.810; specificity, 0.773) and 0.864 (sensitivity, 0.820; specificity, 0.753), respectively. The decision curve indicated that the model had good clinical application value. The lung cancer CT images contained 13 significant regional features of cancer. The best classification function obtained from training and testing data was Levenberg-Marquardt backpropagation. The sensitivity, specificity, and accuracy in the training stage could reach 98.4%, 100%, and 98.6%, respectively, and the corresponding indexes in the test stage reached 90.9%, 100%, and 95.1%, respectively.

Conclusions: The image segmentation algorithm based on the artificial neural network model could extract CT lung cancer lesions efficiently and quasi-determinately, which could be used as an effective tool for radiologists to diagnose lung cancer. The nomogram model based on CT image features and related clinical indicators was an effective method for noninvasive prediction of lymph node metastasis in lung cancer.

Keywords: Computed tomography image (CT images); lymph node metastasis; artificial intelligence; neural network; prediction model

Submitted Aug 23, 2022. Accepted for publication Nov 08, 2022.

doi: 10.21037/jtd-22-1511

View this article at: <https://dx.doi.org/10.21037/jtd-22-1511>

Introduction

Lung cancer is a malignant tumor of the lung and has a high incidence rate and propensity for metastasis. Every year, there are about 1.73 million new cases of lung cancer worldwide, and its death toll accounts for 19% of the global cancer deaths (1). The disease is characterized by uncontrolled cell growth in lung tissue and a high degree of lethality. If lung cancer patients are not treated, the cancer cells will continue to grow and spread to tissues outside the lungs (2). Lung cancer is generally divided into 2 main types, small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). NSCLC is the most common type of lung cancer, accounting for approximately 80% of lung cancer cases (3-6). Although the number of patients with SCLC is small, tumor growth and metastasis are rapid and the degree of malignancy is very high. Therefore, the treatment of SCLC is based on systemic chemotherapy. Compared with SCLC, NSCLC has a relatively slow disease development and is less deadly (7). However, the development rate of the disease varies due to the large number of individual differences found in the lesions of NSCLC patients. Many patients with NSCLC fail to receive appropriate treatment in time due to lack of accurate disease development prediction, resulting in a mortality rate of up to 75% (8). While NSCLC cells grow and divide at a relatively slow rate, patients with NSCLC often miss the best opportunity to receive treatment because symptoms can be nonspecific. About 75% of patients with this type of lung cancer have already reached an advanced stage when the disease is found, resulting in a very low five-year survival rate, despite being a disease that can be alleviated or cured by treatment (9). The establishment of an effective prognostic model to evaluate the survival probability of cancer patients is important for planning follow-up treatment and designing a review plan. Imaging plays a guiding role in the treatment and prognosis of cancer. Through imaging research methods, more comprehensive, systematic, and in-depth information about tumor cells can be obtained (2).

Lymph node metastasis is an important factor affecting the survival and prognosis of lung cancer patients. Only 26–53% of patients with lymph node metastases survive more than 5 years. The gold standard for clinical diagnosis of lymph node metastasis in lung cancer is pathological biopsy (10), including thoracoscopic surgery and bronchial ultrasound, which are invasive methods and may cause serious complications. At present, a technical method

to predict lymph node metastasis of lung cancer by preoperative imaging is lacking. Computed tomography (CT) plays an important role in the routine diagnosis, analysis, and therapeutic evaluation of lung cancer because of its high spatial resolution. The shape, size, orientation, and other clinical information of nodules contained in CT images are of great significance for the diagnosis and prediction of diseases (11,12).

In recent years, radiomics technology has developed rapidly. High-throughput extraction of deep feature information hidden in medical images can accurately reflect the biological characteristics of tumors, thereby allowing for noninvasive classification and discrimination of a patient's tumor (13). By combining radiomics features and clinical indicators of patients, a nomogram model can fully reveal the potential association between features and disease pathology, and it has good clinical applicability. This study used CT images and clinical indicators of lung cancer patients to establish a nomogram model to analyze its effect on predicting lymph node metastasis in order to provide a reference for clinical decision-making. We present the following article in accordance with the TRIPOD reporting checklist (available at <https://jtd.amegroups.com/article/view/10.21037/jtd-22-1511/rc>).

Methods

Analysis of CT image segmentation algorithm

The CT images in this study were obtained retrospectively from 46 males and 36 females and included 48 patients with lung cancer and 34 normal patients from the First Affiliated Hospital of Xi'an Jiaotong University. Inclusion criteria of lung cancer patients: (I) NSCLC; (II) chest CT examination; (III) clinical and pathological data can be obtained after surgical treatment in our hospital. Exclusion criteria: (I) pulmonary metastatic cancer or SCLC; (II) recurrent lung cancer; (III) distant metastasis has occurred. Image preprocessing involved using the Wiener filter to suppress image noise and fuzzy enhancement to improve image contrast. Texture and fractal features were then extracted from the preprocessed images. Finally, salient features were selected and input into the artificial neural network model to complete image segmentation. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the Ethics Committee of the First Affiliated Hospital of Xi'an Jiaotong University (No. 20210040) and individual

consent for this retrospective analysis was waived.

