

Design, Fabrication, and Implementation of Voxel-Based 3D Printed Heterogeneous  
Lung Lesion Phantoms for the Assessment of the impact of CT Imaging Conditions on  
Texture Quantification

by

Yuese Zheng

Department of Biomedical Engineering  
Duke University

Date: \_\_\_\_\_

Approved:

\_\_\_\_\_  
Ehsan Samei, Supervisor

\_\_\_\_\_  
Fan Yuan

\_\_\_\_\_  
Daniele Marin

Thesis submitted in partial fulfillment of  
the requirements for the degree of Master of Science in the Department of Biomedical  
Engineering in the Graduate School  
of Duke University

2016

ABSTRACT

Design, Fabrication, and Implementation of Voxel-Based 3D Printed Heterogeneous

Lung Lesion Phantoms for Assessment of CT Imaging Conditions on Texture

Quantification

by

Yuese Zheng

Department of Biomedical Engineering  
Duke University

Date: \_\_\_\_\_

Approved:

\_\_\_\_\_  
Ehsan Samei, Supervisor

\_\_\_\_\_  
Fan Yuan

\_\_\_\_\_  
Daniele Marin

An abstract of a thesis submitted in partial  
fulfillment of the requirements for the degree  
of Master of Science in the Department of  
Biomedical Engineering in the Graduate School of  
Duke University

2016

Copyright by  
Yuese Zheng  
2016

## Abstract

Realistic virtual lesion models are valuable in medical imaging applications including phantom design and observer studies. Radiologic diagnostic information rarely include lesion texture due to the fact that texture quantification is sensitive to changing imaging conditions. These effects are not well understood, in part, due to a lack of ground-truth phantoms with realistic textures. Internal tumor heterogeneity in nodules can be predictive of lesion malignancy but is not well understood and virtual lesion models will facilitate research in this area. The purpose of this study was to design and fabricate realistic virtual lung nodules with internal heterogeneity characteristics, and assess the variability as well as determine which imaging conditions provides the most accurate texture features compared to voxel-based 3D printed textured lesions for which the true texture features are known.

We propose a texture synthesis method that accounts for the effects of the imaging system to mimic the appearance of texture in real nodules. Modulation Transfer Function blurring effects and noise contamination was included in the texture generation based on a 3D-Clustered Lumpy Background (3D-CLB). The governing parameters of the 3D-CLB were optimized using a Generic Algorithm with an objective function of Mahalanobis distance between synthesized textures and real lesion textures

features. The resultant texture was objectively and visually similar to real nodules of the same heterogeneity category.

The heterogeneous lesion phantoms were designed with three shapes (spherical, lobulated, spiculated), two textures (homogenous, heterogeneous), and two sizes (diameter < 1.5cm, 1.5cm<diameter<3cm), resulting in 24 lesions (replica of each). The lesions were inserted into an anthropomorphic thorax phantom (Multipurpose Chest Phantom N1, Kyoto Kagaku) and imaged using a commercial CT system (SOMATOM Definition Flash, Siemens Healthcare) at three CTDI levels (0.67, 1.42, 5.80mGy), three reconstruction algorithm (FBP, IR-2, IR-4), four reconstruction kernels (standard, soft, edge), and two slice thicknesses (0.6mm, 5mm). Texture features from these images were extracted and compared to the ground truth feature values. High bias and variance was seen for each feature. Significance level of effects of imaging and lesion conditions were explored. Variability related to CT imaging acquisition and reconstruction techniques is a clinically important source of bias and variance during lesion heterogeneity quantification. Lesion size and shape should also be taken into consideration. The features within a feature category should be considered separately since they were affected by the imaging and lesion conditions differently.

## Dedication

To my parents Jun Qian and Xuefeng Zheng, grandparents Liang Zheng, Bowen Wang, Jianwen Qian, and Huili Qiu for raising and supporting me to chase after my dreams. To my aunt and uncle, Qianshan Ray and Alexander W. Ray, for their relentless effort in helping me through my entire undergraduate and graduate years.

