

Automated Classification of Pulmonary Nodules through a Retrospective Analysis of Conventional CT and Two-phase PET Images in Patients Undergoing Biopsy

Atsushi Teramoto^{1*}, Masakazu Tsujimoto², Takahiro Inoue³, Tetsuya Tsukamoto³, Kazuyoshi Imaizumi³, Hiroshi Toyama³, Kuniaki Saito¹, Hiroshi Fujita⁴

¹ Faculty of Radiological Technology, School of Health Sciences, Fujita Health University, Toyoake, Japan

² Fujita Health University Hospital, Toyoake, Japan

³ School of Medicine, Fujita Health University, Toyoake, Japan

⁴ Department of Electrical, Electronic, and Computer Engineering, Faculty of Engineering, Gifu University, Gifu, Japan

ARTICLE INFO

Article type:
Original article

Article history:
Received: 16 Jul 2018
Revised: 14 Sep 2018
Accepted: 15 Sep 2018

Keywords:
Classification
Computer-aided diagnosis
Lung cancer
Pulmonary nodule
PET/CT

ABSTRACT

Objective(s): Positron emission tomography/computed tomography (PET/CT) examination is commonly used for the evaluation of pulmonary nodules since it provides both anatomical and functional information. However, given the dependence of this evaluation on physician's subjective judgment, the results could be variable. The purpose of this study was to develop an automated scheme for the classification of pulmonary nodules using early and delayed phase PET/CT and conventional CT images.

Methods: We analysed 36 early and delayed phase PET/CT images in patients who underwent both PET/CT scan and lung biopsy, following bronchoscopy. In addition, conventional CT images at maximal inspiration were analysed. The images consisted of 18 malignant and 18 benign nodules. For the classification scheme, 25 types of shape and functional features were first calculated from the images. The random forest algorithm, which is a machine learning technique, was used for classification.

Results: The evaluation of the characteristic features and classification accuracy was accomplished using collected images. There was a significant difference between the characteristic features of benign and malignant nodules with regard to standardised uptake value and texture. In terms of classification performance, 94.4% of the malignant nodules were identified correctly assuming that 72.2% of the benign nodules were diagnosed accurately. The accuracy rate of benign nodule detection by means of CT plus two-phase PET images was 44.4% and 11.1% higher than those obtained by CT images alone and CT plus early phase PET images, respectively.

Conclusion: Based on the findings, the proposed method may be useful to improve the accuracy of malignancy analysis.

► Please cite this paper as:

Teramoto A, Tsujimoto M, Inoue T, Tsukamoto T, Imaizumi K, Toyama H, Saito K, Fujita H. Automated Classification of Pulmonary Nodules through a Retrospective Analysis of Conventional CT and Two-phase PET Images in Patients Undergoing Biopsy. *Asia Ocean J Nucl Med Biol.* 2019; 7(1): 29-37. doi: 10.22038/AOJNMB.2018.12014

* *Corresponding author:* Atsushi Teramoto, Faculty of Radiological Technology, School of Health Sciences, Fujita Health University, Toyoake, Japan. Tel: +81562939415; Fax: +81562934595; Email: teramoto@fujita-hu.ac.jp

© 2019 *mums.ac.ir* All rights reserved.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

Lung cancer is a leading cause of cancer mortality among men and women. This disease is a serious public health problem in many countries (1). When a lung nodule is found during cancer screening, it is important to accurately classify the lesion as benign or malignant in order to institute appropriate therapy and improve the survival rate.

Computed tomography (CT) is often used for lung cancer screening (2). According to the results of a national lung screening trial carried out in the United States (3), screening with low-dose CT scans reduced lung cancer mortality by 20%. Therefore, CT is regarded as a suitable diagnostic tool for the early detection of lung cancer. If a suspicious lesion is found by CT examination, positron emission tomography (PET)/CT examination is performed for detailed analysis. In this combined technique, PET images provide functional information while CT images render anatomical information, thereby facilitating a comprehensive analysis of the malignancy of nodules.

However, fluorodeoxyglucose (FDG) PET images of benign nodules, such as those associated with inflammatory diseases, often exhibit high uptake values similar to the images of malignant nodules; furthermore, their anatomical structures are similar. Therefore, it is often difficult to differentiate between benign and malignant nodules (4). In such cases, a bronchoscopic biopsy is performed; however, the procedure is invasive, and the patient faces great physical hardship.

In these cases, if the CT and PET images can be analysed in detail to quantify the degree of malignancy of the nodules, the need for excessive biopsy with its accompanying physical hardship can be reduced. With this background in mind, the present study was focused on the automated analysis of the malignant potential of the pulmonary nodule using PET/CT images.

