# Impact of pixel-based machine-learning techniques on automated frameworks for delineation of gross tumor volume regions for stereotactic body radiation therapy

河田，康雄

Original paper

# Impact of pixel-based machine-learning techniques on automated frameworks for delineation of gross tumor volume regions for stereotactic body radiation therapy

Yasuo Kawata [a], Hidetaka Arimura [b,*], Koujirou Ikushima [a], Ze Jin [a], Kento Morita [c], Chiaki Tokunaga [d], Hidetake Yabu-uchi [b], Yoshiyuki Shioyama [e], Tomonari Sasaki [b], Hiroshi Honda [b], Masayuki Sasaki [b]

[a] Graduate School of Medical Sciences, Kyushu University, 3-1-1, Maidashi, Higashi-ku, Fukuoka 812-8582, Japan
[b] Faculty of Medical Sciences, Kyushu University, 3-1-1, Maidashi, Higashi-ku, Fukuoka 812-8582, Japan
[c] Department of Health Sciences, School of Medicine, Kyushu University, 3-1-1, Maidashi, Higashi-ku, Fukuoka 812-8582, Japan
[d] Department of Medical Technology, Kyushu University Hospital, 3-1-1, Maidashi, Higashi-ku, Fukuoka 812-8582, Japan
[e] Saga Heavy Ion Medical Accelerator in Tosu, 415, Harakoga-cho, Tosu 841-0071, Japan

ABSTRACT

The aim of this study was to investigate the impact of pixel-based machine learning (ML) techniques, i.e., fuzzy-c-means clustering method (FCM), and the artificial neural network (ANN) and support vector machine (SVM), on an automated framework for delineation of gross tumor volume (GTV) regions of lung cancer for stereotactic body radiation therapy. The morphological and metabolic features for GTV regions, which were determined based on the knowledge of radiation oncologists, were fed on a pixel-by-pixel basis into the respective FCM, ANN, and SVM ML techniques. Then, the ML techniques were incorporated into the automated delineation framework of GTVs followed by an optimum contour selection (OCS) method, which we proposed in a previous study. The three-ML-based frameworks were evaluated for 16 lung cancer cases (six solid, four ground glass opacity (GGO), six part-solid GGO) with the datasets of planning computed tomography (CT) and ^{18}F-fluorodeoxyglucose (FDG) positron emission tomography (PET)/CT images using the three-dimensional Dice similarity coefficient (DSC). DSC denotes the degree of region similarity between the GTVs contoured by radiation oncologists and those estimated using the automated framework. The FCM-based framework achieved the highest DSCs of $0.79 \pm 0.06$, whereas DSCs of the ANN-based and SVM-based frameworks were $0.76 \pm 0.14$ and $0.73 \pm 0.14$, respectively. The FCM-based framework provided the highest segmentation accuracy and precision without a learning process (lowest calculation cost). Therefore, the FCM-based framework can be useful for delineation of tumor regions in practical treatment planning.

© 2017 Associazione Italiana di Fisica Medica. Published by Elsevier Ltd. All rights reserved.

## 1. Introduction

Stereotactic body radiotherapy (SBRT) is a promising technique for treatment of lung cancer, which was the leading cause of cancer death in Japan, as of 2011 [1]. The three-year overall survival rate after SBRT was 56.6%, which did not differ significantly between SBRT and surgery in patients with operable Stage I non-small cell lung cancer (NSCLC) [2]. However, SBRT requires fewer uncertainties on many factors, especially the target delineation, because higher doses (e.g., 12 Gy) per fraction are administrated to small

targets ($\leq 5$ cm) rather than to the surrounding normal tissues. The uncertainties on the clinical target volume (CTV) and planning target volume (PTV) in the radiation treatment planning engenders high impacts on the precisions of entire radiation treatment courses, including planning and patient positioning [3]. The CTV and PTV have been determined based on gross tumor volume (GTV). However, the GTV regions are manually delineated on treatment planning computed tomography (CT) images by treatment planners (e.g., radiation oncologists) in current radiation therapy. Consequently, the manual delineation based on the experiences and expertise of treatment planners can cause intra- and inter-observer variations in GTV contours in clinical practice [4,5]. This issue prevents us from performing the multi-institutional planning comparison [6,7]. Therefore, the automated frameworks are substantially necessary to unify GTV contours in the comparison of

multi-institutions. Furthermore, it is time-consuming for treatment planners to manually delineate the GTV contours.

Therefore, several computational approaches have been reported for determining the GTV contours [8–21]. There are two major types of conventional approaches with and without machine learning (ML) techniques: unsupervised and supervised learning. Several researchers have shown the potential of ML techniques to estimate initial regions of GTV contours, i.e., the fuzzy locally adaptive Bayesian approach [15], the Bayesian network model and support vector machine (SVM) model [16], the fuzzy c-means algorithm [17], artificial neural network (ANN)-based thresholding segmentation [19], and SVM-based delineation using an optimum contour selection (OCS) method [21]. One of the drawbacks of supervised ML, however, is that the learning step requires certain knowledge (teaching data, such as reference GTV regions) and longer computational times compared with unsupervised learning. On the other hand, a disadvantage of unsupervised ML is the lack of knowledge. Nevertheless, no studies have been conducted that compared performances of unsupervised and supervised ML techniques for delineation of GTV contours of lung cancer. Furthermore, since there are several lung tumor types (solid, ground glass opacity (GGO), and part-solid GGO), the properties of the ML techniques for the three types of tumor should be explored. A nodule obscured the entire lung parenchyma within them was defined as solid. A nodule that did not completely obscure the entire lung parenchyma within them and there was no solid component was defined as GGO. A non-solid nodule developing an internal solid component was defined as part-solid GGO. The classification of the lung cancer was described in [22–24].

Therefore, the aim of this study was to investigate the impact of ML techniques on an automated framework for delineation of GTV regions of lung cancer for SBRT. The fuzzy-c-means clustering method (FCM) was selected as an unsupervised ML technique, and ANN and SVM were selected as the supervised ML techniques. The three ML techniques were incorporated as pixel-based ML classifiers into an automated delineation framework of GTVs followed by the OCS method, which we proposed in a previous study [21].

## 2. Materials and methods

### 2.1. Clinical cases

This study was performed with the approval of the Institutional Review Board of our university hospital. We selected for this study sixteen lung cancer patients (mean age: 76 years; range: 65–86 years; eight females; eight males; GTV mean effective diameter: 21 mm), who received SBRT at the university hospital. The patient datasets consisted of the planning CT and positron emission tomography (PET)/CT images. The planning CT images were acquired by a four-slice CT scanner (Mx 8000; Philips, Amsterdam, The Netherlands) with a $512 \times 512$ matrix size, a nominal in-plane pixel size of 0.98 mm, and a nominal slice thickness of 2.0 mm. The patients were scanned with free-breathing due to small tumor motion in upper lobe or patient's performance status. The nominal in-plane pixel size in mm is technically determined based on (field of view (FOV) in mm along the anterior-posterior or left-right direction)/(number of pixels). Similarly, the nominal slice thickness is obtained based on (FOV in mm along the superior-inferior direction)/(number of slices). The planning CT images with the anisotropic voxel were transformed into images with matrix sizes of $512 \times 512 \times 266$–389 and a nominal 0.98 mm isotropic voxel using a cubic interpolation method.