To Lance Corporal Parker T. Boden (USMC) who is deployed with the Wasp Amphibious Ready Group on USS San Antonio (2016-2017) for his love and encouragement from across the Atlantic ocean.

To my primary research advisor Dr. Ehsan Samei, for his patience and guidance not only in research but also in life, and to my secondary research advisor Dr. Daniele Marin, for his constant motivation and suggestion.

# Contents

Abstract .....	iv
List of Tables .....	ix
List of Figures .....	xi
1. Introduction .....	1
1.1 Background .....	1
1.1.1 Lung Nodule Biologic Characteristics.....	2
1.1.2 Lung Nodule Characteristics in Computed Tomography.....	3
1.2 Hypothesis.....	3
1.3 Summary of Chapters .....	4
2. Exploring the Effects of CT Imaging Protocol .....	5
2.1 Homogeneous Texture Based Research .....	5
2.2 Heterogeneous Texture Based Research .....	5
2.3 Pros and Cons of Prior Art.....	6
3. Virtual Realistic Lung Nodule Design .....	7
3.1 Real Patient Data Texture Feature Library Generation.....	8
3.2 Texture Features Selection .....	8
3.3 3D-CLB texture synthesis .....	9
3.4 Imaging physics incorporation into texture .....	11
3.4.1 Modulated Transfer Function Blurring.....	11
3.4.2 Noise Addition .....	12

3.4.3 Effects of MTF and Noise Addition .....	13
3.4.4 Optimization of 3D-CLB parameters.....	14
3.4.4.1 Genetic Algorithm .....	14
3.4.4.2 Mahalanobis Distance .....	17
3.4.4.3 Results of Optimization .....	18
4. Fabrication of Virtual Realistic Lung Nodule .....	21
4.1 Shape Generation .....	21
4.2 Texture Generation.....	22
4.3 Voxel Based 3D Printing.....	23
5. Implementation of Virtual Realistic Lung Nodule.....	25
5.1 Imaging Protocol .....	25
5.2 Texture Feature Quantification .....	26
5.3 Statistical Methods .....	27
6. Effects of Imaging Protocol on Texture Quantification.....	28
6.1 Feature measurements Results .....	28
6.2 Bias.....	31
6.3 Analysis of Variance .....	36
6.4 Discussion.....	40
6.4 Conclusion and Future Work .....	43
Works Cited .....	45

## List of Tables

Table 1: List of texture features calculated from synthetic and real lesions. ....	9
Table 2: Genetic Algorithm parameters.....	15
Table 3: 3D-CLB bounds for Genetic Algorithm .....	15
Table 4: Best fit 3D-CLB parameters at 99 <sup>th</sup> generations in Genetic Algorithm .....	20
Table 5: Correlation between ground truth lesion features .....	29
Table 6: Percent relative bias and standard deviation for each feature between heterogeneous lesion error and homogeneous lesion error.....	30
Table 7: Contrast significance against baseline condition for heterogeneous lesions .....	32
Table 8: Correlation significance against baseline condition for heterogeneous lesions ..	32
Table 9: Dissimilarity significance against baseline condition for heterogeneous lesions	33
Table 10: Energy significance against baseline condition for heterogeneous lesions .....	33
Table 11: Homogeneity significance against baseline condition for heterogeneous lesions .....	34
Table 12: Texture Entropy significance against baseline condition for heterogeneous lesions .....	34
Table 13: Variance significance against baseline condition for heterogeneous lesions.....	35
Table 14: Contrast variance significance for heterogeneous lesions.....	36
Table 15: Correlation significance for heterogeneous lesions.....	37
Table 16: Dissimilarity significance for heterogeneous lesions .....	37
Table 17: Energy significance for heterogeneous lesions .....	38
Table 18: Homogeneity significance for heterogeneous lesions.....	38
Table 19: Texture Entropy significance for heterogeneous lesions .....	39