Many studies have investigated the benign/malignant differentiation of pulmonary nodules by image analysis (5-10). For instance, Armato et al. proposed the automated analysis of pulmonary nodule using linear discriminant analysis with characteristic features obtained from CT images (5). The evaluation of 470 CT scans in a study revealed that the area under the receiver operating characteristic (ROC) curve was 0.79. In addition, Shen et al. (9) introduced a deep learning model of the multi-crop convolutional neural network to classify pulmonary nodules. The authors used 880 benign nodules and 495 malignant nodules from the lung image database consortium and image

database resource initiative (LIDC/IDRI) dataset and obtained an accuracy of 87.14% with the model.

Nie et al. developed a semi-automated scheme for distinguishing between benign and malignant pulmonary nodules by integrating PET and CT information. They evaluated three computer-aided diagnosis schemes based on an artificial neural network to distinguish between benign and pulmonary nodules using clinical information and image features. They reported that the combined use of PET and CT rendered a higher diagnostic accuracy, compared to the employment of CT alone or PET alone (10).

However, to the best of our knowledge, the automated calculation of the characteristic values of pulmonary nodules based on PET and CT images have not been developed. The automated classification of pulmonary nodules can have a great practical value. Regarding this, the present study involved the proposition of an automated classification scheme of the pulmonary nodule using CT and PET images. The major objective of our study was to develop the characteristic features using both CT and PET images. Furthermore, we also developed a classification method using random forest, which is a kind of ensemble machine learning technique.

In this paper, first, the architecture of the developed classification method is described. In addition, the effectiveness of the classification of pulmonary nodules as evaluated with the original CT and PET image database is discussed.

Methods

This study was approved by the Institutional Review Board. Informed consent was obtained from all patients under the condition that all data were anonymized (No. HM17-002). The current study was conducted on 36 early and delayed phase PET/CT images obtained from patients with a suspected diagnosis of lung cancer. In addition, conventional CT images at maximal inspiration were analysed. The cases were chosen from those whose differential diagnosis was difficult with diagnostic imaging alone, and final diagnosis was made by bronchoscopy and biopsy analysis.

The PET/CT imaging studies were performed by means of Siemens True Point mCT (Siemens). Both images were obtained with a matrix size of 200×200 pixels (voxel size: 4.07×4.07×2.00 mm³, scan time: 2.0 min/table) with free breathing. Image reconstruction was performed using the 3D-OSEM reconstruction algorithm.

In addition, point-spread function, time-of-

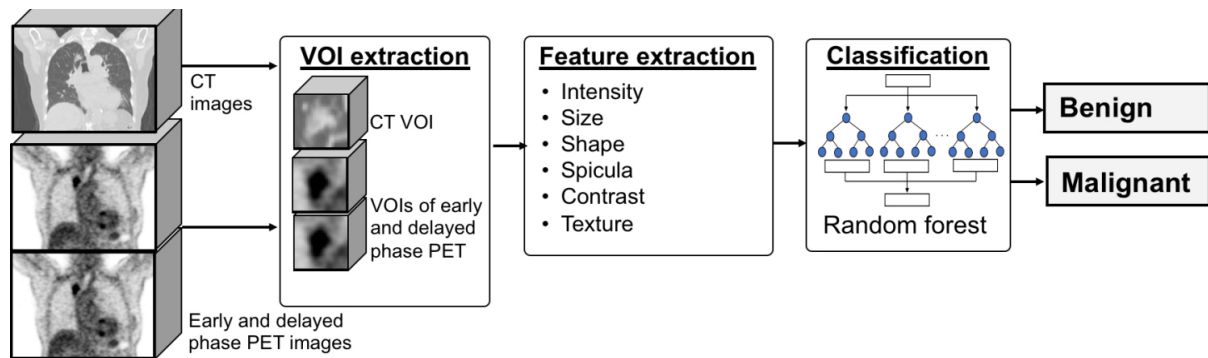


Figure 1. Outline of the proposed method
CT: computed tomography, PET: positron emission tomography

flight correction (PSF+TOF), and attenuation correction were performed by CT images. The PET images were converted to the voxel size of the CT image after reconstruction. Early and delayed PET imaging was performed 60 min and 120 min after the administration of 3.7 MBq/kg of FDG, respectively. These PET and CT images were aligned automatically by the PET/CT scanner.

The conventional CT imaging was performed using Aquilion ONE (Toshiba, Tokyo, Japan) with a matrix size of 512×512 pixels (voxel size: 0.625×0.625×0.500 mm³) with the lung kernel. In case the CT examination was carried out more than once, the images taken with the shortest interval from the PET/CT examination were selected.

Out of a total of 18 benign cases, 13 cases were finally diagnosed by biopsy, and the remaining 5 cases were confirmed to be benign by a follow-up examination of at least 3 years. Out of 25 malignant cases, 1, 4, and 13 cases were small cell carcinoma, squamous cell carcinomas and adenocarcinomas respectively. The mean ages of the patients in the malignant and benign nodule groups were 72.2±7.6 and 65.3±10.2 years, respectively.