The purpose of medical image registration is to realize the spatial transformation of medical images and ensure that they can achieve spatial consistency with the corresponding user interface (UI) points in other modal images. In the 2 medical images participating in registration, the image that remains stationary is the reference image, and the image that is transformed, is denoted as the floating image. The next step of the medical image registration process is to solve the space transformation T .

$$\hat{T} = \arg \min_T C(I_F, T(I_M)) \quad [1]$$

T is the spatial transformation involved in registration and is the objective function. In the process of medical image registration, taking the rigid body space transformation as an example, the transformation formula is as follows:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad [2]$$

Image preprocessing

The quality of CT images is often degraded by noise pollution. The Wiener filter is a classic linear denoising filter and is often used to recover useful signals from additive noise. The filtering method combines a degradation function and statistical noise features for restoration processing. Wiener filtering can be called an optimal filtering system when the minimum mean square error (MSE) between the filter output and the desired output is achieved. In addition, lung CT image features include gray level, size, shape, and other differences, which can cause image boundaries and regions to be blurred, and traditional image enhancement methods have difficulty distinguishing the blurred boundary regions. In this study, fuzzy enhancement algorithm was used to deal with the inaccuracy and fuzzy information of CT images and enhance image contrast.

Image texture feature extraction

Texture is a type of visual feature that shows the homogeneity in an image, reflecting the surface structure organization and arrangement property of the object surface that changes slowly or periodically. Gray cooccurrence

matrix is commonly used to describe the gray level of image pixels.

Image preprocessing is divided into 3 steps: grayscale, negative image, and histogram processing. The spatial enhancement results obtained by gray-scale processing can be expressed as:

$$h(x, y) = L - 1 - F(x, y) \quad [3]$$

$$p(s_k) = \frac{n_k}{n}, \quad k = 0, 1, \dots, L-1 \quad [4]$$

CT image fractal feature extraction

Fractal features are often used to describe the complex and irregular features of medical images. Fractal dimension is an important parameter to describe the complexity, irregularity, and spatial distribution trend of nonlinear images, which is calculated by the differential box dimension method, as follows. For a gray image of size $N \times N$, the 2-dimensional image is regarded as a surface of 3-dimensional space (I, j , $P(I, j)$), where (I, j) represents the spatial location of pixels, and $P(I, j)$ represents the gray value of the corresponding position. Image gray-level change will then be reflected by the roughness of the spatial 3-dimensional surface, so it is not used. The dimension obtained by measuring the stereo surface at the same scale is the fractal dimension of the image.

$$\begin{aligned} n_r(i, j) &= l - k + 1 \\ N(r) &= \sum_{i, j} n_r(i, j) \\ PD &= \frac{\log[N(r)]}{\log(1/r)} \end{aligned} \quad [5]$$

Artificial neural network

The extracted texture and fractal features are not required, while the salient features need to be selected as the input for the artificial neural network model. An artificial neural network is composed of neurons connected with each other and is a mathematical model that simulates the behavior characteristics of a biological neural network and performs distributed parallel data processing. The model is capable of self-learning and self-adaptation, and it can determine the potential rules distinguishing the training samples and the prediction samples in advance, using the rules formed in the training stage to calculate the new input sample data. The most commonly used model is the backward propagation

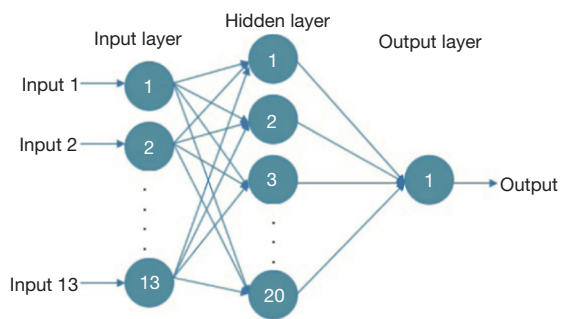


Figure 1 Backward propagation neural network model structure.

neural network model, which can use the errors generated by the output results to carry out backward reasoning so as to gradually optimize the network structure and network parameters. The structure is shown in *Figure 1*, with the input layer on the left, the hidden layer is in the middle, and on the right is the output layer. Each node represents a specific output function, called the excitation function, and the connection between 2 nodes represents a weighted value for the signal passing through the connection (*Figure 1*).

Evaluation index

The MSE, sensitivity, specificity, and accuracy were used to evaluate the performance of the artificial neural network in this study. Target represents the target area, actual represents the measured area, and TN, TP, FN, and FP represent the number of true positives, true negatives, false positives, and false negatives, respectively.

Construction of nomogram model

In combination with clinical indicators [age, gender, smoking history, carcinoembryonic antigen (CEA), neuron-specific enolase (NSE), and *cytokeratin fragment* (CYFRA)], the radiomics label was constructed with the Least Absolute Shrinkage and Selection Operator (LASSO) algorithm. The RMS software package (V.3.5.1) in R was used to establish the nomogram model. Decision curve analysis (DCA) was used to evaluate the clinical application value of the nomogram by calculating the net benefit under different threshold probabilities.

