

Tree-based Methods for Characterizing Tumor Density Heterogeneity

Katherine Shoemaker

*Statistics Department, Rice University
Houston, Texas, 77005, USA*

Brian P. Hobbs

*Biostatistics, MD Anderson Cancer Center
Houston, Texas, 77030, USA
E-mail: bphobbs@mdanderson.org*

Karthik Bharath

*School of Mathematical Sciences, University of Nottingham
Nottingham, NG7 2RD, UK*

Chaan S. Ng

*Radiology, MD Anderson Cancer Center
Houston, Texas, 77030, USA*

Veerabhadran Baladandayuthapani

*Biostatistics, MD Anderson Cancer Center
Houston, Texas, 77030, USA*

Solid lesions emerge within diverse tissue environments making their characterization and diagnosis a challenge. With the advent of cancer radiomics, a variety of techniques have been developed to transform images into quantifiable feature sets producing summary statistics that describe the morphology and texture of solid masses. Relying on empirical distribution summaries as well as grey-level co-occurrence statistics, several approaches have been devised to characterize tissue density heterogeneity. This article proposes a novel decision-tree based approach which quantifies the tissue density heterogeneity of a given lesion through its resultant distribution of tree-structured dissimilarity metrics computed with least common ancestor trees under repeated pixel re-sampling. The methodology, based on statistics derived from Galton-Watson trees, produces metrics that are minimally correlated with existing features, adding new information to the feature space and improving quantitative characterization of the extent to which a CT image conveys heterogeneous density distribution. We demonstrate its practical application through a diagnostic study of adrenal lesions. Integrating the proposed with existing features identifies classifiers of three important lesion types; malignant from benign (AUC = 0.78), functioning from non-functioning (AUC = 0.93) and calcified from non-calcified (AUC of 1).

Keywords: Radiomics, Imaging Features, Heterogeneity, Galton-Watson Trees

© 2017 The Authors. Open Access chapter published by World Scientific Publishing Company and distributed under the terms of the Creative Commons Attribution Non-Commercial (CC BY-NC) 4.0 License.

1. Introduction

One of the critical aspects to the study of solid lesions is intra-tumor heterogeneity (ITH). Solid lesions are often heterogeneous phenotypically, physiologically, and genetically, due to variations in processes such as cell proliferation, cell death, and local environmental factors.¹⁻³ Cellular diagnostic techniques such as biopsies are not only invasive, but they also do not allow for a thorough or complete investigation of the entire tumor environment. In order to get a more comprehensive picture of the entire lesion environment without having to take multiple biopsies or depend on qualitative visual assessments, quantitative imaging features can be mined with analytical techniques often called radiomics.⁴ These radiomic features are objectively assessed and quantitatively descriptive of the lesion phenotypes and can be used to develop models that can be used in prediction, classification or diagnosis. The “radiomics hypothesis” that is central to this strategy is that advanced levels of analytics on imaging data can capture information that would not otherwise be available.⁵ It has been hypothesized that this information on phenotypic patterns are reflective of complementary tumor characteristics at molecular, cellular and genetic levels.⁶

There are many possible ways to extract this radiomic data from routine images: features that describe size and shape, features that describe the intensity of and relationship between pixel values, textures features, and fractal features.⁶ The ability to access a large number of quantitative features from images is now possible due to the progress made in imaging techniques, but issues like high levels of correlation between these features have led to the need to determine which of these features to use in downstream analyses and interpretation. As with any big data problem, when working with such a large number of variables, it is imperative to balance interpretability vs. computational tractability.

Currently the most commonly used radiomic features are texture-seeking. They can be divided into two categories, intensity and texture. Intensity features capture the shape of the histogram of the pixel values, while texture features describe the spatial distribution and pattern of the pixel values.⁷ Texture-seeking methods can be divided into four categories: Non-spatial methods (NSM), Spatial Grey Level methods (SGLM), Fractal Analysis and Filters and Transforms. As NSM and SGLM are the methods used in the majority of analysis, we will focus in on those and give a brief description before highlighting potential new features derived from the tumor heterogeneity trees (THT).

NSM are intensity based features, comprised of first order statistics computed on the image grey level data. These are basic metrics, which include metrics such as the first order features, taken from the grey level image, include minimum and maximum, as well as computations such as range, mean, standard deviation, variance, median, skewness, kurtosis, entropy, root mean square (RMS) and total energy.⁷

The most frequently used of the central moments are variance, skewness and kurtosis.⁸ Variation gives an idea of the size of the spread of the distribution around the mean, skewness is a measure of asymmetry around the mean, and kurtosis is a measure of the sharpness of the histogram. Intensity features give insight into how the pixel densities are distributed, but cannot give insight into their relative spatial positions, which limits their potential for describing the texture features of the image. The size of the images is a confounding factor for

several of these metrics, but the simplicity of these metrics is an advantage, and they contain a nontrivial amount of information about the image.