Figure 1 depicts an overview of the proposed method. In this method, regions designated as suspicious in the PET and CT images by the doctor were analysed using several characteristic features, and then automatically classified as benign or malignant.

Volume of Interest (VOI) extraction

The position and diameter of the nodule to be analysed using conventional CT and PET/CT images was specified by the physician. Accordingly, the segmentation of the volume of interest (VOI) around the pulmonary nodule was carried out on the CT and PET images for analysis. The centre coordinates of the VOIs extracted from the conventional CT and

PET/CT images were manually set while checking the MPR images of the CT images.

First, the trans-axial image with the largest nodule area was located, and its centre coordinates were specified manually. Then, the longest diameter in the image was set as diameter, D_{xy} , in the x-y direction (trans-axial plane) of the nodule. Subsequently, while changing the slice position in the direction of the body axis, the range of the slice in which the nodule was present was obtained and set as D_z . The VOI was extracted using the number of pixels on three sides, namely $2D_{xy}$, $2D_{xy}$, and $2D_z$, from the original image.

Extraction of characteristic features Checkpoints in malignancy diagnosis

Table 1 shows the checkpoints for distinguishing between benign and malignant nodules (11, 12). The physicians created this scheme by referring to the pixel values, such as the uptake value of the PET images and CT values. Furthermore, consideration was given to such factors as nodule components (e.g., ground glass opacity [GGO] or solid), shapes (roundness), clarity of the nodule border, and spiculas. In this study, these points were quantified as characteristic features.

Characteristic features

(i) Pixel intensities of PET and CT images

Many malignant nodules have a high pixel intensity in PET and CT images. Therefore, standardised uptake value (SUV) (13) of early and delayed PET images was defined as $ESUV$ and $DSUV$, respectively. Furthermore, the difference in SUV between the delayed and early phases was defined as ΔSUV . In the measurement of SUV, two methods were introduced, namely SUV_{max} (hottest voxel) and SUV_{peak} (maximum average SUV within

Table 1. Checkpoints of benign and malignant nodules features

	FDG uptake	CT value	Component	Shape	Border	Spicula
Benignity	Low	Low	GGO	Line-like	Clear	Few
Malignancy	High Increase at delayed phase	High	Solid	Ball-like	Unclear	Many

CT: computed tomography, GGO: ground glass opacity, FDG: fluorodeoxyglucose

a 1 cm³ spherical volume). In the CT images, CT value at the centre of the nodule (CT_{centre}) and the maximum CT value inside the nodule (CT_{max}) were calculated.

(ii) Shape

With regard to the shape of the nodules, malignant nodules often have a ball-like shape, while the benign ones have a line-like shape. To evaluate the ball-like and line-like shapes, a method using a Hessian matrix was proposed (14). The Hessian matrix was obtained by taking the second order differential of the three-dimensional image as follows:

$$H = \begin{bmatrix} F_{xx} & F_{xy} & F_{xz} \\ F_{yx} & F_{yy} & F_{yz} \\ F_{zx} & F_{zy} & F_{zz} \end{bmatrix} \quad (1)$$

Then, three eigenvalues (λ_1, e_2, e_3) were obtained from the matrix. Finally, the ball-like and line-like features (L_{mass} and L_{line}) were calculated using the eigenvalues as follows:

$$L_{\text{mass}} = |\lambda_3|/\lambda_1 \quad (2)$$

$$L_{\text{line}} = |\lambda_2|(\lambda_2 - \lambda_3)/|\lambda_1| \quad (3)$$

(iii) Contrast of the nodule border

The border of a malignant nodule is often unclear. Therefore, the contrast of the border was evaluated using the difference between the CT values of the outer and inner borderlines of the nodules. In order to calculate this value, the average CT values at the pixels belonging to the inner edge R1 (CT_{R1}) and the peripheral region R2 (CT_{R2}) were obtained, and the difference between the two values, $|CT_{R1} - CT_{R2}|$, was defined as the contrast, C_b (Figure 2).

To obtain R1 and R2, the image was first binarized, and the contour was extracted by the Sobel operator. The set of pixels on the outline

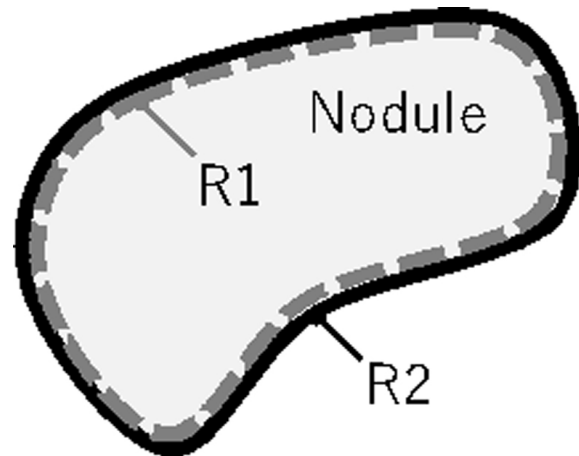


Figure 2. Contrast of nodule border

was defined as R2. Subsequently, the binarized region was shrunk by a morphological operation (erosion) with a structural element having a radius of 1 pixel, and the contour of the reduced region was extracted in the same manner as described above; a set of these pixels was used as R1.

