2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)
Venice, Italy, April 8-11, 2019

NOVEL RADIOMIC FEATURES BASED ON GRAPH THEORY FOR PET IMAGE ANALYSIS

Zhiling Zhou[*1 2], *Ning Guo*[*1], *Jianan Cui*[1], *Xiaxia Meng*[1], *Yiwei Hu*[3], *Han Bao*[1], *Xiang Li*[1], *Quanzheng Li*[1]

[1]Department of Radiology, Massachusetts General Hospital and Harvard Medical School
[2]Department of Computer Science and Technology, Zhejiang University
[3]Department of Computer Science, Yale University *Joint first authors.

ABSTRACT

We proposed a series of new radiomic features for PET image analysis base on graph theory and network analysis. Current PET radiomic features are mostly developed or transferred from CT images analysis which mainly focus on texture information. PET images usually contain functional information with lower resolution. Thus current radiomic features lack interpretability and specificity for PET image quantification. Meanwhile, a large number of texture features have similar definitions which cause severe redundancy for analysis and classification task. We proposed novel radiomic features based on graph theory that can specifically represent PET image characters. Using a set of tools in graph analysis, a new series of PET radiomic features that reveal different attributes of tumor, particularly intratumoral heterogeneity, are extracted. We applied our proposed method to lung cancer diagnosis and prognosis to evaluate performance of new features. Using ANN as classifier, our graph-based features outperformed traditional PET radiomic features. Furthermore, the combination of our features and tradition features can achieve an even better performance. It indicates that our graph-based features reveal significant and unique information of tumor in PET images.

Index Terms — radiomics, ^{18}F-FDG PET, graph theory, network analysis, quantitative analysis

1. INTRODUCTION

Positron emission tomography/computed tomography (^{18}F-FDG PET/CT) is a well-established approach for cancer evaluation owing to its great specificity and sensitivity, and becomes popular in clinical routine and clinical trials [1, 2]. Moreover, the observation and quantification of ^{18}F-FDG uptake shown as the standardized uptake value(SUV) has been employed to reveal different kinds of characteristics of tumors[3]. Recently, radiomic analysis for PET/CT has been widely studied for diagnosis and treatment evaluation of cancer patients owing to its advantage of providing abundant image-derived features. However, most of radiomic features focus on texture information which are more appropriate to quantify CT images [4, 5]. Therefore, current radiomic features are less specific and unexplainable when applied to PET images. At meanwhile, there are several issues and limitations of current PET radiomic analysis. First and the most important one is that, PET image, which has poor resolution, doesn't show much texture details in tumor region. The advantages of PET image rely on the functional information and distribution of radioactive tracer which is contemporary to those high-resolution anatomical images. In consequence, conventional texture analysis is not suitable for PET radiomic analysis. Secondly, the basic texture features are generally extracted from a 2D image slices, which results in the loss of both spatial information and inter-plane interaction for 3D images [6]. Although matrix-based texture analysis revealed some spatial relationship of pixels, the features are usually calculated with statistical method and doesn't contain real 3D relationship of voxels [6]. Another issue of traditional texture analysis is that the large number of features results in high redundancy when fed into the classification. A typical radiomic analysis using texture features may refer to over 100 features, and sometimes with different setting, this number can reach 1000 [7]. The complication of texture features significantly limits its application to clinical practice due to the lack of interpretation and generalizability. Contradictory results and controversies can be resulted from the complex preprocessing steps as well as numerous implementation choices in the workflow of radiomic analysis [8]. Thus, there is an urgent need to develop PET specific radiomic features. However, few PET-specific features were developed for heterogeneity or radiomic analysis so far. Eary et al. developed a 3D ellipsoidal template as prior knowledge to quantify intratumoral heterogeneity and spatial distribution [9]. And our previous work introduced Global Moran's (I) to the quantification of intratumoral heterogeneity specifically for PET image [10].