Table 20: Variance significance for heterogeneous lesions ..... 39

## List of Figures

Figure 1: Demonstration of solid lung nodule (left) and ground glass lung nodule (right). .....	2
Figure 2: Modulated Transfer Function (MTF).....	12
Figure 3: Three dimensional Noise Power Spectrum .....	13
Figure 4: A demonstration of how MTF and NPS affects initial CLB model in the axial direction. Left: initial CLB model generation using the best fit parameter from GA output. Middle: effects of MTF blurring. Right: noise addition onto blurred CLB model. .....	14
Figure 5: Genetic algorithm fitting result. The population median shows the average Mahalanobis distance for each generation, demonstrated by the data points on top. The best populations show the lowest Mahalanobis distance for each generation, demonstrated by the data points below. A log fit was done to each of the dataset separately, showing a linear decline in Mahalanobis distance for both the population median and best population.....	20
Figure 6: A demonstration of the nodule simulation. The image on the left is a real patient nodule. The image in the middle is the nodule simulated with textured background using the in house lesion modeling software. The image on the right is a 3D illustration of the nodule.....	21
Figure 7: Graphic User Interface of the inhouse lesion generation program. ....	22
Figure 8: A comparison between nodules simulated without texture and nodules simulated with texture. The left most image is a real patient nodule. The middle image is a lesion simulated without texture. The right most image is a lesion simulated with texture.....	23
Figure 9: Connex Stratasys 350 3D printing machine .....	24
Figure 10: Demonstration of 3D printed lesions. Left: Cross section of a spiculated lesion. Right: lesions of various shapes and texture.....	24
Figure 11: Absolute error of heterogeneous lesion error and homogeneous lesion error for Energy and Contrast.....	29

Figure 12: Absolute error of heterogeneous lesion error and homogeneous lesion error for Correlation and Variance.....	29
Figure 13: Absolute error of heterogeneous lesion error and homogeneous lesion error for Homogeneity and Texture Entropy. ....	30
Figure 14: Absolute error of heterogeneous lesion error and homogeneous lesion error for Dissimilarity. ....	30
Figure 15: Colored circled plot of bias significance for heterogeneous lesions.....	35
Figure 16: Colored circled plot of variance significance for heterogeneous lesions.....	40

# **1. Introduction**

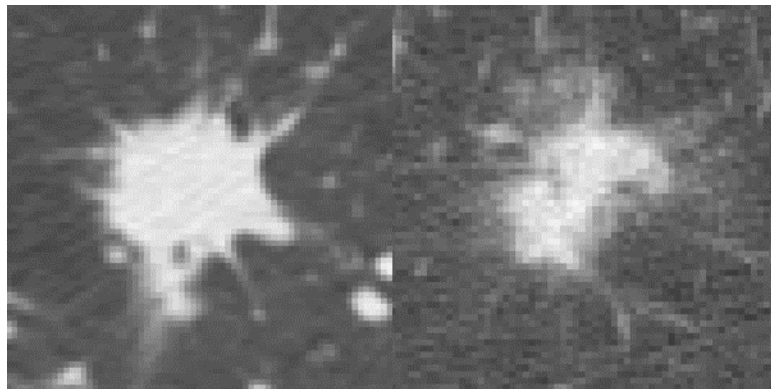
Lung cancer is the leading cause of deaths in cancer related mortalities, surpassing breast cancer related death in both men and women [1]. Lung cancer screening process traditionally relied on chest radiography; however, low-dose CT (LDCT) has demonstrated a 20% reduction in lung cancer mortality and has risen as the primary diagnostic imaging modality [2]. Lung cancer represents itself primarily in lung nodules. Thus, understanding and characterizing lung nodules is a major need in lung cancer research.

## **1.1 Background**

Lung nodules are usually between 5 millimeters to 30 millimeters in size and ones larger than 3 cm are most likely malignant [3]. Many studies have explored the shape and contrast profile characteristics of nodules using lesion phantoms [4-8], including identifying variabilities in shape quantification of phantom lesions under various imaging protocols. The lesion phantoms used in those studies were representative of real lesions in both shape and density but not in internal homogeneity. While internal heterogeneity has been shown to predict response to therapy as well as mortality rate [9], it has not thus far been utilized in clinical practices due to the variations in quantification methods and image settings. There is a need to understand the characteristics of lung nodules in both shape and texture.