(iv) Spicula

The presence of a spicula around the nodule increases the possibility of the nodule malignancy. In this study, spicula in CT images was detected using Gabor filter (15, 16). The use of Gabor filter facilitates the visualization of line patterns and their orientations (Figure 3b and 3c). Radial line patterns were extracted from the two images (Figure 3d), and the number of radial components and their ratios were calculated as the features of spiculas, SP_1 and SP_2 .

(v) Texture features

The texture pattern of the lung lesion is important for the evaluation of malignancy. Out of the several ways to analyse textures, a method that is proposed by Haralick et al. based on the grey level co-occurrence matrix (GLCM) was employed in the current study (17). The matrix element $P(i,j)$ of GLCM is the set of second order statistical probability values for changes between

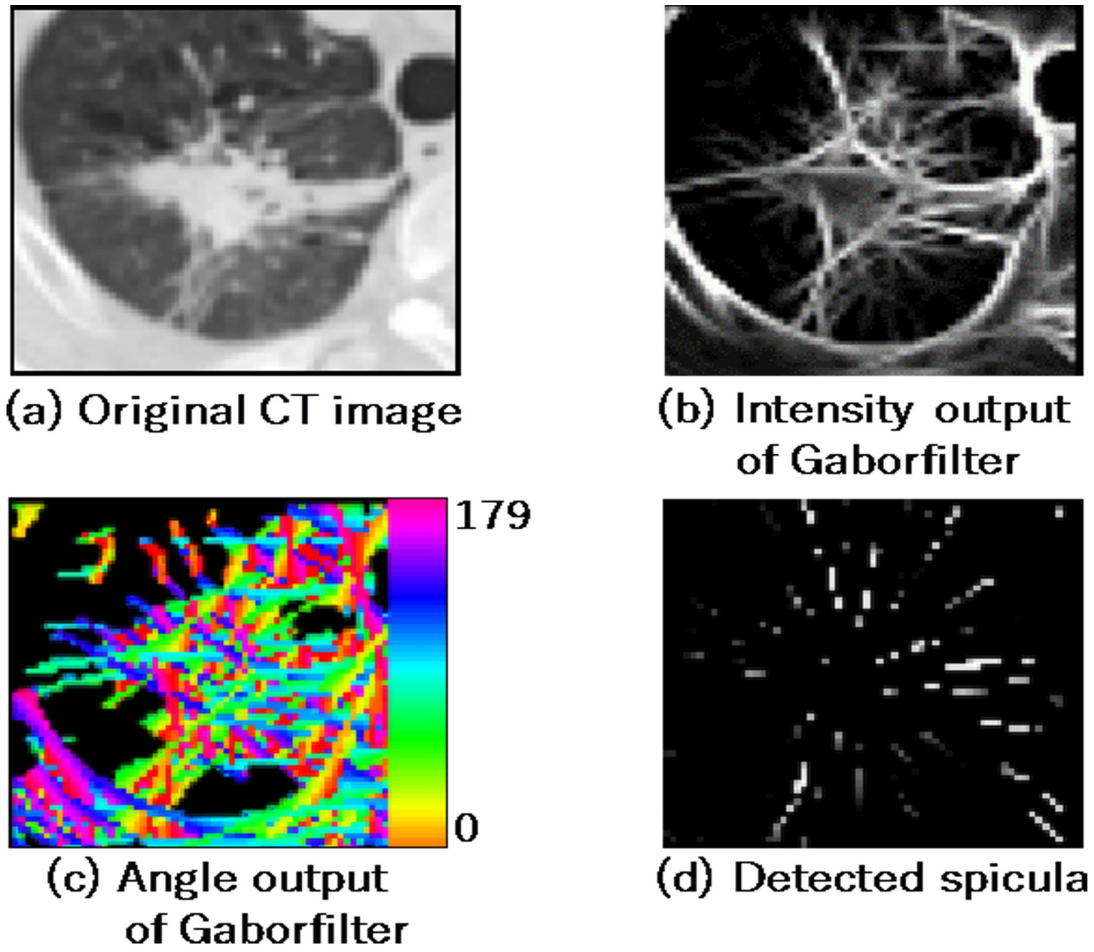


Figure 3. Detection of spicula

grey levels of i and j at a particular distance, d , and angle, θ . In this regard, θ represents the counter-clockwise angle with respect to the X axis.

