

ACCURATE TUMOR SEGMENTATION IN FDG-PET IMAGES WITH GUIDANCE OF COMPLEMENTARY CT IMAGES

Chunfeng Lian^{1,2} Su Ruan² Thierry Denœux¹ Yu Guo³ Pierre Vera^{2,4}

¹ Sorbonne Universités, Université de Technologie de Compiègne, CNRS, Heudiasyc, France

² Normandie Université, Université de Rouen, LITIS-QuantIF, France

³ Department of Biomedical Engineering, Tianjin University, China

⁴ Department of Nuclear Medicine, Centre Henri-Becquerel, France

ABSTRACT

While hybrid PET/CT scanner is becoming a standard imaging technique in clinical oncology, many existing methods still segment tumor in mono-modality without consideration of complementary information from another modality. In this paper, we propose an unsupervised 3-D method to automatically segment tumor in PET images, where anatomical knowledge from CT images is included as critical guidance to improve PET segmentation accuracy. To this end, a specific context term is proposed to iteratively quantify the conflicts between PET and CT segmentation. In addition, to comprehensively characterize image voxels for reliable segmentation, informative image features are effectively selected via an unsupervised metric learning strategy. The proposed method is based on the theory of belief functions, a powerful tool for information fusion and uncertain reasoning. Its performance has been well evaluated by real-patient PET/CT images.

Index Terms— PET/CT Image Segmentation, Unsupervised Learning, Information Fusion, Metric Learning, Belief Functions.

1. INTRODUCTION

Accurate segmentation of target tumor is of great importance in clinical oncology. The integrated positron emission tomography (PET)/computed tomography (CT) scanner effectively combines functional information from PET with anatomical information from CT, which could comprehensively describe tumor volumes for precise tumor delineation. While integrated PET/CT has become a reference imaging technique, many existing automatic methods still segment tumor in high-contrast but low-resolution PET images, without taking into account complementary knowledge from high-resolution but low-contrast CT images.

Existing methods to segment tumor solely in PET images can be categorized into different types [1], where the most common ones include thresholding methods [2], region growing methods [3], statistical methods [4], graph-based method-

s [5], and unsupervised learning methods [6, 7, 8], etc. Unlike supervised learning methods that need a training step, unsupervised methods (e.g., clustering methods) are efficient for PET image segmentation, especially considering positive tissues are usually heterogeneous in PET images with varying contours among different patients [1].

Apart from the above mono-modal methods, some methods have also been proposed to utilize information from CT (or PET) to guide tumor delineation in counterpart PET (or CT) images [9, 10]. For instance, a variant level set method has been proposed in [10] to segment tumor in FDG-PET images, where knowledge from corresponding CT images was adopted to guide the initialization of zero level set. These methods attempted to delineate an unique contour in PET/CT, which neglects the fact that the two distinct modalities describe complementary but not identical characteristics of the same target. Furthermore, they make decision only according to intensities of image voxels, while ignoring other image features that can supplementarily describe spatial context of each voxels.

In this paper, we propose a novel 3-D method based on unsupervised learning to segment tumor in functional PET images, where anatomical knowledge from CT images is integrated to guide the automatic segmentation procedure. The proposed method is developed in the framework of belief functions. As a powerful tool for representing, fusing, and reasoning with uncertain and imprecise knowledge, the theory of belief functions provides multiple ways to reliably fuse information from distinct sources, thus could be efficient for tumor segmentation in noisy and blurry PET images that guided by corresponding CT images. To this end, a special context term is proposed, which iteratively integrates segmentation in CT images to guide the corresponding segmentation in PET images. Considering information from the two different modalities are complementary but also distinct, the proposed context term only drives tumor contours in PET and CT to be consistent, but not force them to be identical. In addition, textural features, e.g., [11, 12, 13, 14], are taken into account in our study, so as to provide complementary information for comprehensive characterization of image

voxels. However, considering a large amount of textures can be extracted without prior knowledge concerning the most informative ones, a feature selection and distance metric adaptation procedure is included in the proposed method. A spatial regularization is also adopted in the proposed to protect local homogeneity during segmentation.

