

RESEARCH

Open Access



Predicting malignant risk of ground-glass nodules using convolutional neural networks based on dual-time-point ^{18}F -FDG PET/CT

Yuhang Liu^{1†}, Jian Wang^{1†}, Bulin Du^{1†}, Yaming Li^{1*} and Xuena Li^{1*}

Abstract

Background Accurately predicting the malignant risk of ground-glass nodules (GGOs) is crucial for precise treatment planning. This study aims to utilize convolutional neural networks based on dual-time-point ^{18}F -FDG PET/CT to predict the malignant risk of GGOs.

Methods Retrospectively analyzing 311 patients with 397 GGOs, this study identified 118 low-risk GGOs and 279 high-risk GGOs through pathology and follow-up according to the new WHO classification. The dataset was randomly divided into a training set comprising 239 patients (318 lesions) and a testing set comprising 72 patients (79 lesions), we employed a self-configuring 3D nnU-net convolutional neural network with majority voting method to segment GGOs and predict malignant risk of GGOs. Three independent segmentation prediction models were developed based on thin-section lung CT, early-phase ^{18}F -FDG PET/CT, and dual-time-point ^{18}F -FDG PET/CT, respectively. Simultaneously, the results of the dual-time-point ^{18}F -FDG PET/CT model on the testing set were compared with the diagnostic of nuclear medicine physicians.

Results The dual-time-point ^{18}F -FDG PET/CT model achieving a Dice coefficient of 0.84 ± 0.02 for GGOs segmentation and demonstrating high accuracy (84.81%), specificity (84.62%), sensitivity (84.91%), and AUC (0.85) in predicting malignant risk. The accuracy of the thin-section CT model is 73.42%, and the accuracy of the early-phase ^{18}F -FDG PET/CT model is 78.48%, both of which are lower than the accuracy of the dual-time-point ^{18}F -FDG PET/CT model. The diagnostic accuracy for resident, junior and expert physicians were 67.09%, 74.68%, and 78.48%, respectively. The accuracy (84.81%) of the dual-time-point ^{18}F -FDG PET/CT model was significantly higher than that of nuclear medicine physicians.

Conclusions Based on dual-time-point ^{18}F -FDG PET/CT images, the 3D nnU-net with a majority voting method, demonstrates excellent performance in predicting the malignant risk of GGOs. This methodology serves as a valuable adjunct for physicians in the risk prediction and assessment of GGOs.

[†]Yuhang Liu, Jian Wang and Bulin Du contributed equally to this work and should be considered co-first authors.

*Correspondence:

Yaming Li

yml2001@163.com

Xuena Li

lixuenacmunm@163.com

Full list of author information is available at the end of the article



Keywords Convolutional neural network, Lung, Ground-Glass nodules, Malignant risk, Dual-time-point ^{18}F -FDG PET/CT

Background

Lung cancer stands as a predominant cause of cancer-related mortality worldwide, accounting for 18% of all cancer-related deaths [1]. Medical imaging plays an important role in the early diagnosis of the lung cancer. Especially, with the widespread adoption of low-dose lung computed tomography in lung cancer screening in recent years, the detection rate of subsolid nodules has significantly increased [2]. Subsolid nodules, also known as ground-glass nodules (GGOs), represent one of the most common manifestations found in lung cancer screening [3]. GGOs comprise pure ground-glass nodules (pGGO) and mixed ground-glass nodules (mGGO) with both ground-glass and solid densities [4, 5]. GGOs encompass a spectrum of lesions, including benign nodules, atypical adenomatoid hyperplasia (AAH), adenocarcinoma in situ (AIS), microinvasive adenocarcinoma (MIA) and invasive adenocarcinoma (IA) [6]. According to latest guideline published by the World Health Organization (WHO) in 2021, based on molecular genetic characteristics, pulmonary lesions were categorized into benign tumors, precursor glandular lesions, and pulmonary adenocarcinomas. Precursor glandular lesions encompass AAH and AIS, while pulmonary adenocarcinomas include MIA and IA [7]. Lung adenocarcinoma excludes AIS from the malignant category, categorizing it as AAH in the new classification [7]. This revised classification introduces challenges in diagnosing and categorizing precursor lesions and microinvasive cancers. Moreover, for GGOs suspected to be AIS, immediate surgical removal is not recommended. Instead, priority is given to follow-up observation until there is evident malignant transformation before considering surgical intervention in the 2023 NCCN lung cancer surveillance guidelines [8]. Therefore, accurate diagnosis and risk prediction of GGOs hold paramount clinical significance for decision-making according to the new WHO classification [9, 10].

Due to the heterogeneity, low density, and uneven density within GGOs tissue, the success rate of pathological biopsy is compromised [11, 12]. Qualitative diagnosis of GGOs through clinical biopsy is challenging [5]. GGOs may exhibit indolent growth, and current clinical practice involves periodic follow-up for diagnosis. Prolonged undiagnosed follow-up imposes substantial psychological burden and stress on patients. Achieving precise diagnosis of GGOs through imaging techniques is a current research focus. Positron emission tomography-computed tomography (PET/CT) with ^{18}F -FDG, integrating PET and CT imaging, provides both anatomical and molecular

information in a single scan, offering a crucial method for predicting malignant risk in pulmonary nodules [13, 14]. However, ^{18}F -FDG PET/CT with visual analysis combined with semi-quantitative indicators still has some limitations in early diagnosis of GGOs [15].

Artificial intelligence has garnered significant attention for its crucial potential in diagnosing GGOs [16–18]. Most previous studies have been conducted based on CT images [19–21], and there is currently a lack of research on artificial intelligence methods for predicting the malignant risk of GGOs based on ^{18}F -FDG PET/CT images, specifically in the context of the new WHO classification for GGOs. The 3D nnU-net, an automatic segmentation method based on convolutional neural networks, exhibits excellent generalization capabilities, adapting its architecture to different tasks and outperforming many existing methods in medical image processing in terms of accuracy, reliability, and efficiency [22–24]. Herein, we employed the 3D nnU-net with a majority voting method to predict the malignant risk of GGOs according to the new WHO classification based on dual-time-point ^{18}F -FDG PET/CT images.

Methods

Patient selection

In this retrospective single-center study, we recruited 316 patients with GGOs who underwent dual-phase ^{18}F PET/CT imaging at the department of nuclear medicine in the First Affiliated Hospital of China Medical University between June 1, 2015, and July 20, 2023. The study focused on patients exhibiting CT imaging characteristics indicative of GGOs. This research received approval from the institutional ethics committee for retrospective analysis and did not require informed consent. Inclusion criteria were as follows: (1) Stage I lung adenocarcinoma; (2) Pulmonary nodules presenting as GGOs; (3) Lesion diameter within the range of [5 mm, 30 mm]; (4) Availability of postoperative pathological diagnosis, and the time interval between PET/CT examination and pathology sample acquisition was less than 90 days, or patients with a clinical follow-up with CT exceeding 3 years. Exclusion criteria were: (1) Antitumor therapy (including neoadjuvant chemotherapy); (2) Incomplete clinical or imaging data.