The GLCM can assess such properties as texture uniformity, directionality, and contrast based on the distribution of the values of the matrix elements. Haralick et al. proposed 14 kinds of characteristic features using GLCM. In our study, it was necessary to limit the number of characteristic features as the number of the analysed cases was small. To obtain the texture features in each direction, the following five types of features were calculated using θ of 0° ($T_{1,0} - T_{5,0}$) and 90° ($T_{1,90} - T_{5,90}$) in the trans-axial plane of the CT images:

Contrast

$$T_1 = \sum_i \sum_j P(i, j) |i - j|^2 \quad (4)$$

Dissimilarity

$$T_2 = \sum_i \sum_j P(i, j) |i - j| \quad (5)$$

Correlation

$$T_3 = \sum_i \sum_j (i - \mu_i)(j - \mu_j)P(i, j) / \sigma_i \sigma_j \quad (6)$$

where,

$$\mu_i = \sum_i \sum_j iP(i, j), \mu_j = \sum_i \sum_j jP(i, j) \quad (7)$$

$$\sigma_i = \sqrt{\sum_i \sum_j P(i, j)(i - \mu_i)^2}, \sigma_j = \sqrt{\sum_i \sum_j P(i, j)(j - \mu_j)^2} \quad (8)$$

Homogeneity

$$T_4 = \sum_i \sum_j P(i, j) / (1 + |i - j|) \quad (9)$$

Energy

$$T_5 = \sqrt{\sum_i \sum_j P(i, j)^2} \quad (10)$$

Classification

Identification of benign or malignant nodules was accomplished by means of the obtained characteristic features. In this study, classification

was performed using the random forest algorithm (18). Random forest is an ensemble learning method for classification and regression that operates by constructing multiple decision trees and outputting the class that is the mode of the classes of the individual trees. Practically, the input for the random forest was the 25 characteristic values, while the output is a judgment result regarding the benignity or malignancy of the nodule. In this study, the maximum number of trees was set at 20.

In order to analyse the distinguishing characteristics of this method, three kinds of classification methods were evaluated. These methods included: 1) classification based on CT images alone, 2) classification based on CT images

and early phase PET images, and 3) classification based on CT images, as well as early and delayed phase PET images.

Results

Characteristic features

In order to confirm whether individual feature values are useful for distinguishing between benign and malignant nodules, the mean, median, and standard deviation of the values in the two groups were calculated.

The effect sizes (19) were also calculated. Furthermore, t-values and p-values were also calculated using the t-test (double-sided test). The results are shown in table 2.

Table 2. Results of basic statistics and t-test

Feature	Benign			Malignant			t-value	p-value	Effect size
	Mean	Median	SD	Mean	Median	SD			
ESUV _{max}	3.31	2.28	2.27	11.66	9.95	6.58	5.242	<0.001	3.685
DSUV _{max}	3.81	2.37	3.08	14.29	11.76	8.3	5.165	<0.001	3.395
ESUV _{peak}	2.31	1.69	1.42	7.66	5.49	4.78	4.684	<0.001	3.764
DSUV _{peak}	2.51	1.56	1.79	9.46	6.83	6.29	4.638	<0.001	3.886
ΔSUV _{max}	0.5	0.21	0.88	2.63	1.96	1.87	4.49	<0.001	2.4
ΔSUV _{peak}	0.2	0.05	0.44	1.8	1.05	2.04	3.348	0.003	3.663
T ₄₋₀	0.0651	0.0602	0.0239	0.045	0.0418	0.0117	3.289	0.003	0.839
T ₄₋₉₀	0.0655	0.0614	0.0234	0.0459	0.0437	0.0114	3.274	0.003	0.834
T ₅₋₀	0.0345	0.0305	0.0192	0.0209	0.0184	0.0084	2.837	0.009	0.708
D _{xy}	17.2	17.5	3.8	22.1	22.8	6.7	2.771	0.01	1.279
T ₅₋₉₀	0.0342	0.0301	0.0179	0.0211	0.0193	0.008	2.92	0.08	0.732
C _b	0.0811	0.0844	0.0305	0.0941	0.0964	0.0137	1.692	0.104	0.425
T ₂₋₀	33.53	33.9	8.27	36.8	37.21	3.09	1.615	0.121	0.394
L _{mass}	56	22.6	104.4	193.2	47.1	365.9	1.574	0.131	1.313
CT _{max}	983.9	661.3	710.7	727.3	535.5	448.8	1.333	0.193	0.361
T ₃₋₀	0.63	0.668	0.154	0.575	0.576	0.104	1.31	0.2	0.362
T ₂₋₉₀	34.36	32.58	8.63	37.02	35.3	4.408	1.198	0.242	0.308
T ₁₋₀	2616.2	2536.8	1096	2885.4	2855.9	395.9	1.009	0.324	0.245
D _z	18.1	16.5	7.9	20.6	19.5	8.1	0.955	0.346	0.312
CT _{centre}	0.3	19.9	151	26.1	35.9	34.1	0.729	0.475	0.171
T ₃₋₉₀	0.612	0.674	0.187	0.58	0.604	0.093	0.671	0.508	0.171
T ₁₋₉₀	2734.4	2500.2	1205.3	2916.3	2761.6	619.5	0.586	0.563	0.15
SP ₂	0.0195	0.0216	0.0104	0.0147	0.0184	0.0161	0.359	0.722	0.101
SP ₁	73.7	42	103.5	82.7	75.5	52.8	0.338	0.738	0.086
L _{line}	42.5	24.4	57.8	43.4	28.9	33.6	0.061	0.952	0.016