Table 1 shows the sixteen lung cancer patients characteristics. There were six solid, four ground glass opacity (GGO), and six

part-solid GGO types of lung cancer. The lung window level and width were set at −600 and 1500 Hounsfield units [HU], respectively, and the mediastinal window level and width were set at 50 and 400 HU, respectively, for determining the tumor types on the planning CT image by a radiologist (H.Y.).

The integrated PET/CT scanner (Discovery STE, General Electric Medical Systems, Milwaukee, WI; and Biograph mCT, Siemens, Berlin, Germany) was used for scanning each patient with their arms down while freely breathing 60 min after FDG injection. The PET images could help treatment planners delineate GTV regions with relevant metabolic information. By using a three-dimensional (3D) ordered subset-expectation maximization (OS-EM) algorithm (VUE Point Plus, GE Healthcare; and ultraHD, Siemens), the PET images were acquired in the 3D mode and reconstructed with correction for attenuation, scatter, decay, random, and dead time. The number of iterations was set to two, the number of subsets was set to 21, and the full width at half maximum of a Gaussian filter was set as 6.0 mm, respectively. The PET images were acquired with 128 × 128 pixels with a nominal in-plane pixel size of 5.5 mm, and a nominal slice thickness of 3.3 mm, or 256 × 256 pixels with a nominal in-plane pixel size of 3.2 mm and a nominal slice thickness of 3.0 mm. In addition, the diagnostic CT images consisted of a $512 \times 512$ matrix, a nominal in-plane pixel size of 0.98 mm, and a nominal slice thickness of 3.3 or 3.0 mm. The diagnostic CT images were acquired by using a 16-slice CT scanner in the PET/CT system.

By using a linear and cubic interpolation method, the PET and diagnostic CT images were converted into isotropic images with a nominal isotropic voxel of 0.98 mm, respectively. The matrix sizes for the PET images were $716 \times 716 \times 793$–920, or $834 \times 834 \times 811$–992. The matrix sizes for the diagnostic CT images were $512 \times 512 \times 811$–992 with a nominal isotropic voxel of 0.98 mm. The linear interpolation method was applied to the PET images to avoid the problem that the edge portions in the PET images were likely to be enhanced by a cubic interpolation method. The isotropic diagnostic CT images were smaller than the isotropic PET images. However, for obtaining the image features at the same voxels in both images, the matrix sizes of both images should have been the same size. Therefore, the isotropic diagnostic CT images were situated on the centers of the isotropic PET images. Then, the isotropic PET images were cropped to the same sizes as the isotropic diagnostic CT images.

The standardized uptake values (SUVs) were used as metabolic features of the PET images to be input to the FCM, ANN and SVM. The SUV is the ratio of the radioactivity concentration of the tissue at one time point to the injected dose of radioactivity concentration at that time point divided by the body weight. The SUV is calculated by the following equation:

$$SUV = \frac{C}{D/W}, \tag{1}$$

where $C$ is the radioactivity concentration in kBq/mL obtained from the pixel value in the PET image multiplied by a cross calibration factor. In addition, $D$ is the injected dose of $^{18}$FDG administered in megabecquerels (MBq; decay corrected), and $W$ is the patient body weight in kilograms.

Two experienced radiation oncologists approved radiation treatment plans by using a commercially available radiation treatment planning (RTP) system (Eclipse v. 6.5 and 8.1, Varian Medical Systems Inc., Palo Alto, USA). Two experienced radiation oncologists (Y.S. and T.S.) delineated the GTV contours based on a consensus between them by using the RTP system on the planning CT images by referencing fusion images of the PET and diagnostic CT images.

**Table 1**
Summary of patient characteristics.

| Case No. | Gender | Age (years) | GTV size[a] (mm) | Tumor location | SUV$_{max}$[b] | Tumor type | Tumor CT imaging characteristics |
|---|---|---|---|---|---|---|---|
| 1 | F[c] | 71 | 18 | RUL[e] | 8.4 | Solid | Homogeneous Irregular |
| 2 | F | 67 | 24 | RUL | 12 | Solid | Homogeneous Irregular Vascular |
| 3 | M[d] | 65 | 25 | RUL | 6.8 | Solid | Inhomogeneous Irregular |
| 4 | M | 75 | 20 | LUL[f] | 8.7 | Solid | Inhomogeneous Irregular Adjacent Pleural |
| 5 | M | 86 | 29 | LUL | 9.7 | Solid | Cavity Irregular |
| 6 | F | 81 | 26 | RUL | 4.4 | Solid | Homogeneous Irregular Pleural Indentation |
| 7 | M | 76 | 18 | LUL | 1.7 | GGO | Irregular Pleural Indentation Vascular |
| 8 | F | 74 | 16 | RLL[g] | 1.3 | GGO | Smooth |
| 9 | M | 81 | 19 | LUL | 2.6 | GGO | Smooth |
| 10 | F | 79 | 21 | RUL | 1.5 | GGO | Irregular |
| 11 | M | 77 | 21 | LLL[h] | 6.6 | Part solid GGO | Inhomogeneous Irregular Cavity |
| 12 | F | 85 | 14 | RUL | 1.7 | Part solid GGO | Irregular |
| 13 | M | 65 | 18 | RUL | 1.3 | Part solid GGO | Smooth Inhomogeneous |
| 14 | F | 84 | 16 | LLL | 1.4 | Part solid GGO | Irregular Pleural Indentation Vascular |
| 15 | F | 78 | 19 | RUL | 1.4 | Part solid GGO | Smooth Pleural retraction |
| 16 | M | 75 | 25 | RUL | 2.2 | Part solid GGO | Smooth |

[a] Effective diameter.
[b] Maximum standardized uptake value.
[c] Female.
[d] Male.
[e] Right upper lobe.
[f] Left upper lobe.
[g] Right lower lobe.
[h] Left lower lobe.