2. METHOD

To segment tumor in PET images using multiple features, the spatial evidential clustering algorithm, e.g., SECM [8], is fully improved by integrating unsupervised distance metric learning and feature selection; then, it will be further extended by including anatomical knowledge from counterpart CT images to guide the segmentation in PET images.

2.1. SECM with Adaptive Distance Metric

First of all, image features are extracted from a volume of interest (VOI). The VOI is a 3-D box defined by users, which fully includes the target tumor. Let $\{X_i^{pt}\}_{i=1}^n$ be feature vectors in \mathbb{R}^p . Using these extracted features, we attempt to find a matrix $\mathbf{M}^{pt} = \{m_i^{pt}\}_{i=1}^n$, where $m_i \in \mathbb{R}^3$ is the mass function [15] for the i th voxel, which quantifies the mass of belief that supports all possible hypothesis with respect to the cluster of this voxel. We assume that all the voxels belong either to the background (i.e. hypothesis ω_1) or to the positive tissue (i.e. hypothesis ω_2), without existence of outliers. Thus, the whole frame of clusters is set as $\Omega = \{\omega_1, \omega_2\}$. The mass function for each voxel obeys $m_i^{pt}(\{\omega_1\}) + m_i^{pt}(\{\omega_2\}) + m_i^{pt}(\Omega) \equiv 1$. Scalar $m_i^{pt}(\Omega)$ measures the ambiguity regarding the clusters ω_1 and ω_2 , thus blurring boundary and severe heterogeneous regions will have large mass on Ω . Finally, crisp segmentation is obtained by making decisions based on \mathbf{M}^{pt} .

Let cluster ω_1 (resp. ω_2) be represented by a center V_1^{pt} (resp. V_2^{pt}) in \mathbb{R}^p . For each nonempty subset $A_j \subseteq \Omega \setminus \emptyset$, we assume that its prototype \bar{V}_j^{pt} is defined as the barycenter of the centers associated to the singletons composing A_j , i.e., $\bar{V}_j^{pt} = \frac{1}{c_j} \sum_{k=1}^2 s_{kj} V_k^{pt}$, where s_{kj} is binary, and it equals 1 if and only if $\omega_k \in A_j$; while $c_j = |A_j|$ denotes the cardinality of A_j . Then, to learn matrix \mathbf{M}^{pt} , the SECM objective function taking into account the spatial prior can be represented as

$$\begin{aligned} \mathcal{J}_{secm}(\mathbf{M}^{pt}) &= \sum_{i=1}^n \sum_{A_j} c_j^2 [m_i^{pt}(A_j)]^2 [d^2(X_i^{pt}, \bar{V}_j^{pt})] \\ &\quad + \eta \sum_{i=1}^n \sum_{A_j} c_j^2 [m_i^{pt}(A_j)]^2 \left[\sum_{t \in \Phi(i)} d^2(X_{i,t}^{pt}, \bar{V}_j^{pt}) \right], \end{aligned} \quad (1)$$

where $\Phi = \{\Phi(i)\}_{i=1}^n$ is a 3-D neighborhood system, in which $\Phi(i) = \{1, \dots, T\}$ is the set of T neighbors of a voxel i , excluding i . Different with the original SECM that uses only intensities, in this study feature vectors for voxels in $\Phi(i)$

are $\{X_{i,1}^{pt}, \dots, X_{i,T}^{pt}\}$. Parameter $\eta > 0$ controls the influence of the spatial regularization, which should be predetermined according to the data at hand. Distances $d^2(X_i^{pt}, \bar{V}_j^{pt})$ and $d^2(X_{i,t}^{pt}, \bar{V}_j^{pt})$ measure clustering distortions of X_i^{pt} and its neighbor $X_{i,t}^{pt}$ to the focal set A_j , respectively.