ESUV_{max}: maximum standardised uptake value of early PET images, DSUV_{max}: maximum standardised uptake value of delayed PET images, ΔSUV_{max}: difference between ESUV_{max} and DSUV_{max}, ESUV_{peak}: peak standardised uptake value of early PET images, DSUV_{peak}: peak standardised uptake value of delayed PET images, ΔSUV_{peak}: difference between ESUV_{peak} and DSUV_{peak}, CT_{centre}: CT value at the centre of the nodule, CT_{max}: maximum CT value inside the nodule, D_{xy} and D_z: nodule diameters in X-Y plane (trans-axial) and z-direction, L_{mass}: ball-like feature, L_{line}: line-like feature, C_b: CT_{R1}-CT_{R2}, SP₁ and SP₂: features of spicula, T₁₋₀~T₅₋₀: texture features for horizontal direction, T₁₋₉₀~T₅₋₉₀: texture features for vertical direction

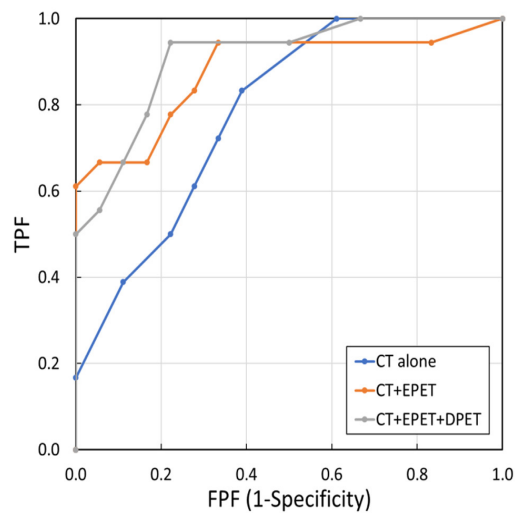


Figure 4. Receiver operating characteristic (ROC) curves of classification

CT: computed tomography, DPET: delayed phase positron emission tomography, EPET: early phase positron emission tomography, FPF: false-positive fraction, ROC: receiver operating characteristics, TPF: true-positive fraction

There was a significant difference between the characteristic features of benign and malignant nodules with regard to SUV. The PET examination resulted in the most significant SUV that was confirmed to be useful for distinguishing between benignity and malignancy. In addition, some texture features showed significant differences between benign and malignant nodules and were effective for distinguishing between these two states. However, difference in features with regard to spicula was of low significance. Furthermore, the average CT_{max} in the benign nodules was higher than that in the malignant nodules.

Results of classification scheme

The nodule classification scheme was evaluated using the receiver operating characteristic (ROC) curve. In the curve, the true positive rate was defined as the ratio of the number of corrected malignant nodules to the total number of malignant nodules. On the other hand, false positive rate was defined as the ratio of the number of misclassified benign nodules to the total number of benign nodules. Furthermore, performance was evaluated by the leave-one-out cross-validation method.

To analyse the distinguishing characteristics of our technique, three methods were evaluated. These methods included: 1) classification based on CT images alone, 2) classification based on CT images and early phase PET images, and 3) classification based on CT images, as well as early and delayed phase PET images. Figure 4 displays