SGLM are texture features that are used to interrogate the spatial relationships between the grey levels of the image. A Grey Level Co-occurrence Matrix (GLCM) is an object that describes the spatial relationship of the grey levels of pixels in an image by counting the number of times two grey level valued pixels appear a specific distance and angle from each other. Before calculating the matrix, the number of grey levels must be chosen. This is done while considering the level of detail desired, the number of unique pixel values available, and the distribution of the density of these pixel values. These grey level “buckets” can be of consistent size, but the choice can alternatively be made to split the pixel value distribution along percentiles. This approach is a potential way to account for outlier pixel values within the pixel distribution. For example, the grey level matrix of angle 0° and distance 1, the cell (i, j) will contain the number of times a pixel of grey level j appears immediately to the right of a pixel of grey level i . This matrix can be made symmetric or not, and can be normalized by dividing each count by the number of pixels in the image. The angle can vary, allowing for comparison vertically or along the diagonal as well as the horizontal example given previously. The distance can be changed as well, with no restriction beyond the size of the image along the chosen angle.

GLCM features were first proposed by Haralick in 1973⁹ and these features are easily computable from the GLCM matrix and include features that measure attributes such as image coarseness, symmetry, energy, and heterogeneity.⁷ These features are often computed on multiple combinations of angle, distance and number of grey levels, leading to a large set of features that can be used in model building and data analysis.

While the above metrics work well in capturing the morphological characteristics of the tumor, they are limited in their characterization of IHT. In consideration of lesion heterogeneity, tree-structured objects offer hierarchical dissimilarity processes that may better reflect the “relationship” between the pixels as a representation of the cellular evolution of cancer and of the lesion as it grows and develops. It is well-established that cancer as a disease starts at a “single point,” a cell, which divides and proliferates outward to an extent that is allowed by the local immune and tissue environments. Each cell division is a biological bifurcation and this process is repeated up to the moment of the diagnostic scan capturing the cross-sectional state of the tumor. Conceptually, malignant cell proliferation is well characterized by a binary decision tree, which describes a hierarchical splitting process that divides iteratively from a common root until arriving at the final state of nodes or leaves. Considering tissue density as a surrogate¹⁰ for the cellular division process, the growth process of a tumor may be well characterized by dissimilarity measures of pixel intensities obtained from tree-structured objects.

With this in mind, it is the goal of this paper to briefly discuss some of the various methods used to interrogate tumor texture, and to present a potential additional method based on tree-based analysis of the lesions, which will be used in conjunction with the currently used methods on a set of solid adrenal lesions to capture various aspects of cancer progression and development. Specifically, we use the feature set to classify benign from malignant, functioning

from non-functioning, and calcified from non-calcified, lesions with encouraging results.

2. Methods

Trees are a data type that is a specific subset of graphs. They are a directed, acyclic set of linked nodes that are connected by edges. The parental node is called the root, while the terminal nodes are called leaves. Depending on construction, they can either start at the leaves and repeatedly combine pairs (in a binary tree) until all leaves are grouped together, or start at the root and divide until each leaf is separate from all others. This branching process shows the relationship between the leaves and the history of how they separated from each other, when and in what order. If this process is applied to pixels from an image, and we consider that a radiological image is a representation the cells present inside the body, the tree can give a “history” of the representation of these cells and how they have divided from an original source, to a reasonable degree. In cancer, the pattern of growth is critical to the lesion development and this growth can be affected by the cellular environment, and is reflective of the ITH. A goal of introducing this feature is capturing this ITH via building the tree-based relationship between the pixels as intermediates for cells. Figure 2 gives a high level summary of the steps for creating tree-based features from radiological images which we describe in ensuing sections.

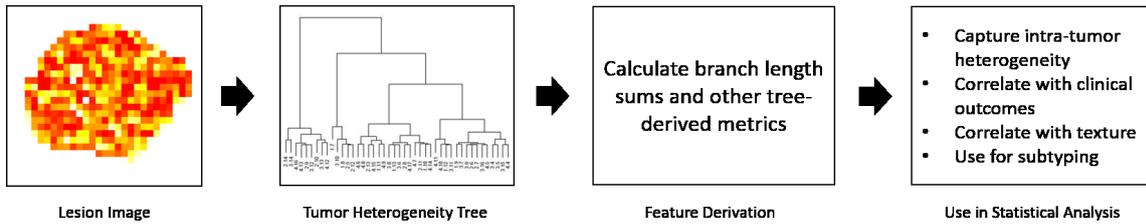


Fig. 1. A pictorial description of the methodology for extracting the tree-based feature metrics from a radiological image: First, a hierarchical tree is built from the image. Then, the branch length sums and other derived metrics are calculated from the tree. These metrics capture ITH and then can be used in further statistical analytics.