## 2.2. Automated framework for delineation of GTV regions

Fig. 1 shows an overall scheme of the automated framework for delineation of GTV regions. The basic concept of this study was to extract GTV regions from the vol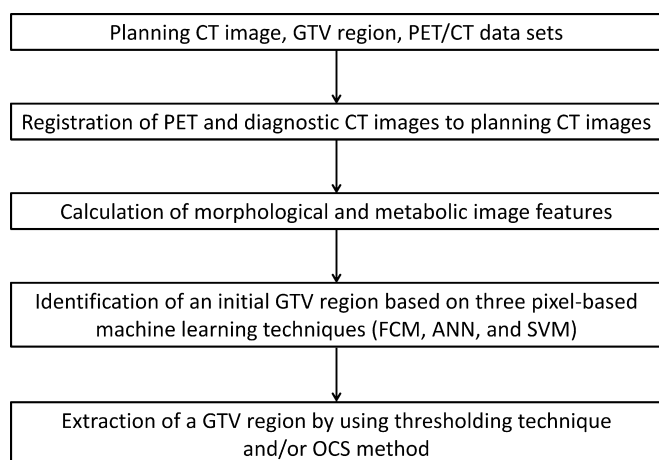ume of interest (VOI) based on the knowledge and experience of radiation oncologists by using the pixel-based ML techniques with six image features. Therefore, the six image features, i.e., the voxel values and the magnitudes of the image gradient vectors on the planning CT and PET/CT image datasets, were used as the input image features for the pixel-based ML techniques.

### 2.2.1. Registration of PET and diagnostic CT images to planning CT images

The PET and diagnostic CT images were registered to the planning CT images by the lung region centroid matching method described in [21] for utilization of the metabolic and morphological information of the PET, diagnostic CT, and planning CT images. The registration of the PET and diagnostic CT images to the planning CT images was performed based on centroid matching of lung regions with bronchi. These lung regions with bronchi were extracted from diagnostic CT and planning CT images. The algorithm for extraction of lung regions with bronchi was described in [21].

### 2.2.2. Calculation of morphological and metabolic image features

The voxel values and the magnitudes of image gradient vectors calculated from PET, diagnostic CT, and planning CT images were used as the image features. The image gradient magnitude at a voxel was derived from the following first-order polynomial within a $5 \times 5 \times 5$ voxel region obtained by use of the least squares method [25]:



Fig. 1. An overall scheme of the automated framework for delineation of GTV regions.

```
Planning CT image, GTV region, PET/CT data sets
        ↓
Registration of PET and diagnostic CT images to planning CT images
        ↓
Calculation of morphological and metabolic image features
        ↓
Identification of an initial GTV region based on three pixel-based machine learning techniques (FCM, ANN, and SVM)
        ↓
Extraction of a GTV region by using thresholding technique and/or OCS method
```

$$f(x, y, z) = ix + jy + kz + l, \tag{2}$$

where $x$, $y$ and $z$ are the coordinates of one of the three types of images, $f(x, y, z)$ is the first-order polynomial, and $i, j, k$ and $l$ are constants. Therefore, gradient magnitude $G$ was calculated by

$$G = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + \left(\frac{\partial f}{\partial z}\right)^2} = \sqrt{i^2 + j^2 + k^2}. \tag{3}$$

### 2.2.3. Identification of initial GTV region based on three pixel-based ML techniques (FCM, ANN, and SVM)

The initial GTV regions were identified by classifying pixels within VOI based on three pixel-based ML techniques, i.e., FCM, ANN, and SVM, with the image features from VOI based on the knowledge and experience of the radiation oncologists. The image features used as input image features for the pixel-based ML techniques were the voxel values and the magnitudes of the image gradient vectors on the planning CT and PET/CT image datasets.

### 2.2.4. Extraction of a GTV region by using thresholding and/or the OCS method

Membership map values of tumor and normal tissue were calculated by applying FCM to each voxel. The membership map values of tumors were used for initial GTVs because membership map values indicate the degree of GTVs.

On the other hand, initial GTV regions and other regions were classified by applying ANN and SVM to each voxel. The initial GTV regions obtained by FCM, ANN, and SVM were smoothed by applying a Gaussian filter with a standard deviation (SD) of three pixels. Finally, the GTV regions were determined by thresholding the smoothed initial GTV regions. In addition, the GTV regions of a solid type obtained by ANN and SVM were determined by thresholding and applying the OCS method to the smoothed initial GTV regions. The post-processing method for the initial GTV regions extracted by ANN and SVM was the same as the method developed by Ikushima et al. [21].

### 2.2.5. Fuzzy c-means clustering method

The FCM method is an unsupervised ML technique that is used to cluster the dataset. In this study, the initial GTV regions were extracted by using the FCM method. A class membership value was assigned to each voxel by the FCM method, depending on the similarity of the pixel to a particular class relative to all other classes [26–29]. The FCM algorithm is shown in Fig. 2. The FCM method was performed by minimizing the following objective

function with respect to the membership function $u_{ik}$ at voxel $k$ and cluster $i$ and the centroid vector of a cluster $\mathbf{v}_i$:

$$J_{FCM} = \sum_{k=1}^{n} \sum_{i=1}^{c} u_{ik}^m \|\mathbf{x}_k - \mathbf{v}_i\|^2, \tag{4}$$

where $\mathbf{x}_k$ is the image feature vector at voxel $k$, $c$ is the number of clusters, $n$ is the number of voxels, and $m$ is a constant that controls the fuzziness of the resulting partition. Additionally, $m$ is set to two in this paper. The membership functions are constrained to be positive and to satisfy the following equation:

$$\sum_{i=1}^{c} u_{ik} = 1. \tag{5}$$

Minimization of $J_{FCM}$ is performed by an iterative calculation of $u_{ik}$ and $\mathbf{v}_i$ using the following equation:

$$u_{ik} = \left[ \sum_{j=1}^{c} \left( \frac{\|\mathbf{x}_k - \mathbf{v}_i\|^2}{\|\mathbf{x}_k - \mathbf{v}_j\|^2} \right)^{\frac{1}{m-1}} \right]^{-1}, \tag{6}$$

where

$$\mathbf{v}_i = \frac{\sum_{k=1}^{m} \left( u_{ik}^m \mathbf{x}_k \right)}{\sum_{k=1}^{n} \left( u_{ik}^m \right)}. \tag{7}$$

The objective function is minimized when high membership values are obtained.

### 2.2.6. Artificial neural network

The ANN is a supervised ML technique that can classify the input data to any categories based on training [30]. The ANN is a framework modeled as a biological neural network with an activation function. In the ANN, the activation function is applied to the sum of the input data multiplied by weights. The weights in the ANN are determined by using a back-propagation of errors between predicted outputs and teacher signals at a learning step. The activation function is expressed by

$$f(x) = 1 - \frac{2}{\exp(\alpha x) + 1}, \tag{8}$$

where $\alpha$ determines the slope of the activation function. In this paper, $\alpha$ was set to 0.5. In addition, the learning rate used for back-propagation was set to 0.001, and the number of nodes in a middle layer was set to two.