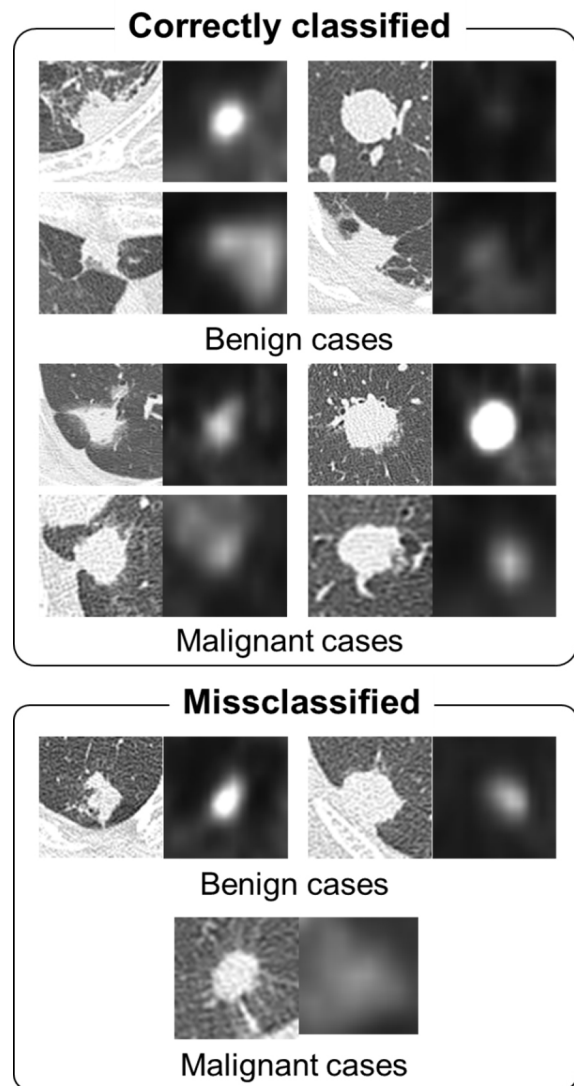


Figure 5. classification results (Images on the left are CT images and images on the right are early phase PET images. The two values under the images are the SUVs of early and delayed PET images).

CT: computed tomography, PET: positron emission tomography, SUV: standardised uptake values

the ROC curves of each of the above methods. The area under the curves (AUC) of methods 1, 2, 3 were 0.730, 0.860, and 0.895, respectively.

Considering the accuracy rate of malignant nodules (0.944), the accuracy rates of benign nodules for methods 1, 2, and 3 were obtained as 0.277, 0.611, and 0.722, respectively. The CT and PET images in trans-axial plane with the accuracy rates of 0.722 and 0.944 for the benign and malignant nodules, respectively are shown in Figure 5. There were significant differences between ROC curves 1 and 2, 2 and 3, as well as 1 and 3 ($P=0.032$, $P=0.104$, and $P=0.021$,

respectively).

Discussion

The findings of the present study revealed a significant difference between the characteristic features of benign and malignant nodules with regard to SUV (Table 2). The SUV obtained by PET examination was confirmed to be useful for distinguishing between benignity and malignancy. However, given the fact that inflammatory diseases can also result in high SUVs, these values should be combined with other features for correct classification. In addition, some texture features were effective for distinguishing between benign and malignant nodules.

Difference in features with regard to spicula was of low significance. This is because many spiculas are observed even in inflammatory diseases. Furthermore, the average CT_{max} in the benign nodules was higher than that in the malignant ones. This is due to the presence of calcification inside the benign nodules. The overlap of the characteristic features between two classes makes it difficult to distinguish between them using a single feature. Therefore, the integration of multiple characteristic features using the classifier would be more effective.

In the ROC curves (Figure 4), the proposed method based on the combined use of CT and two-phase PET images showed an incorrect detection rate of 0.278 for the benign nodules (accuracy rate: 0.722), whereas the accuracy rate of detecting a malignant nodule was 0.944. Target nodules in this study were “difficult nodules” to differentiate based on CT and PET/CT images. Most of the benign cases were not confirmed in follow-up examinations; however, they were confirmed after biopsy. Our results indicated that biopsy examination with its accompanying physical hardship to the patient, especially in benign cases, could be reduced by 72.2%.

As shown in Figure 4, the use of PET plus CT images improves the AUC of the ROC curves, compared to the employment of only CT images. This denotes the effectiveness of using both anatomical and functional information together. Based on the ROC curves, there was a significant difference between the analysis using CT alone and the one using both CT and PET images. In terms of the analysis that was based on PET plus CT images, no significant difference was found between the analysis made based on early phase alone and the one conducted using the early and delayed phases together. However, false-positive

rate at the high true positive rate is very important in malignancy analysis. Based on the above-mentioned results, analysis using two-phase images was still advantageous.

Even with the use of this method, there were some nodules that were misclassified. One of the reasons for this is that the CT and PET images did not show the characteristics that were peculiar to benign or malignant lesions. The future challenge is to introduce the novel characteristic features that reflect benignity and malignancy. Moreover, the further development of this method requires the improvement of the classifiers' performance. Therefore, it is necessary to compare the performance of this method with classifiers, such as linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), support vector machine (SVM), and artificial neural network (ANN).

There are many reports indicating that machine-learning performances are remarkably improved by using deep learning (20, 21). In earlier studies, we applied the deep learning technique for automated nodule detection and classification of lung cancer types (22, 23). In the future, we intend to apply the deep learning technique for the classification of pulmonary nodules.