Datasets

A preliminary cohort study was conducted on 311 patients with GGOs who underwent ^{18}F -FDG PET/CT scans using Siemens Biograph MCT. The scans were performed 60 min after intravenous injection of 5.5 MBq/kg

^{18}F -FDG. Each bed position acquired ^{18}F -FDG PET/CT data over a 90-second interval. Thin-section CT images of the lungs were obtained without contrast. A delayed scan was conducted 120 min after the initial scan. The voxel size for dual-phase PET was set at 4 mm \times 4 mm \times 4 mm, and CT voxel size was 0.68 mm \times 0.68 mm \times 0.7 mm. All scans were supervised by an experienced nuclear medicine technologist with 20 years of chest imaging expertise. The DICOM files were processed using LIFEX software (<https://www.lifexsoft.org/>) for delineation of GGOs in both dual-phase PET and CT images. For PET images, a semi-automatic delineation was performed using the Threshold 41% tool. For CT images, manual delineation was carried out on lung window settings (window level: -450, window width: 1500) using the pencil 2D tool for layer-by-layer delineation.

Based on follow-up and pathological data, the collected data were divided into two groups: the low-risk nodule group (precursor glandular lesion, follow up remains unchanged) and the high-risk nodule group (MIA, IA). Following the initial ^{18}F -FDG PET/CT examination, patients underwent regular lung CT scans during a 3-year follow-up period. For pulmonary GGOs within this follow-up, if there were no changes in the nodule diameter and composition, the nodule was classified as low risk. The final study included a total of 397 GGOs, comprising 118 in the low-risk group and 279 in the high-risk group. The specific number of nodules for each category is detailed in Fig. 1.

Method

The workflow of the model is depicted in Fig. 2 and comprises two primary components: data preprocessing, segmentation GGOs and prediction of malignant risk for GGOs. All experiments were conducted on a single Nvidia GeForce RTX 3090. The model was constructed using PyTorch 1.4.3 and Python 3.8.

Data preprocessing

In the data preprocessing phase, several key steps were undertaken to ensure uniformity and relevance across ^{18}F -FDG-PET, delayed-phases PET, and thin-section CT images. The ^{18}F -FDG-PET, delayed-phases PET, and thin-section CT scans were resampled to achieve consistency in voxel dimensions. The resampling method employed was trilinear interpolation, a technique commonly used to estimate voxel values at non-grid positions based on surrounding grid points. This ensured alignment and uniformity across the different imaging modalities. The central points of the GGOs were identified, and three-dimensional blocks (32 \times 32 \times 32) were extracted from ^{18}F -FDG-PET, delayed-phases PET, and CT images centered around these points. This block size was chosen to focus on the GGOs, removing irrelevant areas outside the region of interest. The CT images were subjected to windowing with a window width of 1500 and a window level of -450. This specific setting was chosen to enhance the visualization of pulmonary details within the CT images, ensuring clarity and highlighting relevant anatomical structures. The entire dataset, encompassing ^{18}F -FDG-PET, delayed-phases PET, and CT images, underwent a z-score standardization process.

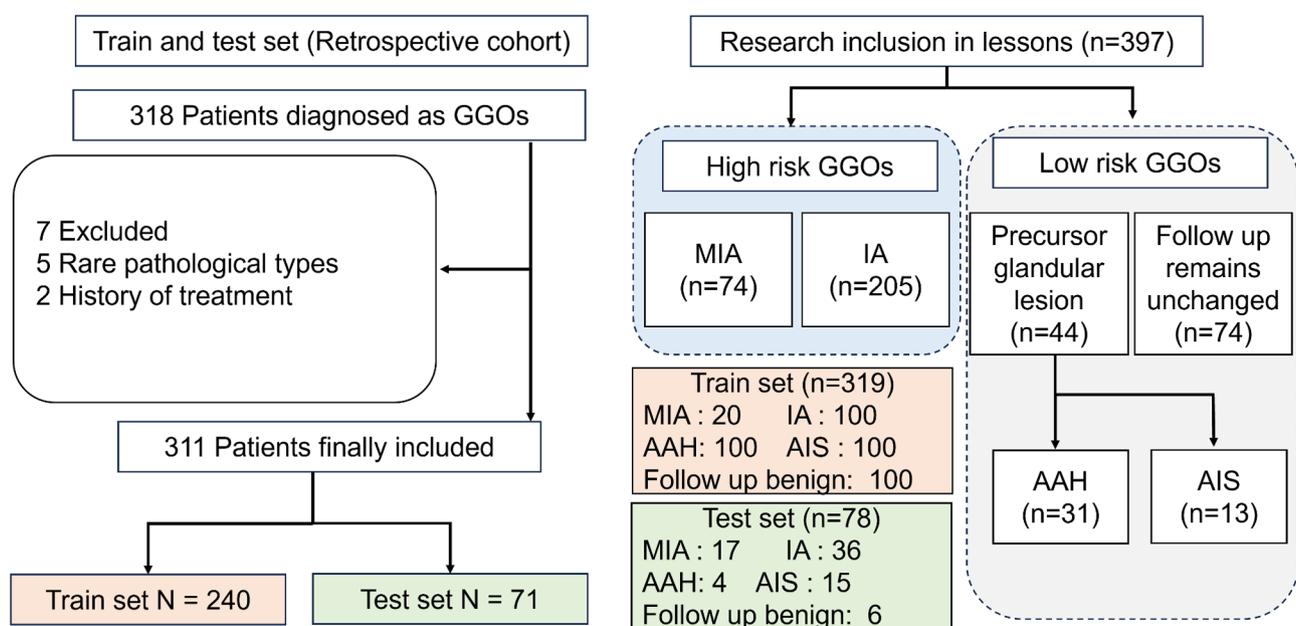


Fig. 1 A flowchart of the patients included and excluded, a grouping diagram of different types of GGOs in this study