2.1. Constructing Trees from pixel-level ROI data

The mathematical objects underlying the first three aspects of Figure 2 are further diagrammed in Equation 1. THT are constructed from radiological images, from which Least-Common Ancestor (LCA) trees are drawn. This sample of LCA trees can then be summarized into metrics to be used as features such as the ones outlined previously, to join these other features in modeling and analysis.

$$\underbrace{\mathcal{I}_i}_{\text{Radiologic Image}} \rightarrow \underbrace{\mathcal{T}_i}_{\text{Tumor Heterogeneity Tree}} \rightarrow \underbrace{c_i(k)}_{\text{LCA Tree}} \rightarrow \underbrace{s_i}_{\substack{\text{Branch length sum,} \\ \text{a tree-based metric}}} \tag{1}$$

To construct the THT, we follow the work of Bharath et al,¹¹ and denote a rooted finite tree with n vertices as τ_n , where τ_n is a point in the space $\mathcal{T}_n \times \mathbb{R}_+^{n-1}$. \mathcal{T}_n is the set of all

finite trees on n vertices. A convenient notation for the tree is $\tau_n = (\mathcal{V}(\tau_n), \mathcal{E}(\tau_n))$, where $\mathcal{V}(\tau_n) = (\text{root}, v_1, \dots, v_{2n-1})$, the set of vertices and $\mathcal{E} = (e_1, \dots, e_{2n-1})$, the set of edges. Note that τ_n denotes a tree with n vertices, including the root and $\tau(n)$ denotes a tree with n terminal vertices.^{11,12} The tree τ_n is not itself a probabilistic structure, so a stochastic process is placed on the growth of the tree in order to build a probabilistic model on the tree-structured data, and further steps are taken to provide a consistent family of densities. A Galton-Watson (GW) process $\{X_n\}_{n \geq 0}$ is a stochastic process that takes on positive integer values in discrete time, often used to model populations. It has an offspring distribution $(\pi_k, k = 0, 1, 2, \dots)$. When this process is conditioned to have n vertices, the resulting tree τ_n is known as a conditioned GW tree.

These conditioned GW trees come from offspring distributions π_k , where k is equal to the number of leaves in the tree. To obtain information about variations in branch structure and to incorporate information about branch lengths, we must move to the Continuum Random Tree (CRT) through weak convergence. The CRT is the asymptotic limit of the GW tree, and in this limit, σ^2 , the variance parameter of the GW tree's offspring distribution appears. Least common ancestor (LCA) trees are randomly chosen binary subtrees of conditioned GW trees, and can be understood as marginals of the CRT and provide dimension reduction. To create an LCA tree from τ_n , choose $k < n$ then uniformly choose k vertices from the n vertices of $\mathcal{V}(\tau_n)$. The density of the family of consistent CRT binary trees $C(k)$ from which these LCA trees with $k < n$ leaves are drawn is shown in equation 2.

$$f_{k, \sigma^2}(c(k)) = \left[\prod_{i=1}^{k-1} \frac{1}{2i-1} \right]^{-1} \frac{1}{2^{k-1}} (\sigma^2)^k s \exp\left(\frac{-s^2 \sigma^2}{2}\right) \quad (2)$$

In summation, properties of the CRT allow us to use this density to approximate the density from which the LCA tree from any conditioned GW tree is drawn. The σ^2 term, gained by taking the CRT of the tree, captures variability in the branching process between different GW trees, while the LCA tree provides the ability to reduce the dimension of the data to s . Thus, for each image's tree \mathcal{T}_i , we create a LCA-tree $c_i(k_i)$ by choosing k_i of the n_i vertices, then calculate the value s_i by taking the sum of the lengths of the branches of LCA-tree. The density above has the kernel of a Gamma distribution with respect to s . Further, we know that s is non-negative, as the branch length components are non-negative, and these CRT branch lengths also asymptotically follow a Gamma distribution in this CRT construction of trees. As the sum of Gamma random variables is also Gamma, this allows for exploration of this feature in a generalized linear model setting. A full reasoning for the choice of the Gamma distribution on the trees can be found in K. Bharath et. al.