Conclusion

In this study, we have developed a machine learning-based analysis of pulmonary nodules using early and delayed phase PET and conventional CT images in patients undergoing biopsy. As a result, 94.4% of the malignant nodules were identified correctly assuming that 72.2% of the benign nodules were judged correctly. These results indicate that the proposed method may be useful to improve the accuracy of malignancy analysis.

Acknowledgements

This research was supported by a Grant-in-Aid for Scientific Research on Innovative Areas (Multidisciplinary Computational Anatomy, No. 26108005) and a Grant-in-Aid for Scientific Research (No. 17K09070), MEXT, Japan.

References

1. American cancer society, cancer facts and figures 2015. Available at: URL: <https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2015/cancer-facts-and-figures-2015.pdf>; Accessed September 14, 2018.

2. Sone S, Takashima S, Li F, Yang Z, Honda T, Maruyama Y, et al. Mass screening for lung cancer with mobile spiral computed tomography scanner. *Lancet*. 1998;351(9111):1242-5.
3. National Lung Screening Trial Research Team, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med*. 2011;365(5):395-409.
4. Gould MK, Maclean CC, Kuschner WG, Rydzak CE, Owens DK. Accuracy of positron emission tomography for diagnosis of pulmonary nodules and mass lesions: a meta-analysis. *JAMA*. 2001;285(7):914-24.
5. Armato SG 3rd, Altman MB, Wilkie J, Sone S, Li F, Doi K, et al. Automated lung nodule classification following automated nodule detection on CT: a serial approach. *Med Phys*. 2003;30(6):1188-97.
6. Way TW, Hadjiiski LM, Sahiner B, Chan HP, Cascade PN, Kazerooni EA, et al. Computer-aided diagnosis of pulmonary nodules on CT scans: segmentation and classification using 3D active contours. *Med Phys*. 2006;33(7):2323-37.
7. Zhang F, Song, Y, Cai W, Lee MZ, Zhou Y, Huang H, et al. Lung nodule classification with multilevel patch-based context analysis. *IEEE Trans Biomed Eng*. 2014;61(4):1155-66.
8. Madero Orozco H, Vergara Villegas OO, Cruz Sánchez VG, Ochoa Domínguez Hde J, Nandayapa Alfaro Mde J. Automated system for lung nodules classification based on wavelet feature descriptor and support vector machine. *Biomed Eng Online*. 2015;14:9.
9. Shen W, Zhou M, Yang F, Yu D, Dong D, Yang C, et al. Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification. *Pattern Recognition*. 2017;61:663-73.
10. Nie Y, Li Q, Li F, Pu Y, Appelbaum D, Doi K. Integrating PET and CT Information to improve diagnostic accuracy for lung nodules: a semiautomatic computer-aided method. *J Nucl Med*. 2006;47(7):1075-80.
11. MacMahon H, Naidich DP, Goo JM, Lee KS, Leung ANC, Mayo JR, et al. Guidelines for management of incidental pulmonary nodules detected on CT images: from the Fleischner Society 2017. *Radiology*. 2017;284(1):228-43.
12. Sim YT, Poon FW. Imaging of solitary pulmonary nodule-a clinical review. *Quant Imaging Med Surg*. 2013;3(6):316-26.
13. Keyes JW Jr. SUV: standard uptake or silly useless value? *J Nucl Med*. 1995;36(10):1836-9.
14. Li Q, Sone S, Doi K. Selective enhancement filters for nodules, vessels, and airway walls in two- and three-dimensional CT scans. *Med Phys*. 2003;30(8):2040-51.
15. Rangayyan RM, Ayres FJ. Gabor filter and phase portraits for the detection of architectural distortion in mammograms. *Med Biol Eng Comput*. 2006;44(10):883-94.
16. Yoshikawa R, Teramoto A, Matsubara T, Fujita H. Automated detection of architectural distortion using improved adaptive Gabor filter. *International Workshop on Digital Mammography*, Springer, Cham; 2014 Jun 29. P. 606-11.
17. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cybernetics*. 1973;3(6):610-21.
18. Breiman L. Random forests. *Machine Learn*. 2001;45(1):5-32.
19. Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale: Lawrence Erlbaum Associates; 1988.
20. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Adv Neural Informat Proc Syst*. 2012;25(2):1106-14.
21. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436-44.
22. Teramoto A, Fujita H, Yamamuro O, Tamaki T. Automated detection of pulmonary nodules in PET/CT images: ensemble false-positive reduction using a convolutional neural network technique. *Med Phys*. 2016;43(6):2821-7.
23. Teramoto A, Tsukamoto T, Kiriya Y, Fujita H. Automated classification of lung cancer types from cytological images using deep convolutional neural networks. *Biomed Res Int*. 2017;2017:4067832.