Model validation method and statistical analysis

The area under curve (AUC), accuracy, specificity, and sensitivity of receiver operating characteristic (ROC) curves were used as the comparison indexes, and the calculation formula was as follows: (I) Accuracy = (number of true positive patients + number of true negative patients)/(true negative number + false negative number + true positive number + false positive number); (II) sensitivity = true positive number/(true positive number + false negative number); and (III) specificity = true negative number/(true negative number + false positive number). The hardware platform for all algorithms in this study was 3.7 GHz Intel I7-8700K CPU and 64GB 3000 MHz DDR4.

Results

Analysis of lung cancer CT image preprocessing

The CT images in this study were obtained from 46 males and 36 females and included 48 patients with lung cancer and 34 normal patients. A total of 1,150 backgrounds, lung cancer regions, and boundaries were extracted, and the image size was 256×256×63 voxels. All simulation experiments were implemented on MATLAB R2013a platform. Image preprocessing results are shown in *Figure 2*, with image noise eliminated by Wiener filtering and image contrast significantly improved after blur enhancement.

Image feature extraction

Based on the values of CT lung image texture and fractal features, it was revealed that only 3 out of 14 texture features (TF1, TF2, TF3) had no overlap between the feature values of normal lung and lung cancer, and 10 out of 12 fractal features (FF11, FF12, FF21, FF22, FF41, FF42, FF51, FF52, FF61, FF62) had no overlapping eigenvalues and could be used as significant features. *Figure 3* and *Figure 4* show the classification images of lung cancer with texture features and fractal features, respectively. It can be seen that the lesions are completely extracted (*Figures 3,4*).

Image segmentation and radiomics feature extraction

The Digital *Imaging* and Communications in Medicine

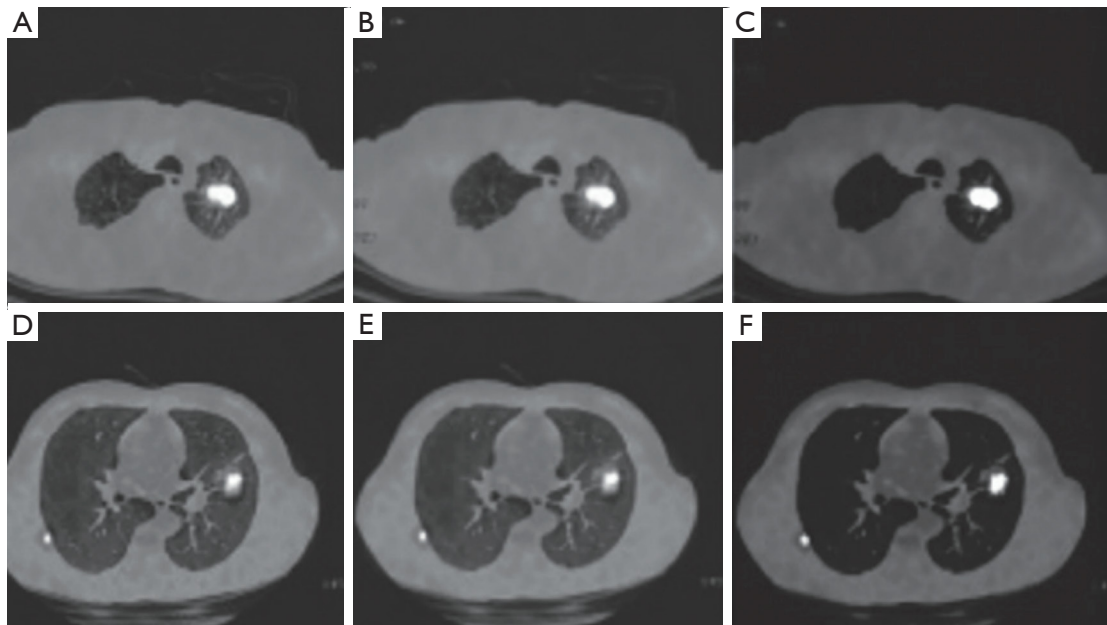


Figure 2 Image preprocessing result.