The trees produced from the images, as well as informative variables derived from these trees, will be the focus of the analysis in this paper. In practice, for each image, \mathcal{I}_i , hierarchical clustering is done on the vector of pixel densities, v_{ij} , where $j = 1, \dots, n_i$, and n_i is equal to the number of pixels in image \mathcal{I}_i . The agglomerative clustering was done using the UPGMA (average) linkage method¹³ and Euclidean distance to produce a tree, \mathcal{T}_i , from each image. Sensitivity analysis to the selection of distance metrics and clustering methods was performed. From this tree \mathcal{T}_i , a LCA tree $c_i(k)$ is randomly sampled and the branch length sum s from

$c_i(k)$ is calculated.

2.2. Deriving metrics of ITH from tree representations

In order to account for the randomness of the selection of leaves in the LCA trees, we randomly sampled 100-fold from the same image. The median value of the sum of the branch lengths and a measure of the spread of these values were collected as the variables of interest. This process is summarized and depicted in Figure 2, where the multi-modality of the empirical distribution highlights the need to take the median as the measure of center.

It is hypothesized that the edge sum value for each lesion can be a feature that is reflective of the ITH. A group of pixels that are more diverse will produce a tree that is taller; a tree that, for example, clusters somewhat quickly into various groups but then those groups do not merge into one cluster until much later. If an image has a large amount of density values that are similar, those will cluster quickly, leading to short branch lengths. A reflection of this hypothesis can be seen in the left hand column of Figure 3, a graph using images from the case study described below. Tumors with a large amount of similarly valued pixels have low branch length sums, while those that have sharp differences have higher median edge sum values. In fact, the lesion with the highest valued median edge sum has a large group of extremely dense pixels, surrounded by more moderately valued pixels. Trees produced from this lesion have very long branches from the split of the group and non-group pixels, which is reflected in its' very large branch sum value. While some of the difference in visual levels of heterogeneity can be explained by the pixel size of the images, there are differences in the small and large valued groups of the median.

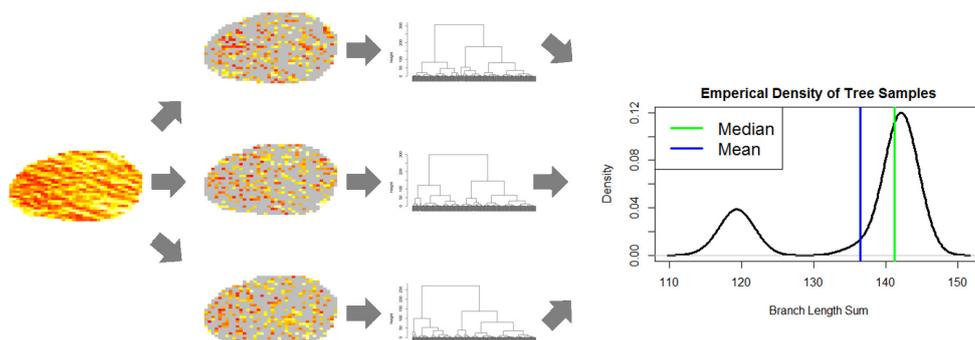


Fig. 2. A depiction of the LCA sampling process. To the left, the original image. Next, a 30% sample of pixels is taken, illustrated by the three images with only 30% of their pixels still in color. From each of these, a hierarchical tree is built, where subtle differences can be seen in the third set of images. The final image to the right is the empirical density plot of the 100 LCA samples taken from this image, and it provides an instance in which taking the median instead of the mean is important.