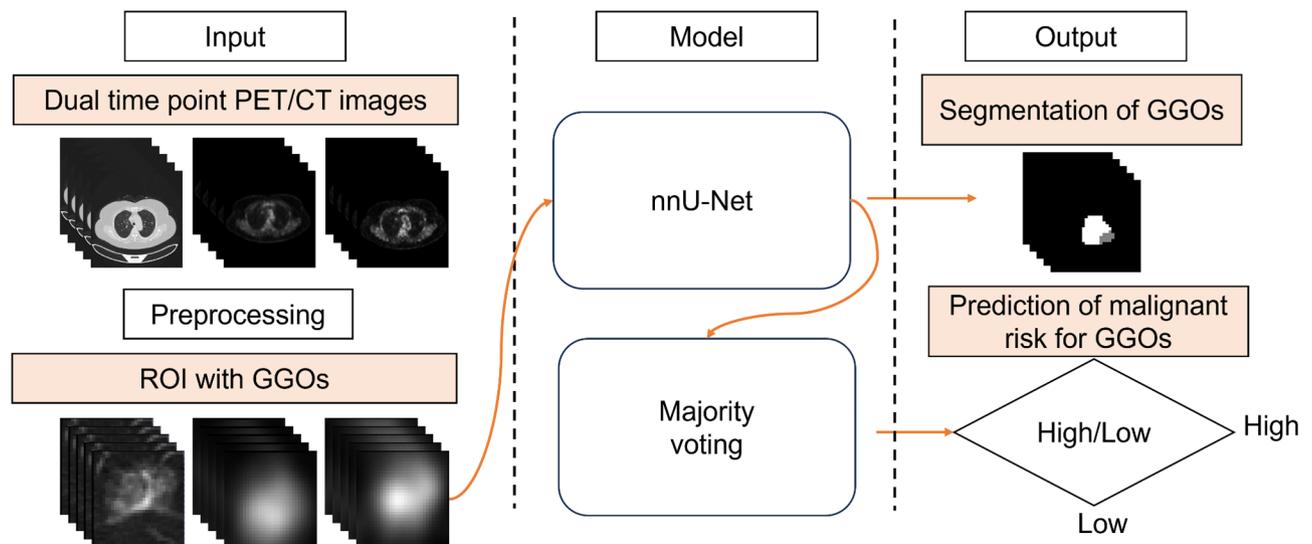


Fig. 2 The workflow of the model for segmenting GGOs and predicting malignant risk

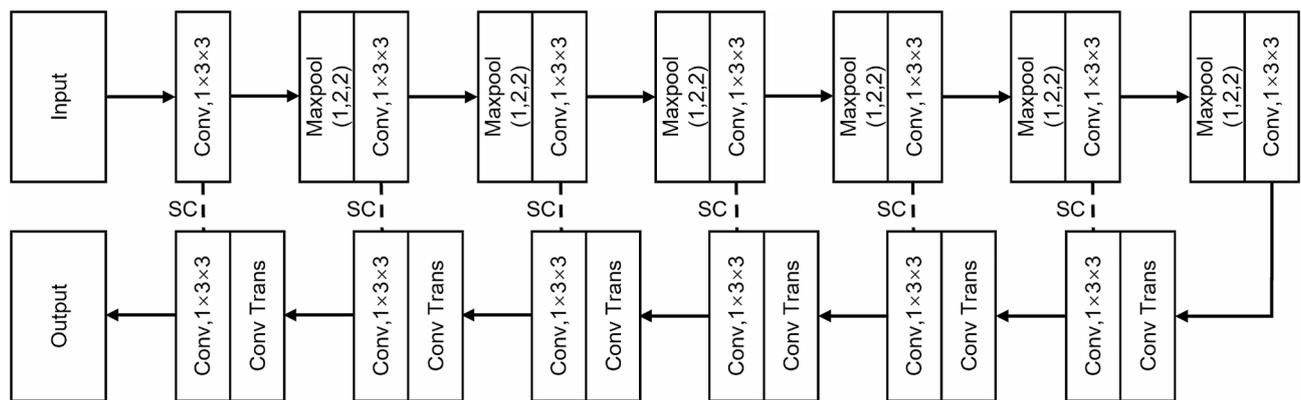


Fig. 3 The architectural framework of the 3D nnU-Net comprises an encoder and a decoder intricately linked through skip connections

This involved subtracting the mean voxel value from each voxel and then dividing by the standard deviation. The z-score standardization ensured that each modality had a mean of 0 and a standard deviation of 1, facilitating model convergence during subsequent training and analysis.

Architecture of model

The classification network for the GGOs is depicted in Fig. 3. The input consists of the 3-channel ROI of the dual-phase ¹⁸F-FDG PET images and thin-section lung CT images, which were extracted in the previous step. The output of the network provides the GGOs segmentation (voxel-wise predicted labels of invasive GGOs and inert GGOs with the same imaging resolution as the input images). The primary architecture of the network is a 3D U-net [25], which belongs to the 3D nnU-net framework [26]. The 3D nnU-net is widely recognized for its high performance on 23 public datasets. It is a deep learning-based segmentation technique that has been

extensively employed for medical imaging segmentation tasks and has demonstrated superiority over numerous existing approaches without requiring manual intervention. The network is composed of an encoder-decoder structure with skip connections linking the two pathways. The encoder includes 5 levels of convolutional layers with strided convolution of the same down-sampling rate. The decoder follows the same design that incorporates transpose convolution up-sampling on the concatenated up-sampled features from the lower level and the skip features from the encoder branch at the same level. After every convolution operation, Leaky ReLU with a slope of 0.01 [27] and batch normalization [28] were applied after every convolution operation. The result of the 3D nnU-net is input into the majority voting module, a straightforward yet effective classification algorithm. High-risk GGOs component prediction voxels exceeding 50% of total voxels indicate malignancy, with values below 50% classifying the lesion as low risk. The

assessment of malignancy risk in the testing set (high-risk GGOs or low-risk GGOs) is further determined.

Training and testing details

A loss function in machine learning and deep learning estimates the disparity between predicted and actual values. It quantifies predictive accuracy during model training, guiding parameter optimization for enhanced performance. Given the use of the 3D nnU-net model for segmenting nodule volumes in this study, the loss function combines Dice loss and cross-entropy loss. Loss computation spans the entire tumor region, yielding the Dice coefficient to assess segmentation outcomes. The 3D nnU-net employs stochastic gradient descent with an initial learning rate of 0.01 and Nesterov momentum of 0.99. Training consists of 1000 epochs, with each epoch's sample size equal to the total training set samples. Data augmentation is applied to each patch during training. The optimal model, determined post-convergence of training loss, is applied to the testing set, generating segmentation results for high-risk and low-risk GGOs components.