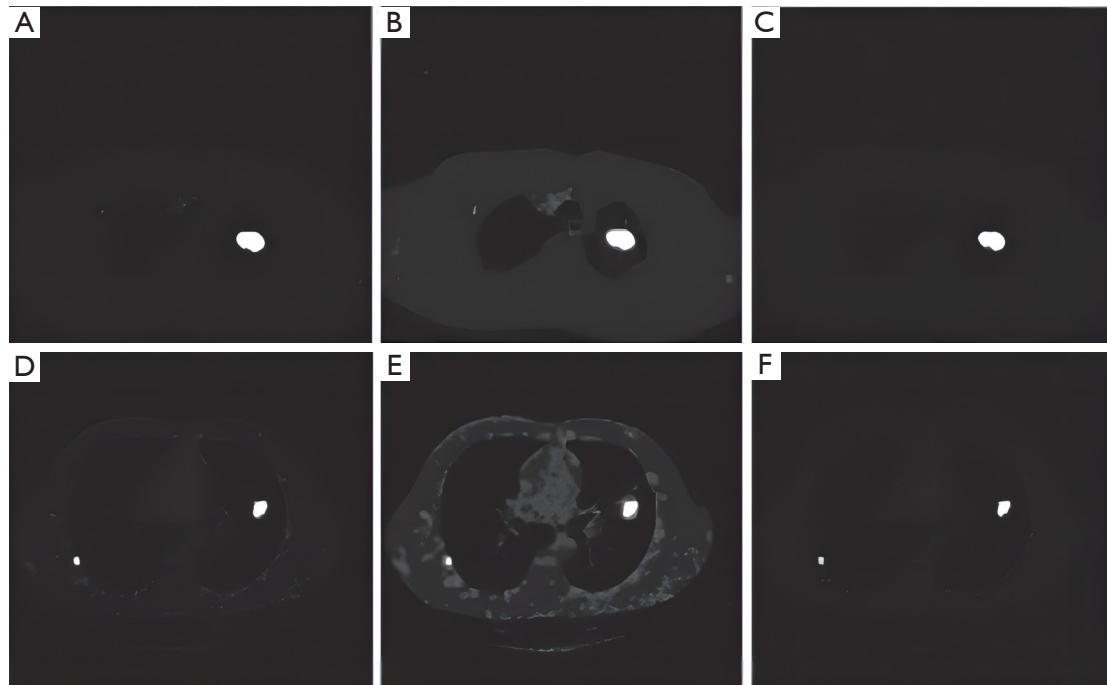


Figure 3 Significant texture features in computed tomography images of lung cancer.

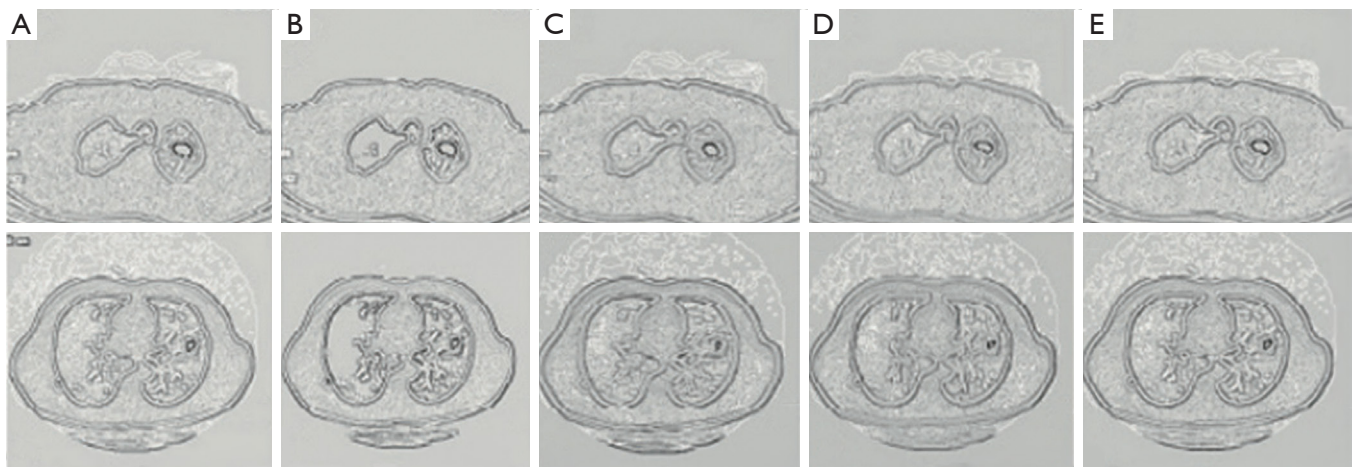


Figure 4 Significant fractal features of computed tomography images of lung cancer.

(DICOM) format CT image data of the included cases were imported into the workstation of α -Discovery, where the work was performed. The station completed image segmentation and radiomics feature extraction. The workstation used a 3D-NET model for nodule segmentation, which was calibrated by the Lung Image Database Consortium (LIDC) dataset. The software automatically identified and segmented pulmonary nodules, and a comparison of automatic segmentation and the manual segmentation of nodules by radiologists showed it had good stability and robustness (Figure 5).

Construction of the nomogram model

A total of 855 radiomics features were extracted from the CT images of patients. The LASSO regression model was used to reduce the dimension of features to obtain 9 features with high correlation, and the radiomics label, namely CT score, was constructed. The nomogram model, consisting of the clinical indicator CEA and radiomics labels constructed by multivariate logistic regression, is shown in the Figure 6. The total score was obtained by adding the score of the radiomics labels and CEA scores together. The total score was positioned on the total number line and the intersection point between a vertical line drawn from top to bottom. The bottom scale is the risk of lymph node metastasis of malignant lung nodules (range, 0.1–0.9) (Figure 6).

ROC curve analysis of lung cancer prognosis model

The AUC of the model training set was 0.859 (sensitivity,

0.810; specificity, 0.773), and the AUC of the test set was 0.864 (sensitivity, 0.820; specificity, 0.753) (Figure 7). The results showed that the nomogram model had good predictive performance in predicting lymph node metastasis of lung cancer. The DCA results showed that with a threshold probability of lymph node metastasis in patients greater than 0.06%, the use of the nomogram model for predicting the patients with lymph node metastasis compared with the other 2 treatments (assuming that all patients had lymph node metastases and assuming that all patients with no lymph node metastasis) could achieve a greater net benefit. This conclusion highlighted the clinical application value of the nomogram model in the prediction of lymph node metastasis (Figure 8).