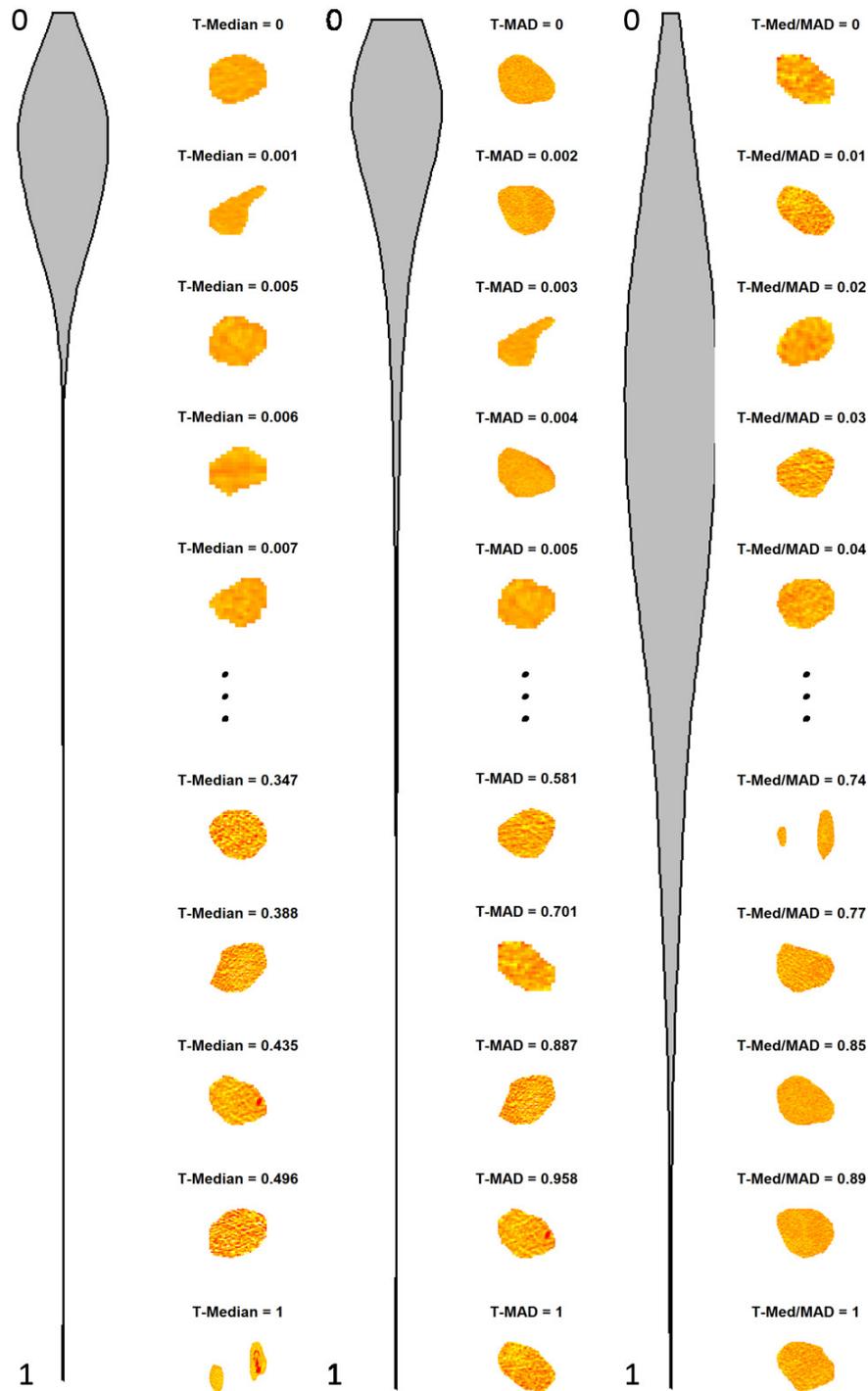


Fig. 3. From left to right, the violin plots show the empirical densities of the Median, Median Absolute Deviation (MAD), and Median/Median Absolute Deviation features for the adrenal lesions used in the case study below, normalized to be between 0 and 1. To the immediate right of the density for each feature are the lesion images for the five highest and lowest values of each, along with their normalized value. The image colors have been scaled so that the individual mean pixel intensities correspond to the same color across all images.

3. Application to Solid Adrenal Lesions

3.1. *Adrenal Lesions*

Adrenal masses are common and can be either functioning or non-functioning. Within each side of this divide, they can be either benign or cancerous. Cancerous lesions can be either first degree, primary, tumors or second degree, metastatic tumors. In patients without known cancer, these masses are often benign adenomas and of little clinical significance, but the proportion that are malignant increases slightly with previous knowledge of cancer¹⁴ as well as with age.¹⁵ A primary non-functioning tumor is a rare malignancy known as an adrenocortical carcinoma (ACC)¹⁶ and a non-functioning benign lesion is an adenoma, which make up approximately 50% of the non-functional lesions.¹⁷ There are many different types of functioning lesions, such as paragangliomas and pheochromocytomas, both of which can present as either benign or malignant.

3.2. *Data*

Our retrospective data consists of 379 CT scans from 356 patients. Of the lesions, 195 are malignant and are 182 benign, 334 are non-functioning lesions and 43 are functioning. For tumors that are metastatic, the information about the primary type of cancer and the timing is available as well. Their pathologies have been verified by the radiologist and are available in table form in the supplementary material, located at http://kas23.web.rice.edu/PSB_supplementary_material.pdf. There are 13 calcified lesions, 13 fatty lesions, 202 heterogeneous lesions and 134 homogeneous ones. The density of lesion size (calculated by pixel count) was heavily skewed to the right, with large lesions presenting as outliers. An unsupervised clustering was performed on the pixel size using k-nearest neighbors in order to produce a distinction between the main group of lesions and the large outlier lesions. To attempt to remove these outlier affects, this cluster of large tumors was not included in calculations based solely on the THT features.