### 1.1.1 Lung Nodule Biologic Characteristics

Shape and texture of lung nodules have strong biological bases. Classification of lung cancer is best done through histological analysis of malignant cells under a microscope. A single specimen obtained by local tumor-biopsy, however, would include only a limited sample of an entire tumor [10], which can be more broadly depicted by noninvasive imaging techniques of the integral tumor mass. For example, the two types of lung cancer, small-cell lung cancer (SCLC) and non-small-cell lung cancer (NSCLC), show distinctive appearance under microscopic imaging. On a macroscopic level, distinctive features are also present between solid lung nodules and ground glass nodules, where the former possesses more defined edges while the latter is opaque and embedded with vascular structures (Figure 1).



**Figure 1: Demonstration of solid lung nodule (left) and ground glass lung nodule (right).**

### **1.1.2 Lung Nodule Characteristics in Computed Tomography**

In clinical practice, the attributes of lung nodules are often evident in CT scans. Diagnosis from CT scans are usually based on shape, spiculation, and in plane measurement of the diameter of the nodules. In recent years, with the emergence of radiomics, more features have been explored to relate imaging information to genetic or histopathological data [11]. Internal heterogeneity has been shown to inversely associate with likelihood of metastasis and resistance to therapy. Tumor stage and textural features of heterogeneity derived from CT are shown to predict survival in Non-Small Cell Lung Cancer (NSCLC) [12,13]. However, textural features are not widely used in clinical practice due to the influence of variations in scanning types and protocols.

## **1.2 Hypothesis**

This work hypothesizes that different CT imaging conditions give rise to variabilities in lung lesion texture feature quantification due to intra-scanner conditions (eg, dose level, reconstruction slice thickness, reconstruction kernel, reconstruction algorithm) as well as inter-scanner variability (different manufacturer and model type). It is also hypothesized that certain trends could be identified between the imaging protocol and texture feature values. It might be anticipated that CT settings for the most accurate texture quantification include a high dose level, thin slice thickness, iterative reconstruction with high strength, and sharp reconstruction kernel. The advent of novel voxel based 3D printed heterogeneous lesions with known texture features can provide

a platform to evaluate and mitigate the variability in quantification of lesion features due to the imaging conditions.

### **1.3 Summary of Chapters**

**Chapter 2** describes the pros and cons of prior research performed on evaluating effects of CT imaging protocol. **Chapter 3** describes the modeling process of virtual realistic lung lesions. **Chapter 4** describes the fabrication process of voxel based 3D printed heterogeneous lesions. Information from **Chapter 3** and **Chapter 4** are based on conference proceeding and abstract for Society of Photo-Optical Engineers (SPIE) 2016 Medical Imaging section *Synthesis of realistic lung nodule texture using a clustered lumpy background. Session Number: PS15, Paper Number: 9783-213*. **Chapter 5** describes the implementation of those lesions to assess the effects of CT imaging protocol settings on texture quantifications. **Chapter 6** describes the effects of imaging protocols on texture features. **Chapter 5** and **Chapter 6** are based on conference proceedings and abstracts for Radiological Society of North America (RSNA) *Measurement of Variability in CT Imaging-based Quantification of Tumor Heterogeneity. Paper Number: SSM20-06* and SPIE Medical Imaging section 2017 *Accuracy and variability of texture-based radiomics features of lung lesions across CT imaging conditions. Session Number: PS10, Paper Number: 10132-198*.

## **2. Exploring the Effects of CT Imaging Protocol**

It is important to understand and optimize CT settings, including x-ray sources, detectors, reconstruction algorithms, as well as CT dose index (CTDI), while maintaining image quality. Overall image quality of CT system has been well studied [14-18]; however, prior studies have noted the need to mimic anatomical texture when evaluating CT system performance. Currently, there is no way to assess the impact that different imaging systems have on measuring nodule texture.