In this work, we propose a novel concept of quantitative modeling for PET image using graph theory. It enables us to extract PET-specific features using network analysis which is a popular tool in graph theory. To the best of our knowledge, this is the first work to interpret tumor image with graph-network features. We hypothesize that the network-based PET features contain comprehensive information of both spatial and functional aspects and perform superior to the traditional radiomic features for PET image analysis. By building up a graph using PET image of each tumor lesion, 26 network-based

978-1-5386-3640-4/19/$31.00 ©2019 IEEE

features are extracted for quantitative analysis. Then the performance of our new series of features was evaluated for lung cancer diagnosis and treatment responses using patient's PET images. The comparison between the traditional radiomic features and proposed features was conducted. Furthermore, we investigated the performance of combined traditional and proposed features in the classification. And the boost of performance indicates that new features reveal more specific information of PET images that is complementary to texture features.

2. METHOD AND MATERIALS

2.1. Patients

We study on a cohort of 133 patients who are diagnosed non-small-cell lung cancer (NSCLC). There are 32 female and 101 male subjects with averaged age 68 ± 8.8 y. For staging task, the patients with NSCLC were grouped as early and advanced malignancy based on their AJCC staging record. AJCC staging system (version 7) is the standard clinical criteria using TNM staging standard [11]. Depending on the TNM stage, the subjects are grouped into four groups, Stage I, II, III and IV. The early malignancy group consists of 104 subjects of stage I and II, while advanced malignancy group contains 24 patients of stage III and IV. Each patient underwent PET/CT imaging prior to the surgery or other treatments. All of the patients were followed up for treatment response after surgery. When evaluating the image features performance for treatment response task, the patients are separated into two groups depending on the appearance of recurrence. There are 28 subjects reported with recurrence including regional, local or distant during the follow-up period. Meanwhile, those subjects who didn't develop recurrence at least one year after surgery were grouped as non-recurrence, which consists of 74 subjects.

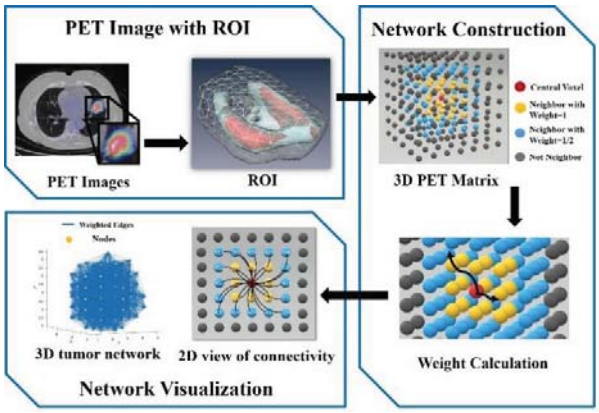

Fig. 1. Schematic of the network construction process. First, 3D ROI is extracted from PET image. Afterwards, an unweighted association matrix is built using adjacent points according to distance threshold. Then the value difference between each two adjacent nodes is used as the weight for this node pair. Finally, the adjacent matrix of current case is constructed. The PET/CT images acquired before the treatment are obtained from The Cancer Imaging Archive (TCIA) [12]. Region of interests (ROIs) are manually contoured on tumor lesions by radiologists referring to both PET and CT images.

2.2. Network

The original image ROI is a 3D matrix where each voxel stands for a node in the graph. As illustrated in Fig 1, for each ROI, a graph was built up by defining the SUV as node value and using the difference of SUV as edges between each pair of nodes. We employed the weighted undirected network (WUN), where each edge has its different weight but no orientation, which reveals more information about the connectivity and relative position between linked nodes [13]. Thus, the constructed network contains both spatial relationship between each pair of nodes and the image gradient represented by SUV difference. Then, a stride of two was used as a threshold to determine valid connectivity between each pair of nodes. Therefore, at most 124 neighboring nodes within threshold will be connected for each voxel in a 3D image. Thus, the weight between each pair of nodes is defined as:

$$\text{weight}(i,j) = \frac{|PET\_value(i) - PET\_value(j)|}{distance(i,j)} \quad \text{.....(1)}$$

2.3. Feature Extraction

With the constructed network, we applied the graph theoretical approaches to analyze the complexity and topological properties of this network including modularity, centrality, hierarchy and distribution of network hubs [14], which potentially refers to intratumoral heterogeneity. All the network-based features investigated are summarized in Table 1.