### 2.2.7. Support vector machine

The SVM is a supervised ML technique that can classify the input dataset into two classes based on a discriminant function [31]. The discriminant function is constructed in a linearly separable space by using a non-linear kernel function with training datasets. In the separable space, a margin is calculated to maximize the distance between the support vector and hyperplane. The margin is calculated by a Lagrangian function based on the training dataset. The following equation is constructed by SVM as the discriminant function:

$$f(\mathbf{x}) = \sum_{i=1}^{N} y_i \alpha_i K(\mathbf{x}, \mathbf{x}_i) + b, \tag{9}$$

where $\mathbf{x}_i$ ($i = 1, \cdots, N$, $N$: number of support vectors) is the support vector, $b$ and $\alpha_i$ are parameters for determining the discriminant function, and $K(\mathbf{x}, \mathbf{x}_i)$ is the non-linear kernel function for mapping a linearly non-separable dataset to a linearly separable dataset. In this study, the Gaussian kernel, i.e., $\exp(-\gamma \|\mathbf{x} - \mathbf{y}\|^2)$, was used as the non-linear kernel function of the SVM. The value $\gamma$ was set to 0.001 for solid and GGO, and to 0.00001 for part-solid GGO. In
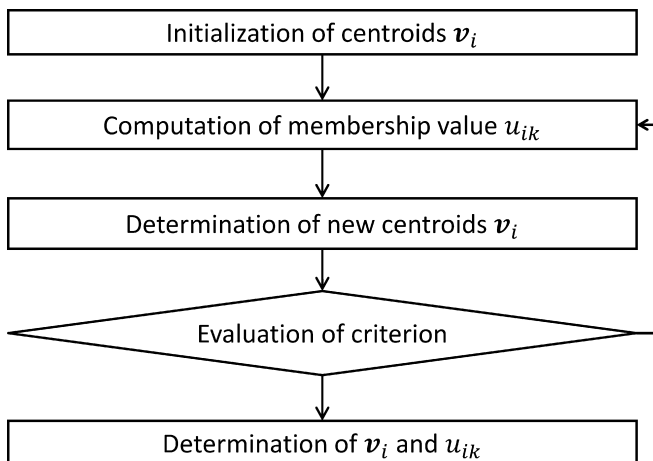


Fig. 2. An algorithm of FCM.

addition, a parameter for the soft margin of SVM was set at ten for solid and part-solid GGO, and to 0.01 for GGO. These SVM parameters were optimized by using image features of the training dataset for each tumor type.

## 2.3. Evaluation of GTV extracted by the ML techniques

The GTV regions extracted by the ML techniques were evaluated by using 3D DSC [32]. The DSC indicates the degree of region similarity between the GTV ground truth region and the GTV region estimated using the ML techniques. The DSC is calculated according to the following equation:

$$DSC = \frac{2n(T \cap D)}{n(T) + n(D)}, \tag{10}$$

where $T$ is the GTV ground truth region determined by two radiation oncologists (Y.S., T.S.), $D$ is the GTV region estimated using the ML techniques, $n(T)$ is the number of pixels in a region $T$, $n$ ($D$) is the number of pixels in a region $D$, and $n(T \cap D)$ is the number of logical AND pixels between regions $T$ and $D$. The DSC ranges from zero (no overlap between $T$ and $D$) to one ($T$ and $D$ are identical).

The GTV regions obtained from the Digital Imaging and Communications in Medicine (DICOM) for radiation therapy (DICOM−RT) files were used as ground truths.

Planning CT ($0.98 \times 0.98 \times 2.0$ mm$^3$), PET ($5.5 \times 5.5 \times 3.3$ or $3.2 \times 3.2 \times 3.0$ mm$^3$), and diagnostic CT images ($0.98 \times 0.98 \times 3.3$ or $0.98 \times 0.98 \times 3.0$ mm$^3$) were converted into images with an isotropic voxel of 0.98 mm to easily perform the image processing techniques developed in this study. To avoid the influence of these large upsampling on the evaluation of segmentation accuracy of the ML techniques, the GTV regions estimated by the ML techniques were downsampled into the original anisotropic voxel using a nearest neighbor method to calculate the DSCs with the ground truths on the planning CT images with the original anisotropic voxel.
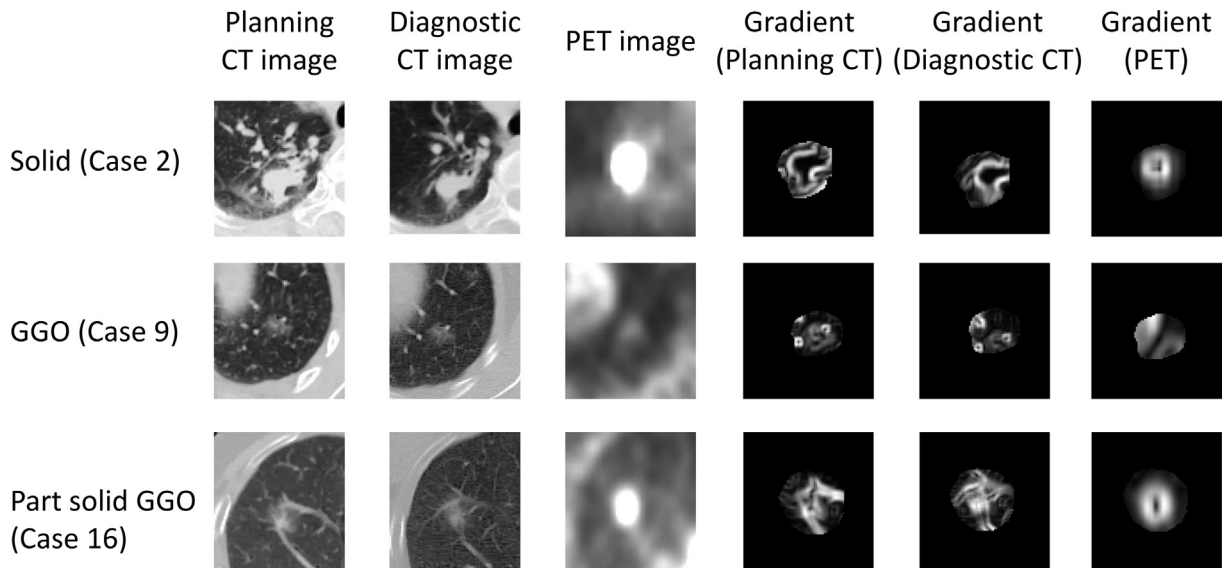