The challenge for reliable quantification of $d^2(X_i^{pt}, \bar{V}_j^{pt})$ and $d^2(X_{i,t}^{pt}, \bar{V}_j^{pt})$ is that a relatively large amount of textures can be extracted, without prior knowledge concerning the most informative ones. Moreover, available high-dimensional feature vectors possibly contain unreliable variables due to noise and blur inherent in PET imaging. To overcome the above difficulties, a feature selection [16] and/or metric learning procedure [17] is desired. In line with our previous supervised learning methods [18, 19], here we propose an unsupervised way to learn a matrix $\mathbf{D}^{pt} \in \mathbb{R}^{p \times q}$, under the constraint $q \ll p$, by which the dissimilarity between any two feature vectors, say X_1^{pt} and X_2^{pt} , can be represented as $d^2(X_1^{pt}, X_2^{pt}) = (X_1^{pt} - X_2^{pt}) \mathbf{D}^{pt} (\mathbf{D}^{pt})^T (X_1^{pt} - X_2^{pt})^T$. Matrix \mathbf{D}^{pt} transforms the original feature space to a low-dimensional subspace, where important input features will have a strong impact when quantifying the dissimilarity.

Based on the above definition, the original objective function (1) integrating unsupervised metric learning is updated as

$$\begin{aligned} \mathcal{J}(\mathbf{M}^{pt}) &= \sum_{i=1}^n \sum_{A_j} c_j^2 [m_i^{pt}(A_j)]^2 [d^2(X_i^{pt}, \bar{V}_j^{pt})] \\ &\quad + \beta \sum_{i=1}^n \sum_{A_j} c_j^2 [m_i^{pt}(A_j)]^2 \left[\sum_{t \in \Phi(i)} d^2(X_{i,t}^{pt}, \bar{V}_j^{pt}) \right] \\ &\quad + \lambda \|\mathbf{D}^{pt}\|_{2,1} - \log(d^2(\bar{X}_{\omega_1}, \bar{X}_{\omega_2})), \end{aligned} \quad (2)$$

where the $\ell_{2,1}$ -norm sparsity regularization $\|\mathbf{D}^{pt}\|_{2,1}$ is included in the third term of (2) to select the most informative input features to reliably quantify the dissimilarity in the feature space. The last term of (2) is used to prevent the objective function being trivially solved with $\mathbf{D} = 0$, which collapses all the features vectors into a single point. Vectors \bar{X}_{ω_1} and \bar{X}_{ω_2} are two automatically predetermined seeds for the positive tissue and the background, respectively.

2.2. Segmentation in PET with CT Guidance

Considering CT images have relatively high-resolution, and can provide complementary anatomical knowledge to PET, segmentation in CT is included in the proposed method to guide the corresponding segmentation in PET images. To this end, after extracting features in each modalities, data in PET are up-sampled to that in CT. Then, we jointly look for two matrices $\mathbf{M}^{pt} = \{m_i^{pt}\}_{i=1}^n$ and $\mathbf{M}^{ct} = \{m_i^{ct}\}_{i=1}^n$, where m_i^{pt} and m_i^{ct} are the mass functions for two corresponding voxels in PET and CT, respectively. The global cost function

Table 1. Average DSC and HD of the proposed method compared to independent segmentation in each single modality.

	PET only	CT only	PET guided by CT
DSC	0.74 ± 0.15	0.51 ± 0.10	0.87 ± 0.04
HD	3.14 ± 1.39	7.44 ± 2.87	2.48 ± 1.00

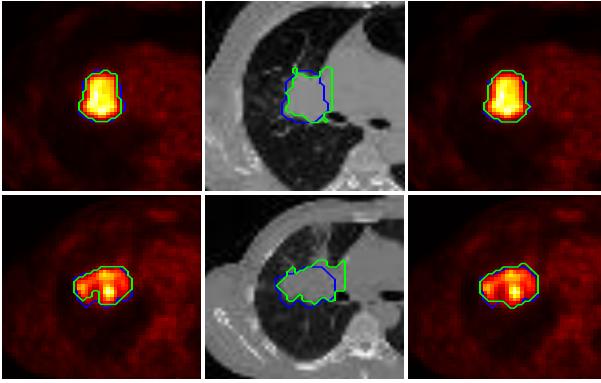


Fig. 1. Comparing segmentation in mono-modalities (first two columns) to segmentation in PET with CT guidance (last column). The two rows correspond to two different examples.