Table 1 Network features and description

Network Feature	Description
Basic Concept & Measure	
Edge Size	Number of edges
Node Size	Number of nodes
Average Weighted Degree	
Minimal/Maximal Spanning Tree	Sum of edge value of minimal/ maximal spanning tree
Maximal Degree	Number of edges incident to the node
Network Density	Density of edges of the graph
Weighted Network Density	Weighted density of links of the graph
Average Degree	The average weighted degree across all nodes
Maximal Weighted Degree	The maximal weighted degree across all nodes
Voxel Number	Number of voxels in ROI
Edge Density	Density of edges in graph
Eigenvalue Related	
Minimal/Maximal Eigenvalue Eigenvalue Distribution	Min/Max eigenvalues of the Laplacian of the graph
Graph Energy	Sum of the absolute values of the real components of the eigenvalues
Eigenvalue Distribution	Proportion of the first 5% eigenvalues in all eigenvalue

Measures of Centrality	
Mean Cluster Coefficient	Duncan J. Watts (1988) [15]
Max Cluster Coefficient	
Min cluster coefficient	
Measures of Segregation	
Global Efficiency	Global efficiency of the network [16]
S Metric	Sum of products of degrees across all edges
Pearson	Pearson degree correlation
Newman-Girvan Modularity	Newman-Girvan community finding algorithm
Classical Newman Girvan Modularity	Community structure in network
Measures of Resilience	
Degree distribution	Standard deviation of the value across all nodes

For the performance comparison, a total of 54 traditional radiomic features were also calculated [17]. The 54 conventional texture features consist of 9 gray level co-occurrence matrix-GLCM features, 13 gray level run length matrix-GLRLM features, 13 gray level size zone matrix-GLSZM features, 5 neighborhoods gray tone difference matrix-NGTDM features, 3 global features and other 14 other basic features including volume and SUVmean & max.

2.4. Classification:

An artificial neural network (ANN) was setup for classification, implemented in Keras framework. The structure of our neural network is illustrated in Figure 2. With each neuron connected to one feature in the input layer, two layers with 32 neurons and 8 neurons respectively were then connected, followed by the sigmoid activation function, to provide the classification output.

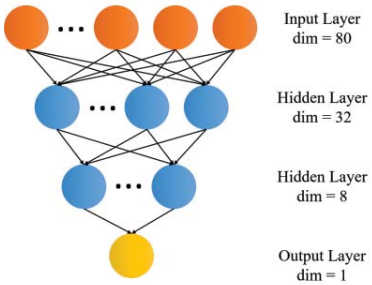

Fig. 2. The structure of our artificial neural network, which consists of one input layer, two hidden layers and one output layer.

Due to the limited sample size, the ANN classifier was trained and validated using five-fold cross-validation to make full use of our data as well as to improve the reliability of our result. To alleviate the model bias caused by class imbalance, Synthetic Minority Over-Sampling Technique (SMOTE) was applied independently to the training and testing samples for each cycle. The dropout rate of hidden layers was set to 0.3 to avoid overfitting [18]. The performance of our network was evaluated by its accuracy, area under curve (AUC), as well as sensitivity and specificity on different tasks including cancer staging and recurrence prediction after treatment. In order to test the comprehensiveness among network features and traditional radiomic features, we combined both feature series and evaluated the performance using same classification setup for the same tasks.

3. RESULTS

Our network features yield satisfying performance in diagnosis and treatment evaluation tasks, as described in the following subsections and summarized in Table 2.

Table 2 Performance of different groups of features.