**Fig. 3.** Original images and gradient vector magnitude images obtained from the planning CT and PET/CT image dataset that were used as image features for the three pixel-based ML techniques.
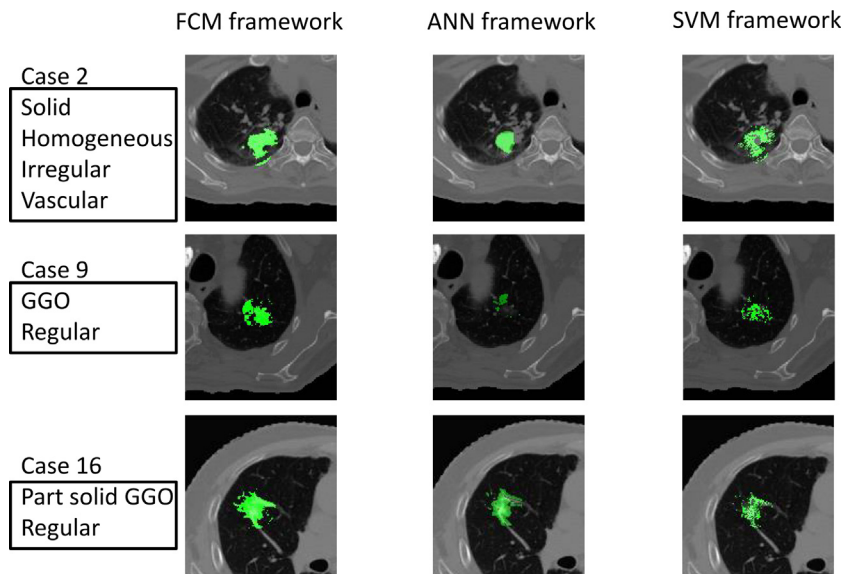


**Fig. 4.** Illustrations of the output images obtained by the three pixel-based ML techniques for solid, GGO, and part-solid GGO lung tumor types.

## 3. Results

Fig. 3 shows the image features, which were given the three ML techniques. The solid (Case 2) and part solid GGO (Case 16) tumors have visible but blurry boundaries in the gradient images, and show bright regions in PET images. In the contrast, the GGO (Case 9) tumor is depicted with subtle edges and dark appearance. Fig. 4 shows the output images obtained by the three pixel-based ML techniques for three lung tumor types. The FCM framework shows the stronger outputs for three types of tumors, but the other frameworks produced weaker outputs for GGO and part solid GGO tumors. Fig. 5 shows extracted GTV regions by the three ML techniques with the planning CT images for the three lung tumor types. For the solid tumor (Case 2), the GTV region of a homogeneous and irregular tumor with vascularization was well extracted

by using the three ML techniques. For the GGO tumor (Case 9), the DSC obtained by the ANN-based framework was lower than the other frameworks. The GTV region of the part solid GGO tumor (Case 16) was well segmented by the FCM-based framework. The ANN- and SVM-based frameworks were not able to segment the blurry region on the part solid GGO tumor for Case 16.

Table 2 shows a comparison of the DSCs among the FCM-, ANN-, and SVM-based frameworks for the sixteen cases. The average 3D DSC between the GTV ground truths contoured by the radiation oncologists and the GTV regions obtained by using the FCM-based framework was $0.79 \pm 0.06$. On the other hand, the average 3D DSC was $0.76 \pm 0.14$ obtained by the ANN-based framework and $0.73 \pm 0.14$ obtained by the SVM-based framework. The average 3D DSC by the FCM-based framework was higher than the average 3D DSCs by SVM-based and ANN-based frameworks
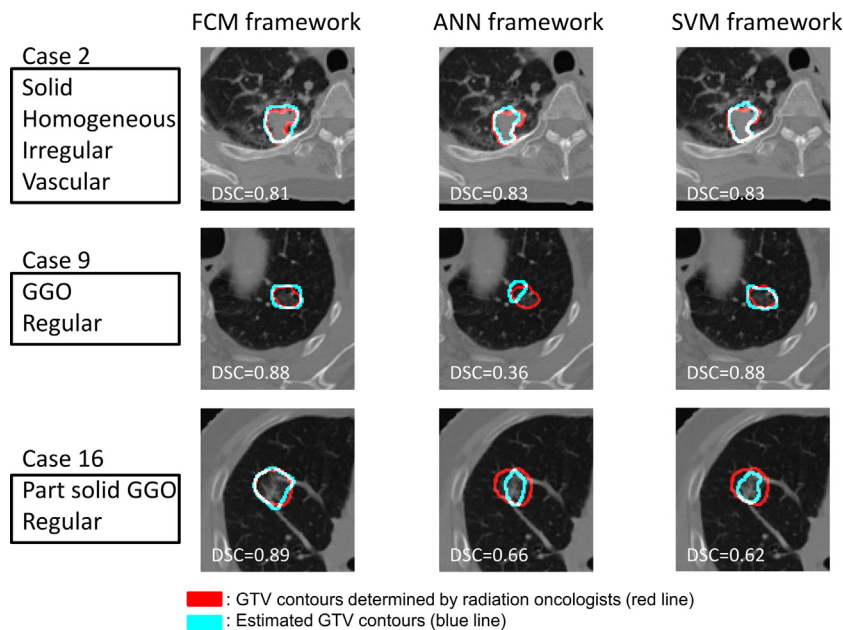


Fig. 5. A comparison of results of the three pixel-based ML framework in terms of tumor CT imaging characteristics.

**Table 2**
Average three-dimensional Dice similarity coefficients (DSCs) obtained by using three pixel-based machine-learning techniques (FCM-based framework, ANN-based framework and SVM-based framework) for 16 cases.

| Case No. | Tumor type | FCM-based framework | ANN-based framework | SVM-based framework |
|---|---|---|---|---|
| 1 | Solid | 0.83 | 0.88 | 0.85 |
| 2 | Solid | 0.82 | 0.80 | 0.84 |
| 3 | Solid | 0.81 | 0.81 | 0.81 |
| 4 | Solid | 0.82 | 0.86 | 0.90 |
| 5 | Solid | 0.90 | 0.82 | 0.87 |
| 6 | Solid | 0.83 | 0.84 | 0.79 |
| 7 | GGO | 0.80 | 0.79 | 0.83 |
| 8 | GGO | 0.84 | 0.75 | 0.77 |
| 9 | GGO | 0.70 | 0.30 | 0.70 |
| 10 | GGO | 0.80 | 0.77 | 0.81 |
| 11 | Part solid GGO | 0.73 | 0.77 | 0.44 |
| 12 | Part solid GGO | 0.74 | 0.84 | 0.66 |
| 13 | Part solid GGO | 0.76 | 0.70 | 0.68 |
| 14 | Part solid GGO | 0.74 | 0.80 | 0.54 |
| 15 | Part solid GGO | 0.71 | 0.76 | 0.72 |
| 16 | Part solid GGO | 0.87 | 0.67 | 0.45 |
| Mean ± SD | | $0.79 \pm 0.06$ | $0.76 \pm 0.14$ | $0.73 \pm 0.14$ |
| P[a] (t-test with FCM) | | – | 0.14 | 0.03 |
| P[b] (F-test with FCM) | | – | 0.002 | $0.94 \times 10^{-3}$ |

[a] The p-value was calculated by t-test with FCM-based framework.
[b] The p-value was calculated by F-test with FCM-based framework.