### ***2.1 Homogeneous Texture Based Research***

Zhao et al. [19] incorporated nodules of various shapes, sizes, and density to assess the effects of slice thickness and reconstruction kernels on characterization of tumors. The nodule phantoms imaged were homogeneous in texture, and the texture feature quantification was heavily influenced by imaging system characteristics. This work was mainly focused on how imaging setting affect shape and volumetric measurements.

### ***2.2 Heterogeneous Texture Based Research***

Mackin et al. [20] considered texture to assess variations in radiomic features calculation when imaged under routine thoracic imaging protocols from various scanners. A cartridge with various textures similar to soft tissue were scanned under several imaging conditions. However, the texture they presented were not representative of the necrotic nature of lesions, with various internal densities. Also, the

cartridge design were not realistic to lung lesions sizes, which are generally on a scale of 1-3 cm.

### ***2.3 Pros and Cons of Prior Art***

Although prior work provide a wide range of radiomic feature values comparable to human tissue, especially NSCLC tumors, they do not have texture emulating real lesions. Also, the textured phantoms used were not molded into a lesion-like morphology. Tumor internal heterogeneity have yet to be modeled realistically based on its complexity in phenotypic variations. Any study of lung nodule attributes, including lesion texture, would require ground truth, which is often very difficult to obtain. Thus, there is a need to create lesions with realistic texture attributes.

### 3. Virtual Realistic Lung Nodule Design

The purpose of this study was to design and build realistic synthetic lung nodules with realistic internal texture and realistic morphology. Furthermore, as CT blurs and distorts biological attributes of the lesion, the study included a process to take into account the properties of the imaging system. Figure 2 shows the main flow chart of the texture generation process.

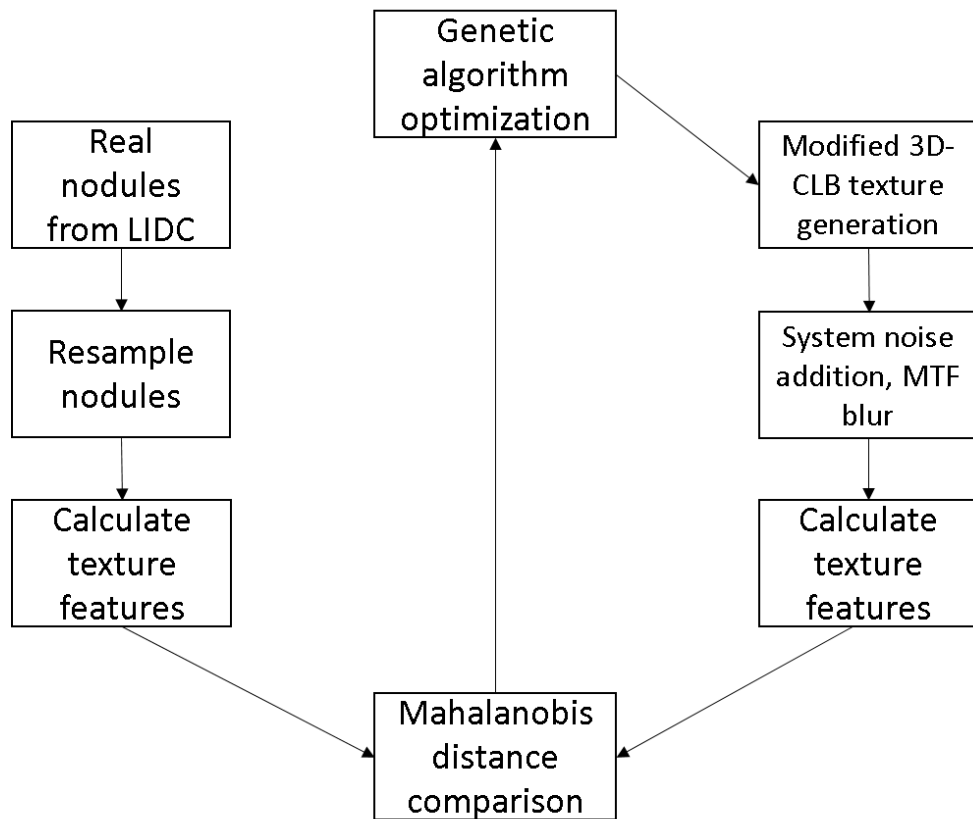


Figure 2: General Flow chart of texture generation.