Label	Feature	Accuracy	AUC	Sensitivity	Specificity
AJCC staging	Network	86.55%	0.86	76.90%	96.19%
	Traditional	82.14%	0.82	66.19%	98.10%
	Combination	92.86%	0.92	88.57%	97.14%
Recurrence	Network	80.43%	0.80	68.86%	92.00%
	Traditional	78.85%	0.74	63.33%	85.14%
	Combination	85.76%	0.85	76.86%	94.67%

3.1. AJCC staging

In differentiating the early from advanced malignant subjects, the network-based features achieved an accuracy of 86.55% with 76.90% sensitivity, 96.19% specificity and AUC of 0.86. At meantime, traditional features yielded 82.14% accuracy with 66.19% sensitivity and 98.10% specificity, and the AUC is 0.82. Thus, network features are superior in the NSCLC staging when compared with traditional radiomic features. The combined features have the best performance (92.86% accuracy with 88.57% sensitivity and 97.14% specificity).

3.2. Recurrence prediction

For recurrence prediction task, the network-based features reached an accuracy of 80.43% with 68.86% sensitivity and 92.00% specificity, and the AUC is 0.80. At meantime, traditional features yielded 78.85% accuracy with 63.33% sensitivity and 85.14% specificity, and the AUC is 0.74. With the combination of both network and traditional features, the accuracy of recurrence prediction is boosted to 85.76% with 76.86% sensitivity and 94.67% specificity, and the AUC is 0.85. The improvement of performance in AJCC staging and recurrence prediction demonstrated that the proposed network features can reveal tumor characteristics represented in PET image from different perspectives, compared with radiomic features, and thus the combination of them are even more powerful.

3.3. Importance of Features

We analyzed the contribution of each features in the combined analysis. It refers to the weight vector W that connects the hidden layer and input layer [19]. $W_{i,j}$ represents the weight between the j^{th} hidden neuron and the i^{th} input feature. We sorted the value $\sum_j W_{ij}$ for each feature when performing the above two tasks. We plot the

ranked features of AJCC staging task as example in Fig. 3. Among the top 9 features with highest importance, 6 of them are network based features. The most important feature is minimal eigenvalue of the Laplacian of the graph. We observed similar pattern for the recurrence predition task. These resutls imply that the network based features carry significant and unique information of PET images. A well-trained ANN reduces the weight of unimportant features automatically, which makes the feature selection unnecessary in our analysis.

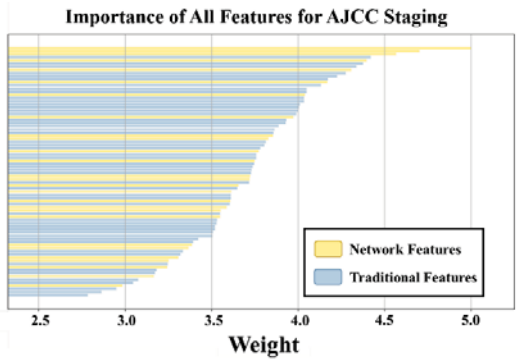

Fig. 3. Importance distribution of combined features in ANN classification for diagnostic staging

4. CONCLUSION

In summary, we presented a series of new radiomic features for PET image analysis based on graph theory and network analysis; these features could be used for AJCC staging and recurrence prediction of lung cancer patients. The superior performance of new features over traditional radiomic features demonstrated that the network-based features are more suitable to characterize the attributions of PET images. Furthermore, when combined with traditional features, the network-based radiomic features can achieve even better performance and thus become very promising tools for cancer diagnosis and prognosis.

5. ACKNOWLEDGEMENT

This work was supported by the National Institutes of Health under grants 1RF1AG052653, 1P41EB022544 and C06 CA059267.

6. REFERENCE

[1] B. E. Hillner *et al.*, "Impact of positron emission tomography/computed tomography and positron emission tomography (PET) alone on expected management of patients with cancer: initial results from the National Oncologic PET Registry," *Journal of Clinical Oncology,* vol. 26, no. 13, pp. 2155-2161, 2008.