(p > 0.05). The SD of DSCs obtained by the FCM-based framework was significantly smaller, compared with those obtained by the SVM-based and ANN-based frameworks (p < 0.05).

Table 3 shows a comparison of the DSCs of the FCM-, ANN-, and SVM-based frameworks for the three tumor types. The highest DSCs for solid, GGO, and part-solid GGO types were obtained by the SVM-, FCM-, and FCM-based frameworks, respectively. Therefore, the suitable ML technique differed for each tumor type. The SDs of the 3D DSCs of the GGO and the part-solid GGO tumor types obtained by the FCM-based framework were significantly smaller (p < 0.05) compared with that of the GGO tumor type obtained by the ANN-based framework and that of the part-solid GGO tumor type obtained by the SVM-based framework.

## 4. Discussion

In this study, we strived to develop an automated framework for delineation of GTV regions, and we compared DSCs of the GTV regions estimated by three pixel-based ML techniques, i.e., FCM, ANN and SVM. As a result, the 3D average DSC by using the FCM-based framework was highest for 16 patients. Table 4 shows a comparison of the computational times for the learning step and test step by the pixel-based ML techniques for the sixteen cases. The computational time by each pixel-based ML framework was measured on a computer (CPU: Intel(R) Core(TM) i7-6700 3.4-GHz quad core; 32 GB of RAM). The average computational time of the FCM-based framework was significantly less than those of the SVM-based and ANN-based framework (p < 0.05).

Fig. 6 indicates the relationship between the DSC and the threshold value used for extraction of the GTV regions, for FCM cluster numbers of 2, 3, and 4. The cluster number and threshold value on each tumor type were determined for the highest DSCs.

**Table 3**
Comparisons of three dimensional DSCs (mean ± SD) for three tumor types obtained by three frameworks.

| Tumor type | FCM-based framework | ANN-based framework | SVM-based framework |
|---|---|---|---|
| Solid | 0.83 ± 0.03 | 0.83 ± 0.03 | 0.84 ± 0.04 |
| GGO | 0.79 ± 0.06 | 0.65 ± 0.24 | 0.77 ± 0.06 |
| Part solid GGO | 0.76 ± 0.06 | 0.76 ± 0.07 | 0.58 ± 0.12 |

**Table 4**
Comparisons of computational time by three frameworks.

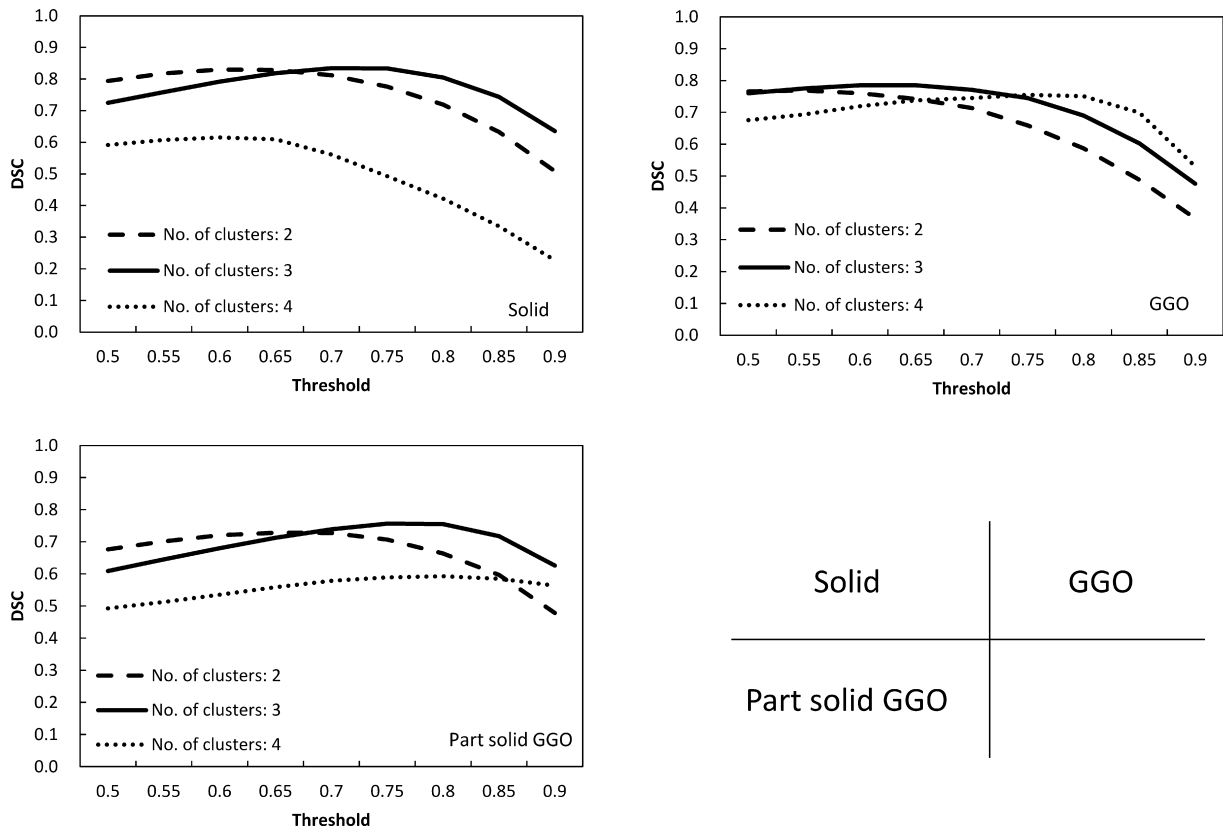| Tumor type | FCM-based framework | | ANN-based framework | | SVM-based framework | |
|---|---|---|---|---|---|---|
| | Learning (sec) | Test (sec) | Learning (sec) | Test (sec) | Learning (sec) | Test (sec) |
| Solid | 0 | 2.3 | 365.6 | 3.4 | 4838.0 | 46.5 |
| GGO | 0 | 1.2 | 98.4 | 1.4 | 30.9 | 7.2 |
| Part solid GGO | 0 | 1.5 | 187.8 | 1.6 | 867.2 | 12.7 |
| Mean (16 cases) | 0 | 1.7 | 232.1 | 2.2 | 2147.2 | 24.0 |
| P value with FCM | – | – | $3.0 \times 10^{-7}$ | 0.02 | $8.0 \times 10^{-4}$ | $1.0 \times 10^{-4}$ |



**Fig. 6.** Relationships between DSC and the threshold value for determining GTV regions of solid, GGO, and part-solid GGO types, for FCM cluster numbers of 2, 3, and 4.