Discussion

In recent years, scholars have devoted themselves to inventing various indicators to predict the prognosis of patients with different diseases (14–17). In this study, an algorithm based on an artificial neural network model for lung cancer CT image segmentation was proposed (18,19). Wiener filtering and fuzzy enhancement could effectively suppress image noise and improve image contrast. The nonoverlapping feature value method could accurately select significant texture features and fractal features, and the square root error could predict the best parameters of the artificial neural network model, which could quickly and accurately extract the lung cancer lesion region (20–22). These results showed that the lung cancer CT image segmentation algorithm proposed in this study was

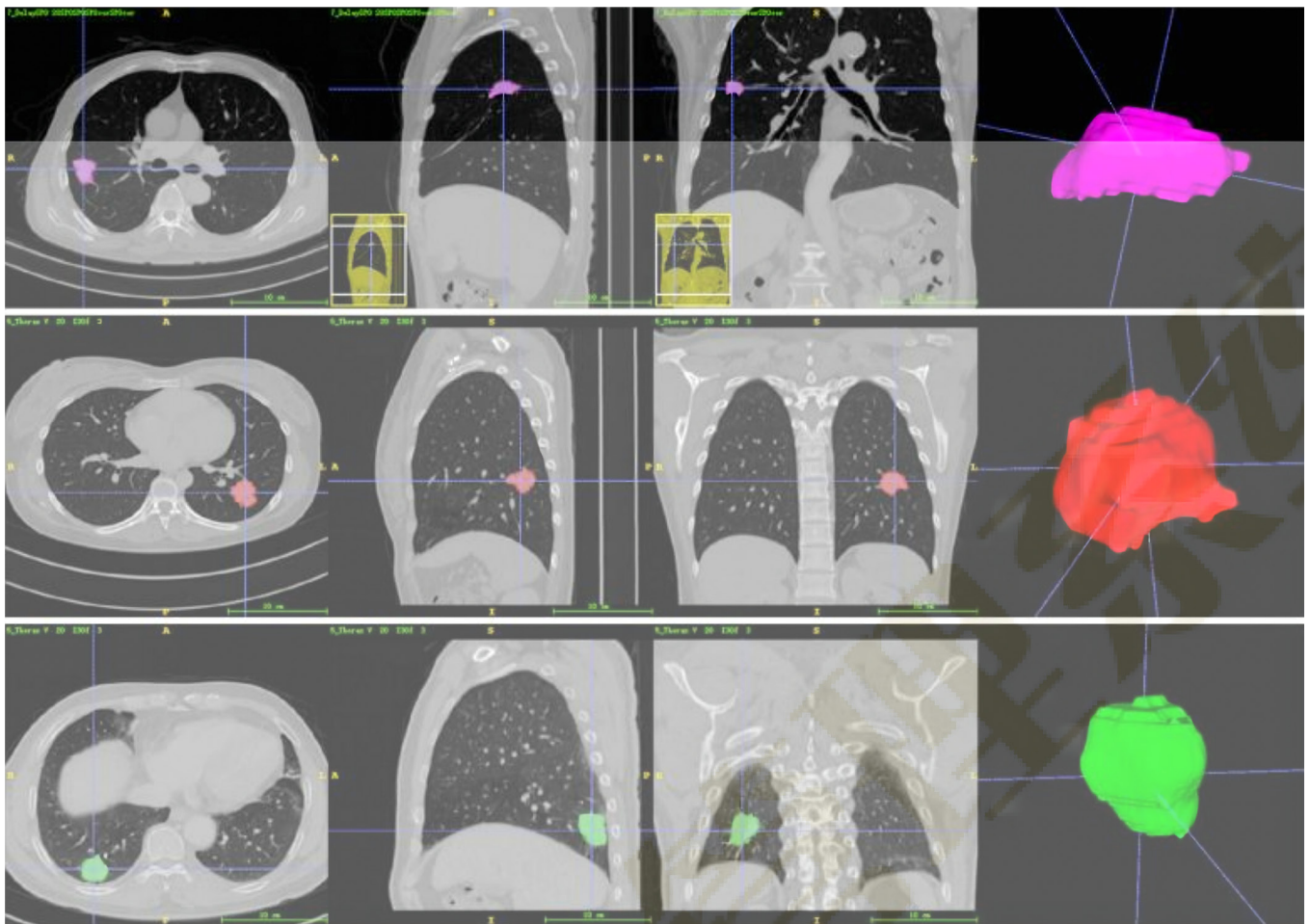


Figure 5 Schematic diagram of automatic 3D segmentation of lung cancer computed tomography images.