[2] M. E. Juweid and B. D. Cheson, "Positron-emission tomography and assessment of cancer therapy," *New England Journal of Medicine,* vol. 354, no. 5, pp. 496-507, 2006.

[3] T. Berghmans *et al.*, "Primary tumor standardized uptake value (SUVmax) measured on fluorodeoxyglucose positron emission tomography (FDG-PET) is of prognostic value for survival in non-small cell lung cancer (NSCLC): a systematic review and meta-analysis (MA) by the European Lung Cancer Working Party for the IASLC Lung Cancer Staging Project," *Journal of Thoracic Oncology,* vol. 3, no. 1, pp. 6-12, 2008.

[4] D. Mackin *et al.*, "Measuring CT scanner variability of radiomics features," *Investigative radiology,* vol. 50, no. 11, p. 757, 2015.

[5] Z. Guo, X. Li, H. Huang, N. Quo, and Q. Li, "Medical image segmentation based on multi-modal convolutional neural network: Study on image fusion schemes," in *Biomedical Imaging (ISBI 2018), 2018 IEEE 15th International Symposium on*, 2018, pp. 903-907: IEEE.

[6] S. Chicklore, V. Goh, M. Siddique, A. Roy, P. K. Marsden, and G. J. Cook, "Quantifying tumour heterogeneity in 18 F-FDG PET/CT imaging by texture analysis," *European journal of nuclear medicine and molecular imaging,* vol. 40, no. 1, pp. 133-140, 2013.

[7] A. Chalkidou, M. J. O'Doherty, and P. K. Marsden, "False discovery rates in PET and CT studies with texture features: a systematic review," *PloS one,* vol. 10, no. 5, p. e0124165, 2015.

[8] M. Hatt, F. Tixier, L. Pierce, P. E. Kinahan, C. C. Le Rest, and D. Visvikis, "Characterization of PET/CT images using texture analysis: the past, the present… any future?," *European journal of nuclear medicine and molecular imaging,* vol. 44, no. 1, pp. 151-165, 2017.

[9] J. F. Eary, F. O'Sullivan, J. O'Sullivan, and E. U. Conrad, "Spatial heterogeneity in sarcoma 18F-FDG uptake as a predictor of patient outcome," *Journal of nuclear medicine: official publication, Society of Nuclear Medicine,* vol. 49, no. 12, p. 1973, 2008.

[10] N. Guo, R.-F. Yen, G. El Fakhri, and Q. Li, "SVM based lung cancer diagnosis using multiple image features in PET/CT," in *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2015 IEEE*, 2015, pp. 1-4: IEEE.

[11] S. Edge, "American joint committee on cancer," *AJCC cancer staging manual,* 2010.

[12] H. J. Aerts *et al.*, "Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach," *Nature communications,* vol. 5, p. 4006, 2014.

[13] M. E. Newman, "Analysis of weighted networks," *Physical review E,* vol. 70, no. 5, p. 056131, 2004.

[14] G. Bounova and O. de Weck, "Overview of metrics and their correlation patterns for multiple-metric topology analysis on heterogeneous graph ensembles," *Physical Review E,* vol. 85, no. 1, p. 016117, 2012.

[15] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world'networks," *nature,* vol. 393, no. 6684, p. 440, 1998.

[16] V. Latora and M. Marchiori, "Efficient behavior of small-world networks," *Physical review letters,* vol. 87, no. 19, p. 198701, 2001.

[17] M. Sollini, L. Cozzi, L. Antunovic, A. Chiti, and M. Kirienko, "PET Radiomics in NSCLC: state of the art and a proposal for harmonization of methodology," *Scientific reports,* vol. 7, no. 1, p. 358, 2017.

[18] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The Journal of Machine Learning Research,* vol. 15, no. 1, pp. 1929-1958, 2014.

[19] J. D. Olden and D. A. Jackson, "Illuminating the "black box": a randomization approach for understanding variable contributions in artificial neural networks," *Ecological modelling,* vol. 154, no. 1-2, pp. 135-150, 2002.