to this end is defined as

$$\begin{aligned} \mathcal{J}_{global}(\mathbf{M}^{pt}, \mathbf{M}^{ct}) = & \mathcal{J}(\mathbf{M}^{pt}) + \mathcal{J}(\mathbf{M}^{ct}) \\ & + \eta \mathcal{J}_{joint}(\mathbf{M}^{pt}, \mathbf{M}^{ct}), \end{aligned} \quad (3)$$

where both the first two terms are quantified by (2), the last term is a context penalty, and parameter η controls the influence of the last term. This context penalty quantifies the disagreement between the segmentation in PET and CT, which is specified by the dissimilarity between \mathbf{M}^{pt} and \mathbf{M}^{ct} . Based on the specific metric defined by Jousselme et al. [20] to calculate conflicts between two distinct sources of information, the last term in (3) is finally designed as

$$\mathcal{J}_{joint}(\mathbf{M}^{pt}, \mathbf{M}^{ct}) = \sum_{i=1}^n (\mathbf{m}_i^{pt} - \mathbf{m}_i^{ct}) \mathbf{Jac} (\mathbf{m}_i^{pt} - \mathbf{m}_i^{ct})^T, \quad (4)$$

where \mathbf{Jac} is a positive definite matrix whose elements are Jaccard indexes, i.e., $\mathbf{Jac}(A, B) = |A \cap B| / |A \cup B|$, $\forall A, B \in 2^\Omega \setminus \emptyset$. It is worth noting that the integration of CT segmentation here only ensures that the two complementary modalities have consistent tumor contours; however, the two contours are not identical, considering the PET and CT present distinct knowledge with respect to the target tumor.

2.3. Iterative Minimization Procedure

To output a pair of \mathbf{M}^{pt} and \mathbf{M}^{ct} , we propose an iterative scheme to minimize the cost function defined in (3), subject to $\sum_{A_j} \mathbf{m}_i^{pt}(A_j) = 1$, $\sum_{A_j} \mathbf{m}_i^{ct}(A_j)$, $\mathbf{m}_i^{pt}(A_j) \geq 0$, and

$m_i^{ct}(A_j) \geq 0$, $\forall i = 1, \dots, n$. More specifically, after initialization by the evidential c -means (ECM) [21], the optimization procedure alternates between the clustering in PET and the clustering in CT. For the clustering in each single modality, the optimization iterates between cluster assignment (i.e. \mathbf{M}^{pt} or \mathbf{M}^{ct} estimation) in the E-step, and both prototype determination (i.e. \mathbf{V}^{pt} or \mathbf{V}^{ct} estimation) and metric learning (i.e. \mathbf{D}^{pt} or \mathbf{D}^{ct} estimation) in the M-step:

The optimization in PET only relates to the minimization of the first term and the last term of (3), which is performed in an EM-like protocol.

E-Step: Given \mathbf{V}^{pt} , \mathbf{D}^{pt} , and \mathbf{M}^{ct} for the last iteration, the updating of \mathbf{M}^{pt} is determined by the first and the last term of (3), which has the analytical solution. Briefly, the calculation of $\mathbf{m}_i^{pt} \in \mathbf{M}^{pt}$, $\forall i = 1, \dots, n$, has the form of

$$\begin{aligned} \mathbf{m}_i^{pt} = & \left[2\eta m_i^{ct} \mathbf{Jac} + \frac{1 - 2\eta m_i^{ct} \mathbf{Jac} (2\mathbf{B} + 2\eta \mathbf{Jac})^{-1} \mathbf{1}^T \mathbf{1}}{\mathbf{1} (2\mathbf{B} + 2\eta \mathbf{D})^{-1} \mathbf{1}^T} \right. \\ & \times (2\mathbf{B} + 2\eta \mathbf{Jac})^{-1}, \end{aligned} \quad (5)$$