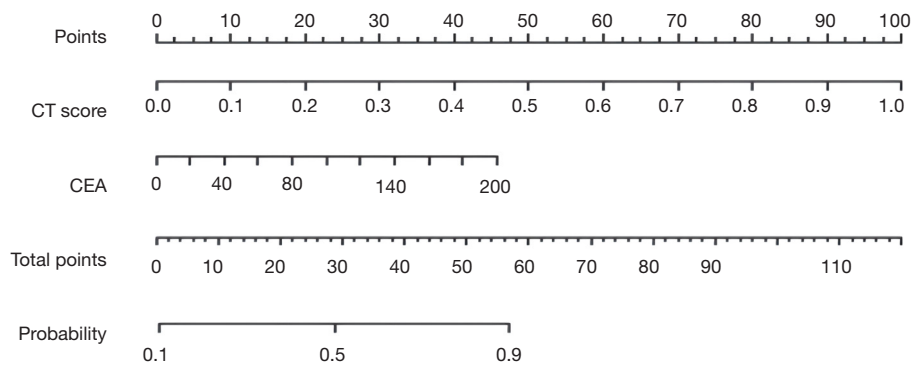


Figure 6 Nomogram model constructed by multifactor logistic regression combined with CEA and radiomics tags. CT, computed tomography; CEA, carcinoembryonic antigen.

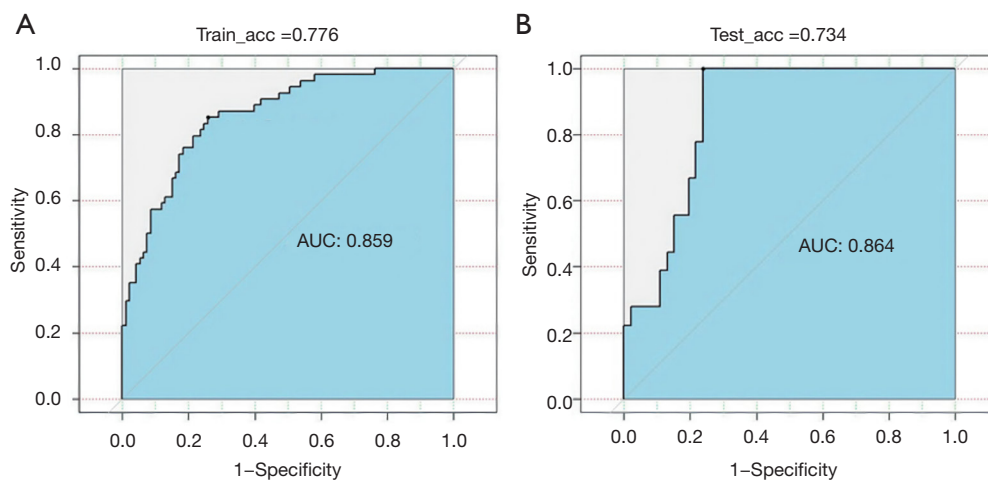


Figure 7 Receiver operating characteristic curves for training and testing sets of nomogram model. AUC, area under the curve.

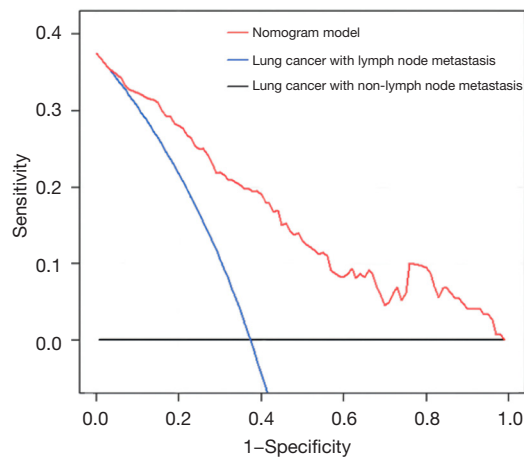


Figure 8 Prediction curves in a nomogram model.

feasible and practical, and it could be used as an auxiliary tool in the diagnosis of lung cancer.

Lymph node metastasis status of lung cancer is a key factor in the choice of surgical treatment and whether to use adjuvant therapy such as lymph node dissection (23,24), and noninvasive preoperative prediction of lymph node metastasis has important clinical significance. ROC curve analysis is the most intuitive and commonly used method to judge the merits of a model. In this paper, ROC curve analysis was used to visually demonstrate the prediction performance of the nomogram model. The AUC was used to accurately predict the probability of patients' disease, which could be used to evaluate the predictive performance of the model, along with specificity and sensitivity. In this

study, the AUC of the nomogram model test set was 0.864, the specificity was 0.753, and the sensitivity was 0.820, which showed excellent ability to predict lymph node metastasis.

Nomogram models are often used in radiomics research as they can evaluate the probability of clinical events by individual characteristics of patients and are a good predictive classification model. Compared with other predictive statistical models, its visual icon results can more intuitively reflect the disease probability of patients and provide individualized prognostic risk assessment. CEA is a serum tumor marker that is often used to detect early lung cancer (25,26). In this study, the clinical indicators related to lung cancer were statistically analyzed, and the results showed that CEA was highly correlated with lymph node metastasis of malignant lung nodules ($P < 0.001$), which was an independent prognostic factor for constructing the nomogram model to predict lymph node metastasis in lung cancer patients, consistent with previous studies (27-30). In this study, the clinical indicator CEA and radiomics labels constructed by LASSO algorithm (the LASSO algorithm has low requirements on processed data and is a widely used data screening method) were used as predictors to construct the nomogram. The AUC of the ROC curve for the test set drawn by the model was 0.864, the sensitivity was 0.820, and the specificity was 0.753. Related studies have established nomograms to predict lymph node metastasis of lung cancer based on tumor/node/metastasis (TNM) grade of lung cancer and radiomics feature label. The AUC of the test set was 0.856, the specificity was 0.821, and the sensitivity was 0.917, indicating that the nomogram model