Table 1 Clinical characteristics for training and testing sets

		Training set (n=239, m=318)	Testing set (n=72, m=798)	p
Gender, n	Male	84	29	
	Female	155	43	0.56
Age (years, mean)		60.87	59.07	0.17
Smoking history, n	Yes	48	16	
	No	191	56	0.82
Lung cancer family history, n	Yes	41	15	
	No	198	57	0.59
GGOs risk Stratification	High risk	226	53	
	Low risk	92	26	0.58
GGOs Position	Left upper lobe	84	18	
	Left lower lobe	41	6	
	Right upper lobe	110	29	
	Right middle lobe	26	7	
	Right lower lobe	57	19	0.53
Early-phase PET SUV _{max} (kBq/ml/MBq/kg)		2.42	1.99	0.13
Delayed-phase PET SUV _{max} (kBq/ml/MBq/kg)		3.03	2.46	0.17
GGOs diameter		17.88	16.67	0.20

Statistical analysis

For the high risk GGOs and low risk GGOs groups, significant differences were assessed by the chi-square test for categorical variables and the t-test for continuous variables. To assess the performance of the model across various modalities, three models were trained: a thin-section lung CT model, an early-phase ¹⁸F-FDG PET/CT model, and a dual-time-point ¹⁸F-FDG PET/CT model, all validated on the testing set. All models adopted the same 3D nnU-net architecture, with adjustments made to the input and layer structures based on the number of modalities in the input images. Simultaneously, to assess the clinical applicability of the models, we compared the results with those of three nuclear medicine diagnostic physicians. These physicians (resident, junior and expert), possessed over 2, 5, and 10 years of PET/CT interpretation experience, respectively. After anonymizing the imaging data, the diagnostic physicians independently interpreted the nature of pulmonary nodules in thin-section lung CT, early-phase PET, and delayed-phase PET images. They subsequently integrated the findings from all three modalities for a comprehensive qualitative assessment of pulmonary nodules. The final qualitative results were compared with the predictions of the dual-time-point ¹⁸F-FDG PET/CT model on the testing set. Performance metrics, including accuracy, specificity, sensitivity, positive predictive value (PPV), negative predictive value (NPV), F1 score, and the area under the ROC curve, were employed to evaluate model efficacy. All statistical analyses were performed using Python (version 3.8).

Results

Characteristics of patients with GGOs

This study enrolled a total of 311 patients with 397 nodules. Clinical data and nodule characteristics were subjected to analysis, encompassing variables such as gender, age, smoking history, family history of lung cancer, nodule size, nodule location, maximum standardized uptake value (SUV_{max}) in the early and delayed phases. The baseline clinical features of each patient group are presented in Table 1. Among the patients, 198 were female, constituting 63.7% of the cohort; 247 were non-smokers, representing 79.4%; and 255 had no family history of lung cancer, accounting for 82%. The mean age of the patients was 60.46 years. There were 279 instances of high-risk GGOs, including 74 MIA cases and 205 IA cases, and 118 instances of low-risk GGOs, comprising 31 AAH cases, 13 AIS cases, and 74 cases with unchanged follow-up lesions. The dataset was randomly divided into a training set comprising 239 patients (318 lesions) and a testing set comprising 72 patients (79 lesions), Clinical characteristics of patients in both sets are summarized in Table 1. Statistical comparisons between the training set

and testing set involved the chi-square test for categorical variables and the t-test for continuous variables. The obtained P-values for each clinical feature in both sets were all above 0.05, indicating an absence of statistically significant differences in the distribution of clinical characteristics between the two sets.

GGOs segmentation and prediction malignant risk

The Dice score for the dual-time-point ^{18}F -FDG PET/CT model was 0.84 ± 0.02 , indicating a close alignment between the model's segmentation results and the labels. The model ultimately provides segmentation results for GGOs on thin-section lung CT, divided into two components: one for the segmentation results of low-risk areas within the nodules and the other for the segmentation results of high-risk areas within the nodules. The segmentation efficacy of the dual-time-point ^{18}F -FDG PET/

CT model is illustrated in Fig. 4. In Low risk 1 and 2, the proportion of model-segmented high-risk areas does not exceed the cutoff value, hence predicting these nodules as low-risk. Conversely, in High risk 1 and 2, the proportion of model-segmented high-risk areas surpasses the cutoff value, leading to their prediction as high-risk nodules. After employing a majority voting method on the segmentation results output, among the testing set samples, 49 were predicted as high-risk nodules with a proportion exceeding the cutoff value, while 30 were classified as low-risk nodules falling below the cutoff value.

Performance of the multi-modal model

A combination of various imaging modalities resulted in the construction of three models: the thin-section lung CT model, early-phase ^{18}F -FDG PET/CT model, and dual-time-point ^{18}F -FDG PET/CT model. The diagnostic