where $\mathbf{1} = [1, 1, 1]$, and $\mathbf{B} = diag(c_1^2 s_1^2, c_2^2 s_2^2, c_3^2 s_3^2)$ with $s_j^2 = d^2(X_i^{pt}, \bar{V}_j^{pt}) + \beta \sum_{t \in \Phi(i)} d^2(X_{i,t}^{pt}, \bar{V}_j^{pt})$, $\forall j \in \{1, 2, 3\}$. It can be found that m_i^{ct} is integrated as a constraint to update \mathbf{m}_i^{pt} .

M-step I: The updating of the prototypes \mathbf{V}^{pt} is only determined by the first term of (3). Further, it relates to the first two terms of (2). Let $f_j = (1 + \beta) \sum_{i=1}^n c_j^2 [\mathbf{m}_i^{pt}(A_j)]^2$ and $g_j = \sum_{i=1}^n c_j^2 [\mathbf{m}_i^{pt}(A_j)]^2 (X_i^{pt} + \beta \sum_{t \in \Phi(i)} X_{i,t}^{pt})$, $\forall j \in \{1, 2, 3\}$, the calculation of \mathbf{V}^{pt} obeys

$$\begin{cases} V_1^{pt} = \frac{2f_2(2g_1+g_3)+f_3(g_1-g_2)}{4f_1f_2+f_3(f_1+f_2)}; \\ V_2^{pt} = \frac{2f_1(2g_2+g_3)+f_3(g_2-g_1)}{4f_1f_2+f_3(f_1+f_2)}. \end{cases} \quad (6)$$

M-step II: The optimization of \mathbf{D}^{pt} relates to the first term of (3). Assume we have \mathbf{V}^{pt} and \mathbf{M}^{pt} obtained by the undergoing iteration, and \mathbf{D}^{pt} for the last iteration. Then, the Beck-Teboulle proximal gradient algorithm [22] is adopted to update \mathbf{D}^{pt} , considering (2) is partly smooth in term of matrix \mathbf{D}^{pt} with a singularity at zero.

Then, the adaptation of \mathbf{M}^{ct} , \mathbf{V}^{ct} , and \mathbf{D}^{ct} , which relates to the last two terms of (3), follows a similar way as that for \mathbf{M}^{pt} , \mathbf{V}^{pt} , and \mathbf{D}^{pt} discussed above. The only difference is that, to update \mathbf{M}^{ct} at the current step, \mathbf{V}^{pt} , \mathbf{D}^{pt} for the last iteration, and \mathbf{M}^{pt} obtained by the undergoing iteration should be substituted into (5).

The whole optimization procedure will terminate when the value of (3) has no significant change.

3. EXPERIMENTAL RESULTS

In this section, the proposed method was evaluated on 14 sets of 3-D FDG-PET/CT images acquired for lung cancer pa-

Table 2. Quantitative results obtained by different segmentation methods on all the 14 sets of 3D PET/CT images. The DSC and HD are presented as mean \pm std.

	3D-LARW	ECM	SECM	FCM-SW	Our method
DSC	0.82 ± 0.07	0.72 ± 0.13	0.77 ± 0.12	0.82 ± 0.11	0.87 ± 0.04
HD	4.43 ± 4.52	8.42 ± 3.71	5.89 ± 4.03	4.69 ± 3.73	2.48 ± 1.00