established by CEA and lung cancer radiomics features had an excellent ability to predict lymph node metastasis. DCA is a statistical method that assesses the utility of a model in supporting clinical decision making. Compared with ROC analysis to evaluate a model's accuracy, DCA is more inclined to evaluate the model from the perspective of clinical benefit, that is, to find a way to maximize the net benefit for patients when false positives and false negatives are unavoidable. In this study, the nomogram decision curve showed that when the threshold probability of lymph node metastasis was greater than 0.06%, the nomogram model was more effective than "no treatment" or "all treatment" strategy in predicting lymph node metastasis in patients, showing good potential for clinical application.

This study had several limitations. First, as all data were obtained from the same hospital, the results were not universal, and thus multicenter data will be introduced in future studies. Secondly, lymph node metastasis of lung cancer is also related to the expression of genes and proteins, which was not included in this study. Finally, the radiomics features in this study relied on the manual delineation of the region of interest (ROI) by 2 doctors. The correlation value of the delineation of the lesion area of the lung nodules by the 2 doctors was 0.954, which was highly consistent but time-consuming. In future research, the sample size will be expanded to include patients from different regions and more clinical information. A more efficient and more accurate segmentation algorithm will be more likely to improve the performance and universality of the model.

In summary, a nomogram model containing radiomics labels of clinical indicators was established using CT images. ROC and DCA results showed that the nomogram model showed good predictive effect and clinical application value in predicting lymph node metastasis of lung cancer. However, further studies were still needed as this study was a retrospective study with limited patients.

Acknowledgments

Funding: This study was supported by the General Projects - Social Development Area (No. 2020SF-285).

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at <https://jtd.amegroups.com/article/view/10.21037/jtd-22-1511/rc>

Data Sharing Statement: Available at <https://jtd.amegroups.com/article/view/10.21037/jtd-22-1511/dss>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://jtd.amegroups.com/article/view/10.21037/jtd-22-1511/coif>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the Ethics Committee of the First Affiliated Hospital of Xi'an Jiaotong University (No. 20210040) and individual consent for this retrospective analysis was waived.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Siegel RL, Miller KD, Fuchs HE, et al. Cancer Statistics, 2021. *CA Cancer J Clin* 2021;71:7-33.
2. Zhang Y, Li Z, Chen Y, et al. Outcomes of Image-Guided Moderately Hypofractionated Radiotherapy for Stage III Non-Small-Cell Lung Cancer. *J Oncol* 2021;2021:2721261.
3. Peng L, Qin BD, Xu S, et al. Risk and Incidence of Infection with Bevacizumab in Non-Small-Cell Lung Cancer Patients: A Meta-Analysis. *Oncol Res Treat* 2022;45:281-90.
4. Little A, Roder D, Zhao GW, et al. Country of birth and non-small cell lung cancer incidence, treatment, and outcomes in New South Wales, Australia: a population-based linkage study. *BMC Pulm Med* 2022;22:366.
5. Greene G, Griffiths R, Han J, et al. Impact of the SARS-CoV-2 pandemic on female breast, colorectal and non-small cell lung cancer incidence, stage and healthcare pathway to diagnosis during 2020 in Wales, UK, using