Using the methodology described above, a GW tree was built for each adrenal lesion image using Matlab, averaging 0.98 seconds per tree. Then 100 branch length samples were taken for each of the lesions. This was done with a C++ program accessed through R. It took, on average, 4 seconds to compute one LCA sample of one tree. This was done on a computer with a 3.3 GHz processor and 16 GBs of RAM. The average number of pixels in an image was 2064 and the median number was 812. Several of the lesions' branch sum distributions were multi-modal, see Figure 2 as an example, so it was decided to take the median in place of the mean as the measure of central tendency to account for this. The median absolute deviation (MAD) of the samples was calculated to capture the spread of these samples.

The median branch length sum empirical density from this truncated group of lesions is plotted vertically in the far left violin bar in Figure 3, along with the densities of the MAD of the sample draws and the normalized feature. For all three, the curve is unimodal and varying degrees of right skewed, but the normalized feature on the right does present a smaller tail. As mentioned previously, a visual difference between the upper and lower groups can be seen, particularly in the median and the MAD features (the left and middle columns, respectively).

Note that when time is referenced, it is not in the usual temporal sense, but rather time within the tree similarity space. The group of lesions with small median values are very smooth with similarly valued pixels, leading to trees that go a long distance without branching out, with pixels tending to stay in the cluster instead of breaking apart. The group with large median values have large differences in color, some even with visible sections of pixels that are isolated and a much different value from the rest. Trees for an image such as this are going to have clusters that break apart quickly due to the large variation of intensity values present, leading to tall trees that have large branch length sums. The extreme outlier lesion at the bottom of the column is a perfect example of this, with a large cluster of high-density pixels that cause there to be a fast division into two primary clusters that stay clustered with themselves for a very long time. Lesions with low MAD values appear to be more homogeneous and uniform, likely from the similar pixels causing the sampled trees to be relatively similar as well, i.e., regardless of random sample taken as in Figure 2, the resulting tree is similar. Higher MAD values correspond to lesions with large visible pixel differences; the sampling of pixels from these lesions can make a large difference in the height of the resulting tree. While the normalized feature has less of a visible differential between the high and low image groupings, it contains the information from both other features and has the advantage that it is less correlated with the commonly used radiomics features.

3.3. Connection with Other Radiomics Features

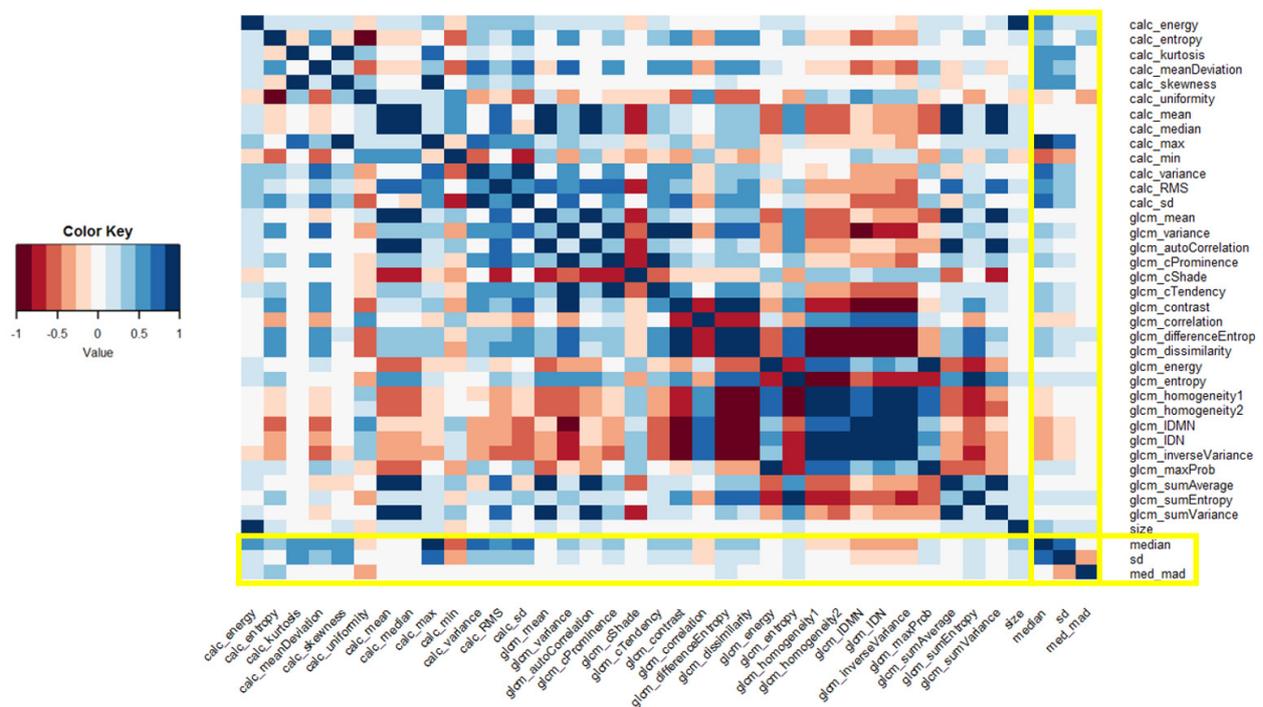


Fig. 4. A heatmap showing the correlation between a set of the NSM and SGLM features with the THT based features (highlighted in yellow).