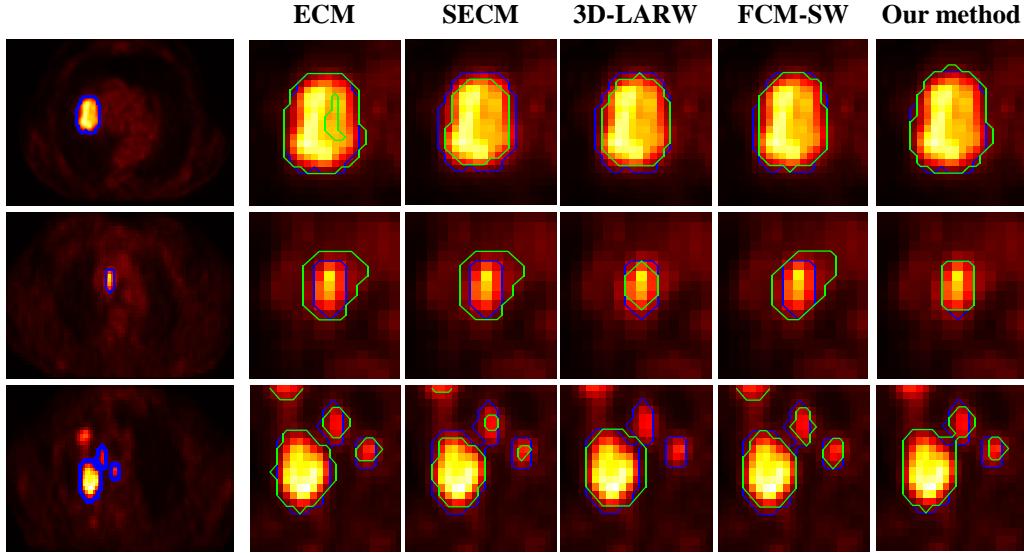


Fig. 2. Contours delineated by different methods (from the second column to the last column) for three different tumor volumes shown in the axial plane. The first column represents the input images with contours delineated by expert clinicians. The delineation by the five algorithms (green line) is compared with that by clinicians (blue line) in the following columns.

tients. A 28-dimensional feature vector was extracted for each voxel, which consists of 6 SUV-based, 7 GLSZM-based [13], and 15 GLCM-based [11] features. The performance of the proposed method was compared with that obtained in each single modality without guidance from the counterpart modality. In addition, its performance was also compared with that of other PET segmentation methods, namely 3D-LARW [5], FCM-SW [6], SECM [8], and also the original ECM [21]. Experimental results were quantified by the Dice coefficient (DSC) and the Hausdorff distance (HD), using manual delineation by clinicians in PET as the reference.

First of all, the proposed method was compared with that using solely PET and CT images. The independent segmentation in PET and CT was performed by directly minimizing (2). The average quantitative results for all the 14 sets of PET-CT images is summarized in Table 1. We can find that, including anatomical knowledge from CT to guide segmentation in PET led to the best DSC and HD. Two illustrative results are shown in Fig. 1, from which we can also find that the proposed method outperformed solely segmentation in mono-modalites in both two cases.

Then, the segmentation performance of the proposed method on all the 14 sets of PET-CT images was also compared with that of 3D-LARW [5], the original ECM [21],

SECM [8], and FCM-SW [6]. The quantitative comparison is shown in Table 2, from which we can find that the proposed method led to better performance, both DC and HD, than the other four algorithms in this experiment. To be more comprehensive, the visual examples obtained by these methods are also presented in Fig. 2. The first column of Fig. 2 presents the axial slices of three different tumors, where the first row is a slice corresponds to a large tumor, the second row represents a small tumor, while the last row represents a heterogenous tumor. The second column to the last column of Fig. 2 compare the contours delineated by the five different methods (green line) with that delineated by clinicians (blue line). We can find that the contours delineated by the proposed method (the last column) are more in consistent with the reference contours in this experiment.

4. CONCLUSION

In this paper, a belief function based method has been proposed for the delineation of tumor in PET using corresponding CT images as the guidance. A specific context term and unsupervised metric learning procedure are included in the proposed method to ensure its performance. The experimental results obtained on 14 real-patient FDG-PET/CT stacks have shown the effectiveness of the proposed method.