### **3.1 Real Patient Data Texture Feature Library Generation**

The image data were obtained from the online database of Lung Image Database Consortium (LIDC-IDRI), which contains a large number of CT examinations in patients with radiologists' consented diagnosis on nodule malignancy and manual tracings. Images containing nodules were selected based on one scanner model (GE LightSpeed series), a range of malignancy scores ( $>3$  on a 1-5 scale, 1 being benign and 5 being malignant), and a size range ( $>3$ mm in plane diameter). From 103 patients, 190 nodules were identified and then grouped into three categories based on their internal heterogeneity characteristics on a scale of 1 (homogeneous) to 3 (heterogeneous) by two radiologists in consensus. The radiologists' criteria for determining internal heterogeneity were based on both neighboring pixel and regional intensity variation.

### **3.2 Texture Features Selection**

The nodule images were resampled to have uniform isotropic voxel size,  $0.6 \times 0.6 \times 0.6$  mm for consistency of texture feature measurement. The features in Gray-Level Co-Occurrence Matrix (GLCM) and Gray -Level Run- Length Matrix (GLRLM) were selected and expanded to be calculated in 3D. GLCM features are associated with prediction in radiation treatment [13, 21] and are able to observe intensities at different positions relative to each other; GLRLM features have the ability to discern regional variances [22]. For each nodule, GLCM and GLRLM features [23] shown in Table 1 were measured and grouped according to the nodule's heterogeneity category.

**Table 1: List of texture features calculated from synthetic and real lesions.**

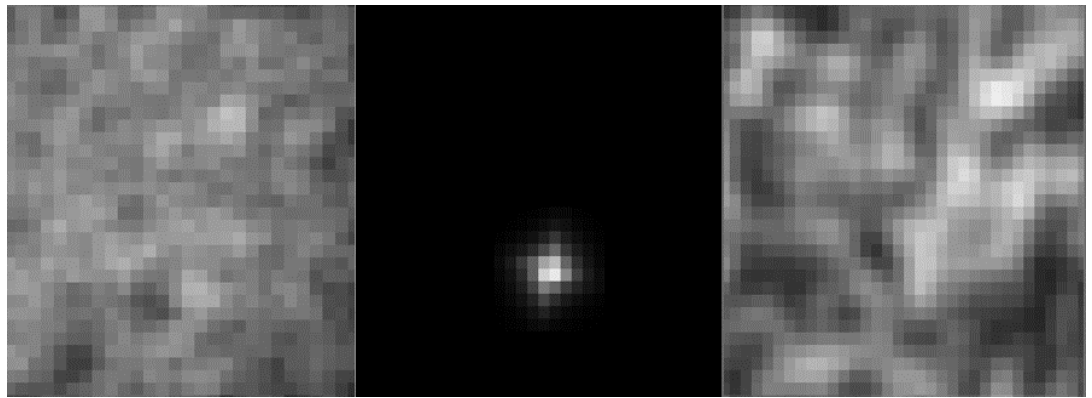
Second Order Statistics	Texture Features
GLCM	Energy, contrast, correlation, homogeneity, variance, sum average, texture entropy, dissimilarity
GLRLM	Short Run Emphasis, Long Run Emphasis, Gray-Level Non-uniformity, Run Length Non-uniformity, Run Percentage, Low Gray-Level Run Emphasis, High Gray-Level Run Emphasis, Short Run Low Gray-Level Emphasis, Short Run High Gray-Level Emphasis, Long Run Low Gray-Level Emphasis, Long Run High Gray-Level Emphasis, Gray-Level Variance, and Run-Length Variance