When the group of small lesions' branch length sums are compared against each of the set of preexisting radiomics features, as in Figure 4, various relationships and lack of relationships can be observed. Less informative features such as minimum and maximum appear dramatically related with the branch lengths, but more expository features such as contrast, which measures local intensity variation, and entropy, which can be used to distinguish tissue with structure, also appear to be correlated with branch length. As can be seen in Figure 4, the median is slightly correlated with the other radiomics features. This correlation led to the decision to divide the median by the calculated MAD, and a normalized feature that was much less correlated with the other radiomics features emerged. This feature is used as the THT feature in further statistical analysis. Besides the THT feature computed for each lesion, the first order NSM features as well as the second order SGLM features were computed. For simplicity and ease of computation, only one GLCM was used for these features, the GLCM with distance 1 and angle of 0° . Thus, for each lesion, there were a total of 37 features collected using a combination of NSM, SGLM, and THT methods.

3.4. *Characterization and Classification*

In order to search for separation caused by the groups of features, a Principal Component Analysis (PCA) was performed on these 37 features in order to construct an orthogonal set of features. The components produced were further used in a 5-fold cross validated logistic regression on three qualitative features of the data set to determine the discriminatory abilities of these features when modeled conjointly. The feature loadings as well as the coefficients of the logistic regression can be found in Section 2 of the supplementary material.

We decided to use the first 6 PCs, as this was sufficient to explain 90% of the variance. The scree plot of the variances can be found as Figure S1 in the supplementary material. When each of these 6 PCs are plotted against each of the other components, the points cluster in a line with the outliers scattered to the side. As can be seen in Figure S2, especially in the 4th and 5th PCs, when compared against the various qualitative information available about the lesions, it was found that the majority of these outlier lesions were denoted by the radiologist as calcified. Calcification is typically distinguishable on CT scans, but doesn't signify one subtype of lesion over any other.¹⁸ This delineation of the calcified lesions is likewise apparent in the results of the logistic regression, presented in Figure 5. The set of PCs has a very high level of accuracy for characterizing calcified lesions from non-calcified lesions (AUC = 1) and functioning from non-functioning (AUC = 0.93) and does moderately well at distinguishing malignant from benign (AUC = 0.78).

At a qualitative level, as can be seen above in Figure 3, a difference between the textural and visual presentation of the high and low value images along the top of the density plot can be seen for the median. This leads to the conclusion that this feature is capturing some aspect of the ITH. When looking at the lesions in the right column of Figure 3, there appears to be a difference in ITH. In general, while the lower valued images have larger particles of density clusters, the ones with higher values have a much finer grain of texture. At a quantitative level, the feature set was able to make a perfect characterization of calcified tumors (AUC of 1) and was highly accurate for determining functioning tumors (AUC of 0.93).

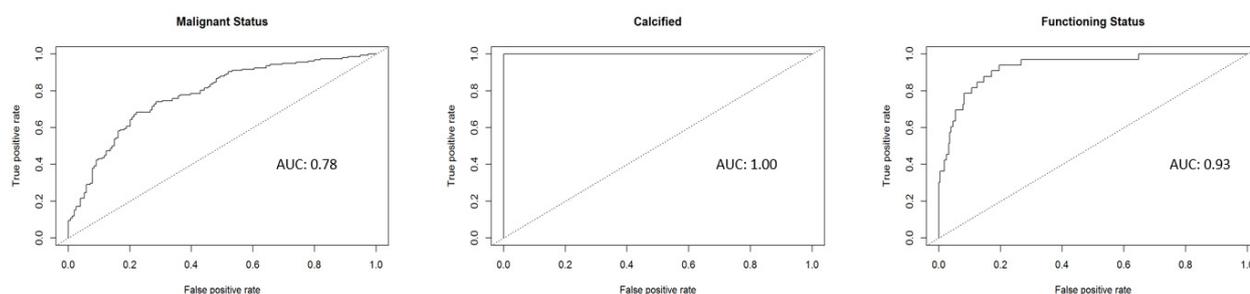


Fig. 5. Receiver Operating Curves for three endpoints chosen for discrimination. Perfect stratification was achieved in the calcified lesions, and near perfection for the functioning status.