5. REFERENCES

- [1] Brent Foster et al., “A review on segmentation of positron emission tomography images,” *Comput. Biol. Med.*, vol. 50, pp. 76–96, 2014.
- [2] Sébastien Vauclin et al., “Development of a generic thresholding algorithm for the delineation of 18FDG-PET-positive tissue: application to the comparison of three thresholding models,” *Phys. Med. Biol.*, vol. 54, no. 22, pp. 6901–6916, 2009.
- [3] Hua Li et al., “A novel PET tumor delineation method based on adaptive region-growing and dual-front active contours,” *Med. Phys.*, vol. 35, no. 8, pp. 3711–3721, 2008.
- [4] Michalis Aristophanous et al., “A Gaussian mixture model for definition of lung tumor volumes in positron emission tomography,” *Med. Phys.*, vol. 34, pp. 4223–4235, 2007.
- [5] DP Onoma et al., “Segmentation of heterogeneous or small FDG PET positive tissue based on a 3D-locally adaptive random walk algorithm,” *Comput. Med. Imag. Graph.*, vol. 38, no. 8, pp. 753–763, 2014.
- [6] Saoussen Belhassen et al., “A novel fuzzy c-means algorithm for unsupervised heterogeneous tumor quantification in PET,” *Med. Phys.*, vol. 37, no. 3, pp. 1309–1324, 2010.
- [7] Brent Foster et al., “Segmentation of PET images for computer-aided functional quantification of tuberculosis in small animal models,” *IEEE TBME*, vol. 61, no. 3, pp. 711–724, 2014.
- [8] Benoît Lelandais et al., “Fusion of multi-tracer PET images for dose painting,” *Med. Image Anal.*, vol. 18, no. 7, pp. 1247–1259, 2014.
- [9] Julien Wojak et al., “Joint variational segmentation of ct-pet data for tumoral lesions,” in *ISBI*. IEEE, 2010, pp. 217–220.
- [10] Wei Mu et al., “A segmentation algorithm for quantitative analysis of heterogeneous tumors of the cervix with 18 F-FDG PET/CT,” *IEEE-TBME*, vol. 62, no. 10, pp. 2465–2479, 2015.
- [11] Leen-Kiat Soh et al., “Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices,” *IEEE-TGRS*, vol. 37, no. 2, pp. 780–795, 1999.
- [12] Jun Zhang et al., “Local energy pattern for texture classification using self-adaptive quantization thresholds,” *IEEE-TIP*, vol. 22, no. 1, pp. 31–42, 2013.
- [13] Guillaume Thibault et al., “Advanced statistical matrices for texture characterization: application to cell classification,” *IEEE-TBME*, vol. 61, no. 3, pp. 630–637, 2014.
- [14] Jun Zhang et al., “Scale invariant texture representation based on frequency decomposition and gradient orientation,” *Pattern Recogn. Lett.*, vol. 51, pp. 57–62, 2015.
- [15] Glenn Shafer, *A mathematical theory of evidence*, Princeton University Press, 1976.
- [16] Mingxia Liu et al., “Pairwise constraint-guided sparse learning for feature selection,” *IEEE-TCyb*, vol. 46, no. 1, pp. 298–310, 2016.
- [17] Kilian Q Weinberger et al., “Distance metric learning for large margin nearest neighbor classification,” *JMLR*, vol. 10, no. Feb, pp. 207–244, 2009.
- [18] Chunfeng Lian et al., “Robust cancer treatment outcome prediction dealing with small-sized and imbalanced data from FDG-PET images,” in *MICCAI*, 2016, pp. 61–69.
- [19] Chunfeng Lian et al., “Dissimilarity metric learning in the belief function framework,” *IEEE-TFS*, vol. 24, no. 6, pp. 1555–1564, 2016.
- [20] Anne-Laure Jousselme et al., “A new distance between two bodies of evidence,” *Information Fusion*, vol. 2, no. 2, pp. 91–101, 2001.
- [21] Marie-Hélène Masson et al., “ECM: An evidential version of the fuzzy c-means algorithm,” *Pattern Recognition*, vol. 41, no. 4, pp. 1384–1397, 2008.
- [22] Amir Beck and Marc Teboulle, “A fast iterative shrinkage-thresholding algorithm for linear inverse problems,” *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.