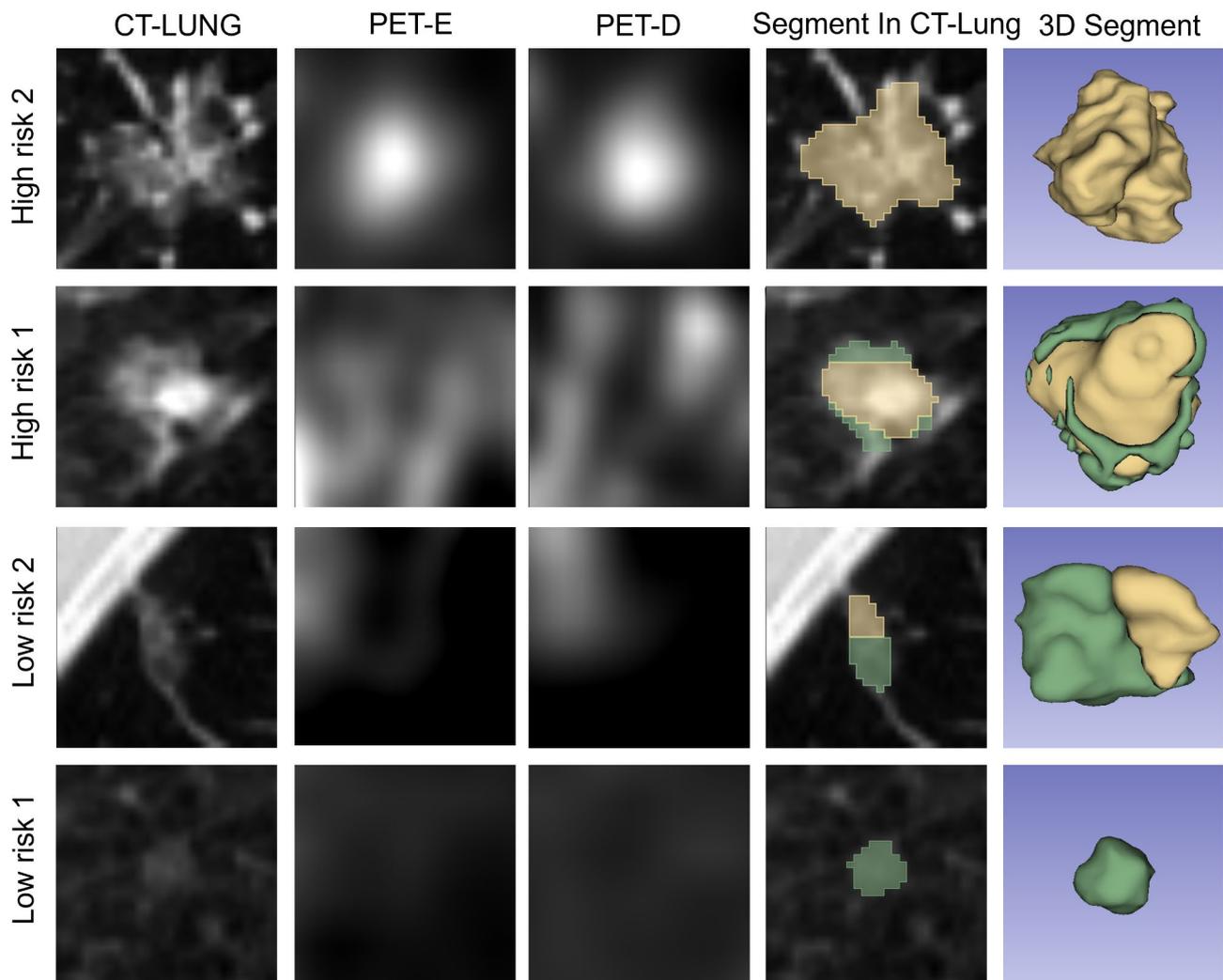
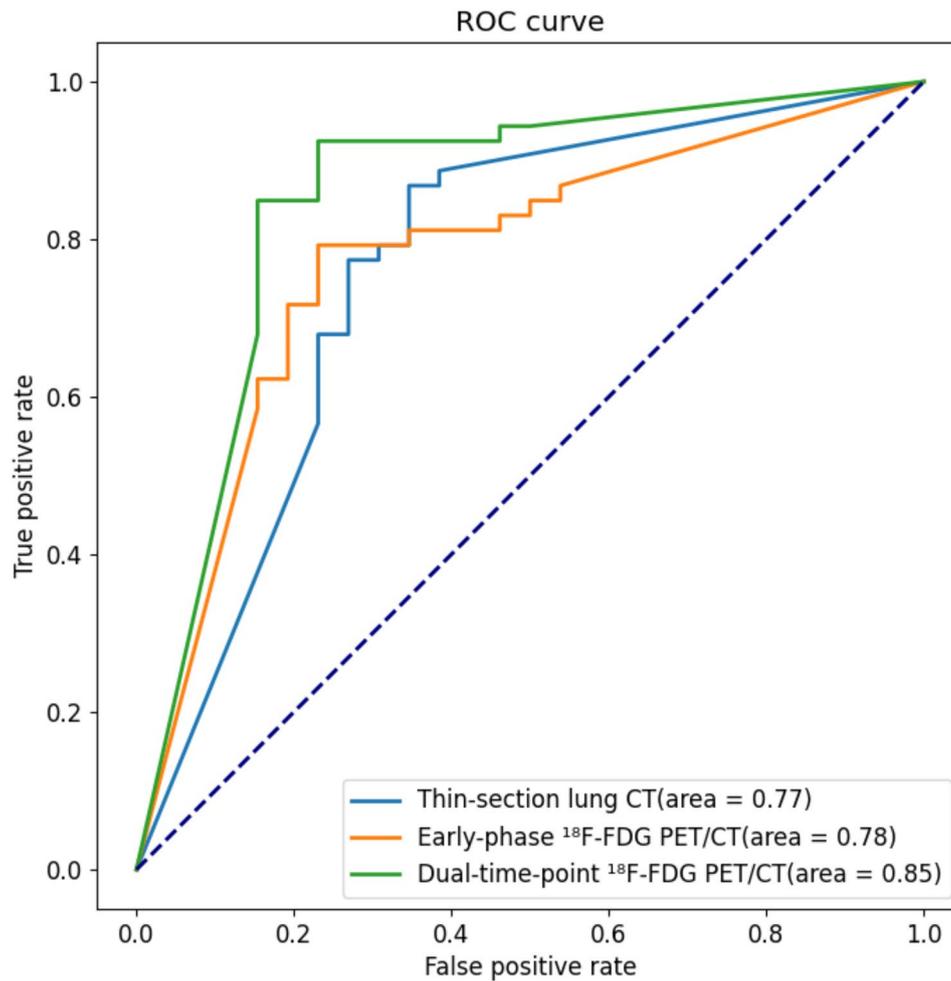


Fig. 4 Segmentation Results of High-Risk and Low-Risk GGOs High risk1 and High risk2 represent high-risk nodules, while Low risk1 and Low risk2 depict low-risk nodules. "Segment In CT-Lung" displays the segmentation results overlaid on thin-section lung CT images, and "3D Segment" provides a three-dimensional visualization of the segmented regions. In the visualizations, the green areas represent the low-risk regions segmented by the model, while the yellow areas indicate the high-risk regions segmented by the model

Table 2 The performance of the thin-section lung CT model, the early-phase ^{18}F -FDG PET/CT model and the dual-time-point ^{18}F -FDG PET/CT model was assessed using 3D nnu-net within the testing set

	Accuracy	Specificity	Sensitivity	PPV	NPV	F1 score	AUC
Thin-section lung CT	73.42%	73.08%	73.58%	84.78%	57.57%	0.79	0.77
Early-phase ^{18}F -FDG PET/CT	78.48%	76.92%	79.25%	87.50%	64.52%	0.83	0.78
Dual-time-point ^{18}F -FDG PET/CT	84.81%	84.62%	84.91%	91.84%	73.33%	0.88	0.85

**Fig. 5** Performance of 3D nnU-net in the prediction of high risk and low risk. AUCs for the thin-section lung CT model (orange curve), the early-phase ^{18}F -FDG PET/CT model (green curve) and the dual-time-point ^{18}F -FDG PET/CT model (blue curve) were 0.77, 0.78, and 0.85, respectively

performance indicators and ROC curves for each model on the testing set are presented in Table 2; Fig. 5. Comparing these models with the thin-section lung CT model, it is evident that, with the addition of PET-E and PET-D images, the performance of each model sequentially improves. The accuracy and area under the ROC curve for the thin-section lung CT model are 73.42% and 0.77, respectively. The early-phase ^{18}F -FDG PET/CT model demonstrates an accuracy of 78.48% and an area under the ROC curve of 0.78. Notably, the dual-time-point ^{18}F -FDG PET/CT model achieves the highest performance metrics, with an accuracy of 84.81% and an area under the ROC curve reaching 0.85.