- a national cancer clinical record system. *Br J Cancer* 2022;127:558-68.
6. Beachler DC, Lamy FX, Kolitsopoulos F, et al. Incidence of safety events after immune checkpoint inhibitor initiation for advanced-stage non-small-cell lung cancer: a real-world study. *Future Oncol* 2022;18:2891-901.
 7. Snee M, Cheeseman S, Thompson M, et al. Trends in the prescription of systemic anticancer therapy and mortality among patients with advanced non-small cell lung cancer: a real-world retrospective observational cohort study from the I-O optimise initiative. *BMJ Open* 2021;11:e043442.
 8. Sun X, Men Y, Wang J, et al. Risk of cardiac-related mortality in stage IIIA-N2 non-small cell lung cancer: Analysis of the Surveillance, Epidemiology, and End Results (SEER) database. *Thorac Cancer* 2021;12:1358-65.
 9. Jiang KT, Huang DZ. The risks and trends of cardiac-specific mortality associated with chemotherapy or radiotherapy in a large cohort of non-elderly patients with non-small cell lung cancer. *Transl Cancer Res* 2021;10:842-53.
 10. Ding Y, Li J, Li X, et al. Impact of preoperative biopsy on tumor spread through air spaces in stage I non-small cell lung cancer: a propensity score-matched study. *BMC Pulm Med* 2022;22:293.
 11. Wei SH, Zhang JM, Shi B, et al. The value of CT radiomics features to predict visceral pleural invasion in ≤ 3 cm peripheral type early non-small cell lung cancer. *J Xray Sci Technol* 2022. doi: 10.3233/XST-221220.
 12. Xie P, Zhang Y, He L. Prognostic Evaluation of CT Imaging Big Data-Assisted Arterial Chemoembolization Combined with 125I Seed Implantation for Non-Small-Cell Lung Cancer. *Comput Math Methods Med* 2022;2022:3472982.
 13. Wang D, Zhang X, Liu H, et al. Assessing dynamic metabolic heterogeneity in non-small cell lung cancer patients via ultra-high sensitivity total-body [18F]FDG PET/CT imaging: quantitative analysis of [18F]FDG uptake in primary tumors and metastatic lymph nodes. *Eur J Nucl Med Mol Imaging* 2022;49:4692-704.
 14. Chen Y, Wang J, Zhang X, et al. Correlation between apparent diffusion coefficient and pathological characteristics of patients with invasive breast cancer. *Ann Transl Med* 2021;9:143.
 15. Qi A, Li Y, Yan S, et al. Effect of postoperative chemotherapy on blood glucose and lipid metabolism in patients with invasive breast cancer. *Gland Surg* 2021;10:1470-7.
 16. Qi A, Li Y, Sun H, et al. Incidence and risk factors of sexual dysfunction in young breast cancer survivors. *Ann Palliat Med* 2021;10:4428-34.
 17. Teng D, Xia S, Hu S, et al. miR-887-3p Inhibits the Progression of Colorectal Cancer via Downregulating DNMT1 Expression and Regulating P53 Expression. *Comput Intell Neurosci* 2022;2022:7179733.
 18. Park D, Oh D, Lee M, et al. Importance of CT image normalization in radiomics analysis: prediction of 3-year recurrence-free survival in non-small cell lung cancer. *Eur Radiol* 2022. [Epub ahead of print]. doi: 10.1007/s00330-022-08869-2.
 19. Chen W, Hou X, Hu Y, et al. A deep learning- and CT image-based prognostic model for the prediction of survival in non-small cell lung cancer. *Med Phys* 2021;48:7946-58.
 20. Song Z, Liu T, Shi L, et al. The deep learning model combining CT image and clinicopathological information for predicting ALK fusion status and response to ALK-TKI therapy in non-small cell lung cancer patients. *Eur J Nucl Med Mol Imaging* 2021;48:361-71.
 21. Wang B, Wang DQ, Lin MS, et al. Accumulation of the delivered dose based on cone-beam CT and deformable image registration for non-small cell lung cancer treated with hypofractionated radiotherapy. *BMC Cancer* 2020;20:1112.
 22. Yu L, Tao G, Zhu L, et al. Prediction of pathologic stage in non-small cell lung cancer using machine learning algorithm based on CT image feature analysis. *BMC Cancer* 2019;19:464.
 23. Qiao J, Zhang X, Du M, et al. 18F-FDG PET/CT radiomics nomogram for predicting occult lymph node metastasis of non-small cell lung cancer. *Front Oncol* 2022;12:974934.
 24. Huang Y, Jiang X, Xu H, et al. Preoperative prediction of mediastinal lymph node metastasis in non-small cell lung cancer based on 18F-FDG PET/CT radiomics. *Clin Radiol* 2022. [Epub ahead of print]. pii: S0009-9260(22)00524-4. doi: 10.1016/j.crad.2022.08.140.
 25. Zheng J, Wang Y, Hu C, et al. Predictive value of early kinetics of ctDNA combined with cfDNA and serum CEA for EGFR-TKI treatment in advanced non-small cell lung cancer. *Thorac Cancer* 2022;13:3162-73.
 26. Wang H, Zhou X, Wang Z, et al. Clinical Efficacy of Osimertinib in Patients with Advanced Non-Small Cell Lung Cancer and Its Effect on Serum CEA and VEGF Expression. *Evid Based Complement Alternat Med* 2022;2022:3032087.

27. Li Z, Zhao J. Clinical efficacy and safety of crizotinib and alectinib in ALK-positive non-small cell lung cancer treatment and predictive value of CEA and CA125 for treatment efficacy. *Am J Transl Res* 2021;13:13108-16.
28. Zhao T, Mao G, Chen M. The Role of Change Rates of CYFRA21-1 and CEA in Predicting Chemotherapy Efficacy for Non-Small-Cell Lung Cancer. *Comput Math Methods Med* 2021;2021:1951364.
29. Clevers MR, Kastelijin EA, Peters BJM, et al. Evaluation of Serum Biomarker CEA and Ca-125 as Immunotherapy Response Predictors in Metastatic Non-small Cell Lung Cancer. *Anticancer Res* 2021;41:869-76.
30. Mishra A, Singh N, Shyam H, et al. Differential expression profiling of transcripts of IDH1, CEA, Cyfra21-1, and TPA in stage IIIa non-small cell lung cancer (NSCLC) of smokers and non-smokers cases with air quality index. *Gene* 2021;766:145151.

Cite this article as: Geng M, Geng M, Wei R, Chen M. Artificial intelligence neural network analysis and application of CT imaging features to predict lymph node metastasis in non-small cell lung cancer. *J Thorac Dis* 2022;14(11):4384-4394. doi: 10.21037/jtd-22-1511