4. Discussion

One of the foundational ideas behind radiomics is that analytics can detect nuances in tumors that a human eye might be unable to distinguish. By capturing unique aspects of ITH, as demonstrated through lack of correlation with existing quantitative image features, as well as yielding accurate tissue characterization in integrative analysis, THT-derived features represent a valuable contribution to the parameter domain of radiomics. By their formulation, tree-structured objects offer the potential to better reflect the biological and evolutionary processes that give rise to solid lesions. Summing the branch lengths of GW trees, the feature demonstrated in our case study, and interrogating their distributions under repeat sampling is straightforward and highly interpretable. THTs potentially access much broader information, however, whether that be features from the empirical distribution of the tree samples, the point process inherent to the branch breaking pattern, or using the rationale of the statistical grounding of the known distribution for the branch length sums. More exploration can be done, on this data set or others, to determine the full extent of the THT for producing features for characterization, classification or more.

Translation and dissemination of code is ongoing, available upon request. For further detail about the selection of the random LCA trees, see www.github.com/pkambadu/DyckPaths, where it is available under a BSD-style license

Acknowledgements

This work was supported by NIH R01-CA194391, NIH R01160736 and NSF 1463233 (to VB). KS was partially supported by NIH grant T32 - CA09652.

References

1. L. Alic, W. J. Niessen and J. F. Veenland, *PLOS ONE* **9**, 1 (10 2014).
2. M.-C. Asselin, J. P. OConnor, R. Boellaard, N. A. Thacker and A. Jackson, *European Journal of Cancer* **48**, 447 (2012).

© 2017 The Authors. Open Access chapter published by World Scientific Publishing Company and distributed under the terms of the Creative Commons Attribution Non-Commercial (CC BY-NC) 4.0 License.

3. A. Sottoriva, I. Spiteri, S. G. M. Piccirillo, A. Touloumis, V. P. Collins, J. C. Marioni, C. Curtis, C. Watts and S. Tavar, *Proceedings of the National Academy of Sciences* **110**, 4009 (2013).
4. H. J. Aerts, E. R. Velazquez, R. T. Leijenaar, C. Parmar, P. Grossmann, S. Cavalho, J. Bussink, R. Monshouwer, B. Haibe-Kains, D. Rietveld *et al.*, *Nature communications* **5** (2014).
5. P. Lambin, E. Rios-Velazquez, R. Leijenaar, S. Carvalho, R. G. van Stiphout, P. Granton, C. M. Zegers, R. Gillies, R. Boellard, A. Dekker and H. J. Aerts, *European Journal of Cancer* **48**, 441 (2012).
6. S. S. F. Yip and H. J. W. L. Aerts, *Physics in medicine and biology* **61**, R150 (2016).
7. C. Parmar, E. Rios Velazquez, R. Leijenaar, M. Jermoumi, S. Carvalho, R. H. Mak, S. Mitra, B. U. Shankar, R. Kikinis, B. Haibe-Kains, P. Lambin and H. J. W. L. Aerts, *PLOS ONE* **9**, 1 (07 2014).
8. N. Aggarwal and R. K. Agrawal, *Journal of Signal and Information Processing* **3**, 146 (2012).
9. R. M. Haralick, K. Shanmugam *et al.*, *IEEE Transactions on systems, man, and cybernetics* , 610 (1973).
10. M. Fassnacht, M. Kroiss and B. Allolio, *The Journal of Clinical Endocrinology & Metabolism* **98**, p. 4551 (2013).
11. K. Bharath, P. Kambadur, D. K. Dey, A. Rao and V. Baladandayuthapani, *Journal of the American Statistical Association* (2016).
12. D. Aldous, *Ann. Probab.* **19**, 1 (01 1991).
13. R. R. Sokal and C. D. Michener, *University of Kansas Science Bulletin* **38**, 1409 (1958).
14. J. C. Miller, M. A. Blake and G. W. L. Boland, *BMJ* **338** (2009).
15. G. W. L. Boland, M. A. Blake, P. F. Hahn and W. W. Mayo-Smith, *Radiology* **249**, 756 (2008), PMID: 19011181.
16. E. Duregon, M. Volante, E. Bollito, M. Goia, C. Buttigliero, B. Zaggia, A. Berruti, G. V. Scagliotti and M. Papotti, *Human Pathology* **46**, 1799 (2015).
17. A. B. Grossman, *Nonfunctional Adrenal Masses*. Merck Manual.
18. P. J. Kenney and R. J. Stanley, *Urologic radiology* **9**, 9 (Dec 1988).