Comparison of assessment results with clinical doctors

Comparison of the interpretation results from three nuclear medicine diagnostic physicians with the predictive outcomes of the dual-time-point ^{18}F -FDG PET/CT model for the testing set is presented in Table 3. The diagnostic accuracy of the dual-time-point ^{18}F -FDG PET/CT model was 84.81%. In comparison, the accuracy for resident physician, junior physician, and expert physician was 67.09%, 74.68%, and 78.48%, respectively. The diagnostic accuracy of the dual-time-point ^{18}F -FDG PET/CT model significantly surpassed that of the nuclear medicine diagnostic physicians. The diagnostic sensitivity of the dual-time-point ^{18}F -FDG PET/CT model was 84.91%,

Table 3 Comparison of the performance between nuclear medicine physicians' assessments and the dual-time-point ^{18}F -FDG PET/CT model

	Accuracy	Specificity	Sensitivity	PPV	NPV	F1 score	AUC
Resident physician	67.09%	84.62%	58.49%	88.57%	86.67%	0.70	0.72
Junior physician	74.68%	76.92%	73.58%	86.67%	58.82%	0.80	0.75
Expert physician	78.48%	84.62%	75.47%	90.91%	62.86%	0.82	0.80
Dual-time-point ^{18}F -FDG PET/CT	84.81%	84.62%	84.91%	91.84%	73.33%	0.88	0.85

whereas the sensitivity for resident physician, junior physician, and expert physician was 58.49%, 73.58%, and 75.47%, respectively. Notably, the diagnostic sensitivity of the dual-time-point ^{18}F -FDG PET/CT model was markedly higher than that of the nuclear medicine diagnostic physicians.

Discussion

In this study, we employed the 3D nnU-net with a majority voting method to predict the malignant risk of GGOs according to the new WHO classification based on dual-time-point ^{18}F -FDG PET/CT images. The 3D nnU-net architecture, initially proposed by Isensee et al. [26], is a deep learning segmentation method that automatically configures itself, covering preprocessing, network architecture, model training, and post-processing to address any new task. Several studies have also employed the 3D nnU-net for classification tasks. Yin et al. [22] utilized it to differentiate pelvic and sacral osteosarcoma from Ewing's sarcoma, Cheng et al. [23] used it for predicting the molecular subtypes of posterior fossa ependymomas, and Cai et al. [24] applied it to classify extramural vascular invasion status. In each case, they achieved excellent performance. Our study employs the 3D nnU-net in conjunction with a majority voting approach to predict the malignant risk of nodules. This involves the automated segmentation of GGOs on dual-time-point ^{18}F -FDG PET/CT images, followed by classification via majority voting without manual intervention, forming an end-to-end process. Considering that deep learning models typically require large datasets, this voxel prediction method based on the majority principle greatly optimizes the accuracy of this classification task, enhancing model stability and efficiency [29]. With this model, the automatic discrimination of GGOs malignant risk is achieved by inputting raw images without manual delineation or any intervention. Furthermore, it can extract contextual information from consecutive slices, better capturing spatial features in lung nodules [30]. Additionally, the multi-scale feature extraction of 3D nnU-net enables it to adapt to nodules of different sizes, eliminating the disadvantages of conventional models in predicting small lung nodules [31]. Finally, 3D nnU-net reduces overfitting and training time through weight fine-tuning via transfer

learning, improving algorithm efficiency and achieving high model accuracy with a smaller sample size.

In terms of GGOs segmentation, the obtained Dice coefficient of 0.84 surpasses previously reported DICE coefficients for GGOs segmentation using CT (Dice scores ranging from 0.73 to 0.82) [32, 33], underscoring the clinical potential of the model in GGOs segmentation. The interpretation of the segmentation results reveals that the model delineates the 3D morphology of lung nodules, indicating the location and proportion of high-risk and low-risk regions. The high-risk areas are concentrated in the solid density, lobulation, spiculation, and pleural traction areas of lung nodules, showing increased FDG uptake and delayed-phases PET imaging uptake (as shown in Fig. 4). This matches the malignant characteristics judged in clinical practice [34, 35], providing support for the improved efficacy of the model.

Regarding model predictive performance, the dual-time-point ^{18}F -FDG PET/CT model in this study achieved an accuracy, sensitivity, specificity, and AUC of 84.81%, 84.91%, 84.62%, and 0.85, respectively. Compared to the early-phase ^{18}F -FDG PET/CT model and thin-section lung CT model, it outperformed in all metrics, and with the addition of delayed-phases PET, the model performance sequentially increased. Particularly, the inclusion of delayed-phases PET demonstrated the greatest enhancement in model performance, indicating its value in predicting GGOs malignant risk. Additionally, when comparing our dual-timepoint PET/CT model with existing CT-based methods for GGO malignant risk assessment (AAH and AIS vs. MIA and IA) reported by Xu et al. [36] (accuracy: 82.2%, AUC: 0.831) and Wang et al. [37] (accuracy: 73.4%, AUC: 0.813), our approach demonstrated increased accuracy and AUC. This gain may be related to the upregulation of glucose consumption by malignant cells to obtain more proliferative energy, leading to the concentration of ^{18}F -FDG in delayed-phases PET [38, 39]. In contrast to other convolutional neural networks, the diagnostic performance of the dual-time-point ^{18}F -FDG PET/CT model is more balanced, with other convolutional neural networks exhibiting lower overall specificity, indicating limited efficacy in predicting low-risk GGOs. Furthermore, when compared with nuclear medicine diagnostic physicians,

the dual-time-point ^{18}F -FDG PET/CT model achieved higher diagnostic efficacy, especially surpassing more experienced physicians, confirming its clinical application value. Previous studies using radiomics and deep learning methods for GGOs classification under early-phase ^{18}F -FDG PET/CT achieved an accuracy of around 80% [40, 41]. However, these models showed significant variations in sensitivity and specificity, and their results were not validated against the diagnostic efficacy of clinical physicians [42, 43].

Clinically, the interpretation of PET/CT images for pulmonary GGOs relies on subjective judgment supported by the experience of diagnostic physicians. Comparing the model with manual readings, significant differences in diagnostic efficacy were observed between low-experience physicians, high-experience physicians, and the model. Moreover, the majority of GGOs exhibits characteristics of indolent growth, with approximately 20% of pure GGOs and 40% of mixed GGOs gradually increasing or adding solid components over time, while the majority remains unchanged for several years. For distinguishing between growing and non-growing GGOs, a reasonable benchmark is a volume doubling time (VDT) of approximately 600 to 900 days for pure GGOs and 300 to 450 days for mixed GGOs during a 3-year follow-up observation [44]. The specific time intervals and frequency of follow-up in clinical practice vary among individuals, requiring physicians to perform comparative analyses of nodules at different time points to determine their nature.