As a result, the cluster number was determined to be three, and the threshold value for the smoothed initial GTVs was determined to be 70% for the solid type, 65% for the GGO type, and 75% for the part-solid GGO type.

The suitable pixel-based ML differed depending on the tumor type. The GTV region of the solid tumor type was well extracted by the SVM-based framework compared to the other pixel-based ML approaches. The GTV regions of the GGO tumor type had the highest DSC when using the FCM-based framework. The part-solid tumor type had the highest DSC, which was extracted by the ANN-based framework. Therefore, as a combined framework, we applied the SVM-based framework to the solid tumor type, the FCM-based framework to the GGO tumor type, and the ANN-based framework to the part-solid tumor type, as shown in Fig. 7. Table 5 shows the results of the FCM-based and the combined frameworks. The average DSCs obtained by the FCM-based framework and the combined approach were $0.79 \pm 0.06$ and $0.80 \pm 0.06$, respectively. There were no significant differences in the mean and SD of the average 3D DSCs. However, the computational time of the FCM-based framework was significantly shorter than that of the combined framework ($p < 0.05$). Therefore, the FCM-based framework can be considered more useful for clinical practice because of its robustness and shorter computational time.

The planning CT images used in this study were acquired under a free-breathing condition due to small tumor motion or patient's performance status. Patients' breathing motions or tumor motion may influence the accuracy at all steps of the treatment process due to image blurring, dose blurring, interplay effects [33,34]. In particular, the image blurring effect, which causes ambiguous outlines of tumors, could affect the segmentation accuracy of the automated frameworks. In fact, the FCM-based framework provided an average DSC of $0.79 \pm 0.06$ for all cases, but $0.80 \pm 0.06$ for upper lobe tumors and $0.77 \pm 0.06$ for lower lobe tumors. Therefore, as a future work, we should

improve the framework to be robust for mitigating the impact of the breathing motion or tumor motion on the segmentation accuracy of the frameworks.

In future work, other image features should be considered to improve the segmentation accuracy. However, it will be important to select the useful image features depending on the tumor type. In addition, the proposed approach should be applied to large databases with various types of tumors to improve the performance. Furthermore, the proposed approach should be evaluated by observer tests to further improve its robustness.
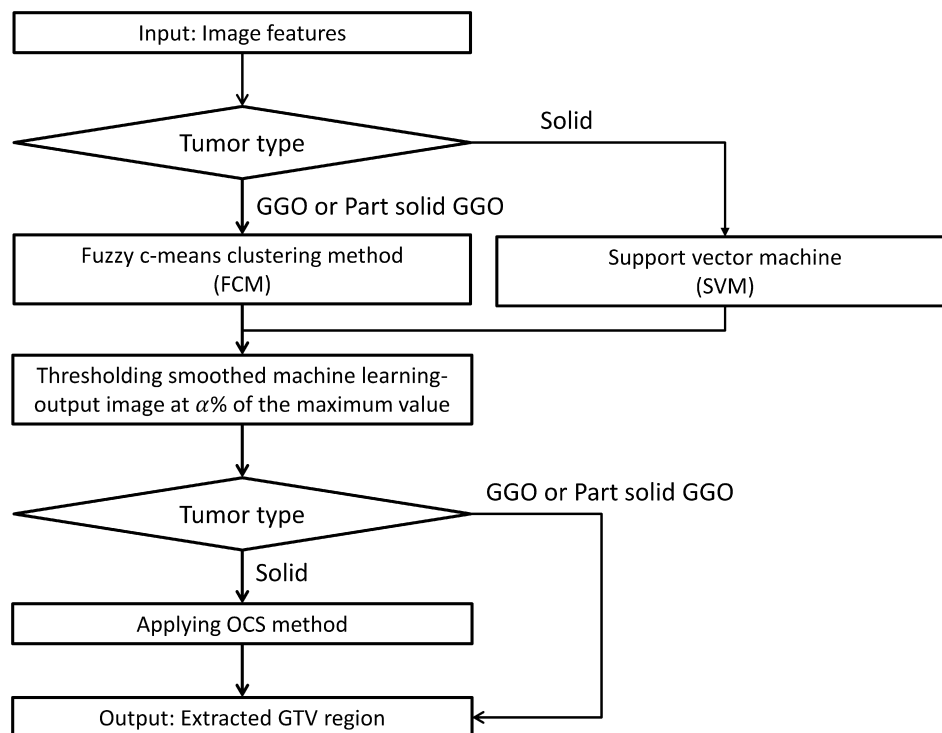
**Table 5**
Comparison of three-dimensional DSC obtained by FCM-based framework and combinational framework.

| Case No. | Tumor type | FCM-based framework | Combinational framework |
|---|---|---|---|
| 1 | Solid | 0.83 | 0.85 |
| 2 | Solid | 0.82 | 0.84 |
| 3 | Solid | 0.81 | 0.81 |
| 4 | Solid | 0.82 | 0.90 |
| 5 | Solid | 0.90 | 0.87 |
| 6 | Solid | 0.83 | 0.79 |
| 7 | GGO | 0.80 | 0.80 |
| 8 | GGO | 0.84 | 0.84 |
| 9 | GGO | 0.70 | 0.70 |
| 10 | GGO | 0.80 | 0.80 |
| 11 | Part solid GGO | 0.73 | 0.73 |
| 12 | Part solid GGO | 0.74 | 0.74 |
| 13 | Part solid GGO | 0.76 | 0.76 |
| 14 | Part solid GGO | 0.74 | 0.74 |
| 15 | Part solid GGO | 0.71 | 0.71 |
| 16 | Part solid GGO | 0.87 | 0.87 |
| Mean ± SD | | $0.79 \pm 0.06$ | $0.80 \pm 0.06$ |
| P (t-test) | | 0.30 | |
| P (F-test) | | 0.78 | |



Fig. 7. An algorithm of the combined framework.

## 5. Conclusions

We proposed an automated framework using pixel-based ML, i.e., FCM, ANN, SVM, and a combined framework for extracting GTV regions of lung cancer using datasets for planning CT and FDG-PET/CT images. The highest average DSC and its SD were obtained by the FCM-based framework. The DSC obtained by the FCM-based framework was $0.79 \pm 0.06$, whereas DSCs of the ANN- and SVM-based frameworks were $0.76 \pm 0.14$ and $0.73 \pm 0.14$, respectively. The FCM-based framework showed the significant robustness, and the computational time was significantly reduced to 77% for ANN-based framework and 7.1% for SVM-based framework. Therefore, the FCM-based framework can be useful for radiation oncologists to delineate tumor regions during radiation treatment planning.

## Conflicts of interest

The authors have no conflicts of interest to declare. The authors alone are responsible for the contents and writing of the paper.

## Acknowledgements

## References

[1] Matsuo Y, Shibuya K, Nagata Y, et al. Prognostic factors in stereotactic body radiotherapy for non-small-cell lung cancer. Int J Radiat Oncol Biol Phys 2011;79:1104–11.
[2] Zheng X, Schipper M, Kidwell K, et al. Survival outcome after stereotactic body radiation therapy and surgery for stage I non-small cell lung cancer: a meta-analysis. Int J Radiat Oncol Biol Phys 2014;90:603–11.
[3] Weiss E, Hess CF. The impact of gross tumor volume (GTV) and clinical target volume (CTV) definition on the total accuracy in radiotherapy theoretical aspects and practical experiences. Strahlenther Onkol 2003;179:21–30.
[4] International Commission on Radiation Units & Measurements (ICRU) Prescribing, Recording and Reporting Photon Beam Therapy. Supplement to ICRU Report 50, ICRU Report 62. Bethesda, MD, ICRU, USA, 1999, 833–4.
[5] Persson GF, Nygaard DE, Hollensen C, et al. Interobserver delineation variation in lung tumor stereotactic body radiotherapy. Br J Radiol 2012;85:e654–60.
[6] Esposito M, Maggi G, Marino C, et al. Multicentre treatment planning inter-comparison in a national context: the liver stereotactic ablative radiotherapy case. Phys Med 2016;32:277–83.
[7] Giglioli FR, Strigari L, Ragona R, et al. Lung stereotactic ablative body radiotherapy: a large scale multi-institutional planning comparison for interpreting results of multi-institutional studies. Phys Med 2016;32:600–6.
[8] Rousson M, Khamene A, Diallo M, et al. Constrained surface evolutions for prostate and bladder segmentation in CT images. Lect Notes Comput Sci (LNCS) 2005;3765:251–60.
[9] Nestla U, Kremp S, Schaefer-Schuler A, et al. Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. J Nucl Med 2005;46:1342–8.
[10] Zhang T, Tachiya Y, Sakaguchi Y, et al. Phantom study on three-dimensional target volume delineation by PET/CT-based auto-contouring. Fukuoka Acta Med 2010;101:238–46.
[11] Zhang G, Han D, Ma C, et al. Gradient-based delineation of the primary GTV on FLT PET in squamous cell cancer of the thoracic esophagus and impact on radiotherapy planning. Radiat Oncol 2015;10:11.
[12] Day E, Betler J, Parda D, et al. A region growing method for tumor volume segmentation on PET images for rectal and anal cancer patients. Med Phys 2009;36:4349–58.
[13] Aristophanous M, Penney BC, Martel MK, et al. A Gaussian mixture model for definition of lung tumor volumes in positron emission tomography. Med Phys 2007;34:4223–35.
[14] Geets X, Lee JA, Bol A, et al. A gradient-based method for segmenting FDG-PET images: methodology and validation. Eur J Nucl Med Mol Imaging 2007;34:1427–38.
[15] Hatt M, Cheze Le Rest C, Albarghach N, et al. PET functional volume delineation: a robustness and repeatability study. Eur J Nucl Med Mol Imaging 2011;38:3663–72.
[16] Jayasurya K, Fung G, Yu S, et al. Comparison of Bayesian network and support vector machine models for two-year survival prediction in lung cancer patients treated with radiotherapy. Med Phys 2010;37:1401–7.
[17] Belhassen S, Zaidi H. A novel fuzzy C-means algorithm for unsupervised heterogeneous tumor quantification in PET. Med Phys 2010;37:1309–24.
[18] Niyazi M, Landrock S, Elsner A, et al. Automated biological target volume delineation for radiotherapy treatment planning using FDG-PET/CT. Radiat Oncol 2013;8:180.
[19] Kerhet A, Small C, Quon H, et al. Application of ML methodology for PET-based definition of lung cancer. Curr Oncol 2010;17(1):41–7.
[20] El Naqa I, Yang D, Apte A, et al. Concurrent multimodality image segmentation by active contours for radiotherapy treatment planning. Med Phys 2007;34:4738–49.
[21] Ikushima K, Arimura H, Jin Z, et al. Computer-assisted framework for machine-learning-based delineation of GTV regions on datasets of planning CT and PET/CT images. J Radiat Res 2017;58:123–34.
[22] Cl Henschke, Yankelevitz DF, Mirtcheva R, et al. CT screening for lung cancer: frequency and significance of part-solid and nonsolid nodules. AJR Am J Roentgenol 2002;178(5):1053–7.
[23] Li F, Sone S, Abe H, et al. Malignant versus benign nodules at CT screening for lung cancer: comparison of thin-section CT findings. Radiology 2004;233(3):793–8.
[24] Henschke C, Yip R, Smith JP, et al. CT screening for lung cancer: part-solid nodules in baseline and annual repeat rounds. AJR Am J Roentgenol 2016;207(6):1179–84.
[25] Tokunaga C, Arimura H, Yoshiura T, et al. Automated measurement of three-dimensional cerebral cortical thickness in Alzheimer's patients using localized gradient vector trajectory in fuzzy membership maps. J Biomed Sci Eng 2013;6(03):327–36.
[26] Bezdek JC. A convergence theorem for the fuzzy ISODATA clustering algorithms. IEEE Trans Pattern Anal Mach Intell 1980;2(1):1–8.
[27] Pham DL, Prince JL. Adaptive fuzzy seg-mentation of magnetic resonance images. IEEE Trans Med Imaging 1999;18:737–52.
[28] Miyamoto S. Fuzzy c-means clustering. How to Clustering Analysis. Tokyo: Morikita Publishing Corporation; 1999. p. 27–33.
[29] Pham DL. Spatial models for fuzzy clustering. Comput Visi Image Underst 2001;84:285–97.
[30] Rumelhart DE, Hinton GE, Williams RJ. Learning Internal Representations by Error Propagation. Parallel Distributed PROCESSING. USA: Mit Press Cambridge; 1986. p. 318–62.
[31] Vapnik VN. The nature of statistical learning theory. 2nd ed. New York: Springer; 1999.
[32] Crum WR, Camara O, Hill DL. Generalized overlap measures for evaluation and validation in medical Image analysis. IEEE Trans Med Imaging 2006;25:1451–61.
[33] Schwars M, Gattaneo GM, Marrazzo L. Geometrical and dosimetrical uncertainties in hypofractionated radiotherapy of the lung: a review. Phys Med 2017;36:126–39.
[34] Giglioli FR, Clemente S, Esposito M, et al. Frontiers in planning optimization for lung SBRT. Phys Med 2017. https://doi.org/10.1016/j.ejmp.2017.05.064 (in press).