In lung cancer screening, the occurrence of multifocal GGOs is highly prevalent [45]. For cases involving multiple GGOs that necessitate clinical intervention, guidelines recommend a preference for surgical procedures. The principle is to prioritize the treatment of primary lesions while addressing secondary lesions [46]. But performing pathological evaluation on every lesion is neither feasible nor resource-efficient. Furthermore, the presence of multiple lesions within a single patient may introduce a “clustering effect,” potentially leading to overfitting to patient-specific characteristics in predictive models [47]. Clarifying these issues through future research could refine risk stratification and guide more nuanced clinical decision-making for multifocal GGO management.

However, it is essential to acknowledge certain limitations in this study. The use of a retrospective single-center dataset resulted in a relatively small sample size. Further extensive prospective multicenter research is necessary to evaluate the generalizability of the established model in clinical practice.

Conclusions

In conclusion, our study revealed the effectiveness of utilizing dual-time-point ^{18}F -FDG PET/CT images along with the 3D nnU-net and a majority voting method for predicting the malignant risk of GGOs based on the new WHO classification. This methodology serves as a valuable adjunct for physicians in the risk prediction and assessment of GGOs.

Abbreviations

GGOs	Ground-glass nodules
AAH	Atypical adenomatoid hyperplasia
AIS	Adenocarcinoma in situ
MIA	Microinvasive adenocarcinoma
IA	Invasive adenocarcinoma
WHO	World Health Organization
PET/CT	Positron emission tomography-computed tomography
PPV	Positive predictive value
NPV	Negative predictive value
SUV_{max}	Maximum standardized uptake value

Acknowledgements

Not applicable.

Author contributions

All authors contributed to the study conception and design. Material preparation, data collection, and analysis was performed by Yaming, L., X.L. and J.W. The first draft of the manuscript was written by Yuhang, L., B.D. and J.W., and all the authors commented on following updated versions of the manuscript. All the authors read and approved the final version of the manuscript.

Funding

This work was supported by National Natural Science Foundation of China (grant no. 82272043 and 82171987).

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the First Hospital of China Medical University (2024/23). The requirement for patients' informed consent was waived in light of the retrospective nature of the study.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Nuclear Medicine, The First Hospital of China Medical University, No. 155 Nanjing St, Shenyang 110001, China

Received: 8 May 2024 / Accepted: 4 February 2025

Published online: 18 February 2025

References

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer statistics 2020: GLOBOCAN estimates of incidence and Mortality Worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021;71(3):209–49.

2. Munden RF, Black WC, Hartman TE, et al. Managing incidental findings on thoracic CT: lung findings. A White Paper of the ACR Incidental findings Committee. *J Am Coll Radiol*. 2021;18(9):1267–79.
3. Yoon HY, Bae JY, Kim Y, et al. Risk factors associated with an increase in the size of ground-glass lung nodules on chest computed tomography. *Thorac Cancer*. 2019;10(7):1544–51.
4. MacMahon H, Naidich DP, Goo JM, et al. Guidelines for management of Incidental Pulmonary nodules detected on CT images: from the Fleischner Society 2017. *Radiology*. 2017;284(1):228–43.
5. Gould MK, Donington J, Lynch WR et al. Evaluation of individuals with pulmonary nodules: when is it lung cancer? Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*, 2013. 143(5 Suppl): e93S-e120S.
6. Zhang Y, Fu F, Chen H. Management of Ground-Glass opacities in the Lung Cancer Spectrum. *Ann Thorac Surg*. 2020;110(6):1796–804.
7. Board WC. o.T.E., Thoracic Tumours. 5 ed. WHO Classification of Tumours. Vol. 5. 2021: International Agency for Research on Cancer, January 2021.
8. NCCN. *Lung Cancer Screening (Version 2.2024)*. October 18, 2023; Available from: <http://www.nccn.org>
9. Yoshida C, Yokomise H, Ibuki E, et al. High-grade tumor classified by new system is a prognostic predictor in resected lung adenocarcinoma. *Gen Thorac Cardiovasc Surg*. 2022;70(5):455–62.
10. Tsao MS, Marguet S, Le Teuff G, et al. Subtype classification of lung Adenocarcinoma Predicts Benefit from Adjuvant Chemotherapy in patients undergoing complete resection. *J Clin Oncol*. 2015;33(30):3439–46.
11. Liu S, Wang R, Zhang Y, et al. Precise diagnosis of Intraoperative Frozen section is an effective method to Guide Resection Strategy for Peripheral Small-Sized Lung Adenocarcinoma. *J Clin Oncol*. 2016;34(4):307–13.
12. Yeh YC, Nitadori J, Kadota K, et al. Using frozen section to identify histological patterns in stage I lung adenocarcinoma of ≤ 3 cm: accuracy and interobserver agreement. *Histopathology*. 2015;66(7):922–38.
13. Kim HS, Lee KS, Ohno Y, et al. PET/CT versus MRI for diagnosis, staging, and follow-up of lung cancer. *J Magn Reson Imaging*. 2015;42(2):247–60.
14. Kim TJ, Kim CH, Lee HY, et al. Management of incidental pulmonary nodules: current strategies and future perspectives. *Expert Rev Respir Med*. 2020;14(2):173–94.
15. Chen H, Pang Y, Meng T, Yu X, Sun L. ^{18}F -FDG and ^{68}Ga -FAPI PET/CT evaluation of Ground-opacity nodule. *Clin Nucl Med*. 2021;46(5):424–6.
16. Schwyzer M, Martini K, Benz DC, et al. Artificial intelligence for detecting small FDG-positive lung nodules in digital PET/CT: impact of image reconstructions on diagnostic performance. *Eur Radiol*. 2020;30(4):2031–40.
17. Apostolopoulos ID, Pintelas EG, Livieris IE, et al. Automatic classification of solitary pulmonary nodules in PET/CT imaging employing transfer learning techniques. *Med Biol Eng Comput*. 2021;59(6):1299–310.
18. Park YJ, Choi D, Choi JY, et al. Performance evaluation of a Deep Learning System for Differential diagnosis of Lung Cancer with Conventional CT and FDG PET/CT using transfer learning and Metadata. *Clin Nucl Med*. 2021;46(8):635–40.
19. Huang H, Zheng D, Chen H, Wang Y, Chen C, Xu L, Li G, Wang Y, He X, Li W. Fusion of CT images and clinical variables based on deep learning for predicting invasiveness risk of stage I lung adenocarcinoma. *Med Phys*. 2022;49(10):6384–94.
20. Zhu M, Yang Z, Wang M, Zhao W, Zhu Q, Shi W, Yu H, Liang Z, Chen L. A computerized tomography-based radiomic model for assessing the invasiveness of lung adenocarcinoma manifesting as ground-glass opacity nodules. *Respir Res*. 2022;23(1):96.
21. Yu Y, Wang N, Huang N, Liu X, Zheng Y, Fu Y, Li X, Wu H, Xu J, Cheng J. Determining the invasiveness of ground-glass nodules using a 3D multi-task network. *Eur Radiol*. 2021;31(9):7162–71.
22. Yin P, Wang W, Wang S, Liu T, Sun C, Liu X, Chen L, Hong N. The potential for different computed tomography-based machine learning networks to automatically segment and differentiate pelvic and sacral osteosarcoma from Ewing's sarcoma. *Quant Imaging Med Surg*. 2023;13(5):3174–84.
23. Cheng D, Zhuo Z, Du J, Weng J, Zhang C, Duan Y, Sun T, Wu M, Guo M, Hua T, Jin Y, Peng B, Li Z, Zhu M, Imami M, Bettgeowda C, Sair H, Bai HX, Barkhof F, Liu X, Liu Y. A fully automated Deep-Learning Model for Predicting the Molecular subtypes of posterior Fossa Ependymomas using T2-Weighted images. *Clin Cancer Res*. 2024;30(1):150–8.
24. Cai L, Lambregts DMJ, Beets GL, Mass M, Pooch EHP, Guérendel C, Beets-Tan RGH, Benson S. An automated deep learning pipeline for EMVI classification and response prediction of rectal cancer using baseline MRI: a multi-centre study. *NPJ Precis Oncol*. 2024;8(1):17.
25. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: learning dense volumetric segmentation from sparse annotation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2016: 19th International Conference, Athens, Greece, October 17–21, 2016, Proceedings, Part II 19* (pp. 424–432).
26. Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat Methods*. 2021;18(2):203–11.
27. Maas AL, Hannun AY, Ng AY. Rectifier nonlinearities improve neural network acoustic models. *Proc. icml*. 2013; 30:1.
28. Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *International conference on machine learning*. 2015: 448–456.
29. Sohsah GN, Ibrahimzada AR, Ayaz H, et al. Scalable classification of organisms into a taxonomy using hierarchical supervised learners. *J Bioinform Comput Biol*. 2020;18(5):2050026.
30. Zhao C, Han J, Jia Y et al. Lung nodule detection via 3D U-Net and contextual convolutional neural network. in. *2018 International conference on networking and network applications (NaNA)*. 2018. IEEE.
31. Falk T, Mai D, Bensch R, et al. U-Net: deep learning for cell counting, detection, and morphology. *Nat Methods*. 2019;16(1):67–70.
32. Chen H, Liu J, Lu L, et al. Volumetric segmentation of ground glass nodule based on 3D attentional cascaded residual U-Net and conditional random field. *Med Phys*. 2022;49(2):1097–107.
33. Dong T, Wei L, Ye X, et al. [Segmentation of ground glass pulmonary nodules using full convolution residual network based on atrous spatial pyramid pooling structure and attention mechanism]. *Sheng Wu Yi Xue Gong Cheng Xue Za Zhi*. 2022;39(3):441–51.
34. Qiu ZX, Cheng Y, Liu D, et al. Clinical, pathological, and radiological characteristics of solitary ground-glass opacity lung nodules on high-resolution computed tomography. *Ther Clin Risk Manag*. 2016;12:1445–53.
35. Evangelista L, Panunzio A, Scagliori E, et al. Ground glass pulmonary nodules: their significance in oncology patients and the role of computer tomography and ^{18}F -fluorodeoxyglucose positron emission tomography. *Eur J Hybrid Imaging*. 2018;2(1):2.
36. Xu Y, Li Y, Yin H, Tang W, Fan G. Consecutive serial non-contrast CT scan-based Deep Learning Model facilitates the prediction of Tumor Invasiveness of Ground-Glass nodules. *Front Oncol*. 2021;11:725599.
37. Wang J, Chen X, Lu H, et al. Feature-shared adaptive-boost deep learning for invasiveness classification of pulmonary subsolid nodules in CT images. *Med Phys*. 2020;47(4):1738–49.
38. Higashi T, Saga T, Nakamoto Y, et al. Relationship between retention index in dual-phase (^{18}F -FDG PET, and hexokinase-II and glucose transporter-1 expression in pancreatic cancer. *J Nucl Med*. 2002;43(2):173–80.
39. Zhuang H, Pourdehnad M, Lambright ES, et al. Dual time point ^{18}F -FDG PET imaging for differentiating malignant from inflammatory processes. *J Nucl Med*. 2001;42(9):1412–7.
40. Shao X, Niu R, Shao X, et al. Application of dual-stream 3D convolutional neural network based on (^{18}F -FDG PET/CT in distinguishing benign and invasive adenocarcinoma in ground-glass lung nodules. *EJNMMI Phys*. 2021;8(1):74.
41. Shao X, Niu R, Shao X, et al. Value of (^{18}F -FDG PET/CT-based radiomics model to distinguish the growth patterns of early invasive lung adenocarcinoma manifesting as ground-glass opacity nodules. *EJNMMI Res*. 2020;10(1):80.
42. Chen KN. The diagnosis and treatment of lung cancer presented as ground-glass nodule. *Gen Thorac Cardiovasc Surg*. 2020;68(7):697–702.
43. Hammer MM, Palazzo LL, Eckel AL, et al. A decision analysis of Follow-Up and Treatment algorithms for Nonsolid Pulmonary nodules. *Radiology*. 2019;290(2):506–13.
44. Kobayashi YT, Mitsudomi. Management of ground-glass opacities: should all pulmonary lesions with ground-glass opacity be surgically resected? *Transl Lung Cancer Res*. 2013;2(5):354–63.
45. Liu Y, Luo H, Qing H, et al. Screening baseline characteristics of early lung cancer on low-dose computed tomography with computer-aided detection in a Chinese population. *Cancer Epidemiol*. 2019;62:101567.
46. Jiang G, Chen C, Zhu Y, et al. [Shanghai Pulmonary Hospital experts Consensus on the management of Ground-Glass nodules suspected as Lung Adenocarcinoma (Version 1)]. *Zhongguo Fei Ai Za Zhi*. 2018;21(3):147–59.

47. Pan Z, Hu G, Zhu Z, et al. Predicting invasiveness of lung adenocarcinoma at chest CT with Deep Learning Ternary classification models. *Radiology*. 2024;311(1):e232057.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.