### 3.3 3D-CLB texture synthesis

Previous works [24-27] suggest various parametric approaches to represent internal texture. Among those methods, the Clustered Lumpy Background (CLB) model has been used successfully to simulate mammogram backgrounds and other anatomical textures [14, 24, 25]. Further statistical descriptions of CLB model have been investigated by Kupinsky et al.[28]. Thus, a 3D Clustered Lumpy Background (CLB) texture synthesis method was chosen as the basis of generating realistic nodule textures. Using the CLB formalism, textures were created by the sum of randomly positioned anisotropic exponential functions as

$$I(\vec{x}) = \sum_i \exp\left[-\alpha \frac{|R_i(\vec{x} - \vec{c}_i)|^\beta}{R_i L}\right],$$

where  $\vec{x}$  denotes a 3D spatial coordinate,  $\vec{c}_i$  is the center of the  $i$ th exponential,  $R_i$  is the rotation matrix of the  $i$ th exponential,  $\vec{L}$  is the length of exponential in each direction, and  $\alpha, \beta$  are coefficients that affect the shape of the exponential. The number of initial seed points and sub-seed points (placed around initial seed points) are governed by  $K_{\text{mean}}$  and  $N_{\text{mean}}$ , which are two Poisson distributed random variables. To compensate for variabilities in the volume of the nodules to be modeled,  $K_{\text{mean}}$  was replaced by  $K_{\text{density}}$ , the number of initial seed points per  $\text{mm}^3$ . The radial distances between sub-seed points and one initial seed point were controlled by  $\sigma$ , and to compensate for 3D generation,  $\sigma$  was extended to  $\sigma_x, \sigma_y, \sigma_z$ , representing the distances in each direction. Thus, the CLB texture was parameterized by  $K_{\text{density}}, N_{\text{mean}}, \sigma_x, \sigma_y, \sigma_z, \alpha, \beta, L_x, L_y,$  and  $L_z$ . Figure 3 is an example of the CLB's ability to generate different textures based on various parameter values.



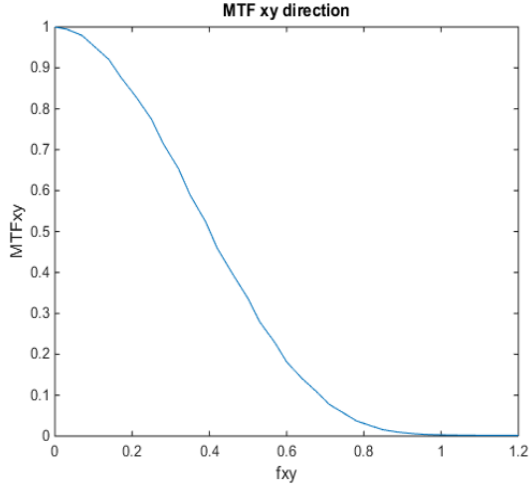
**Figure 3: CLB generations. Left is CLB texture created with lower bound, middle is CLB texture created with upper bound, Right is CLB texture created with parameters of [5 30 10 10 10 10 10 10 10].**

### **3.4 Imaging physics incorporation into texture**

3D-CLB initial texture generation was un-blurred without noise. However, the real lesion texture data were extracted based on scanned patient images, which are affected by the imaging system characteristics. It was necessary to find the parameter values that resulted in texture most similar to the real lung nodules, while attempting to account for the influence of spatial blurring and quantum noise (both magnitude and texture).

#### **3.4.1 Modulated Transfer Function Blurring**

For linear shift invariant systems, all image details are affected equally by blurring and in depth anatomical structure may be lost. The MTF describes how well an imaging system performs in differentiating fine structures with minimal blur [31]. Both in plane (x-y plane) and z-direction resolution was measured in terms of a task transfer function (TTF). A radial edge technique was applied to the inserts in the phantom for measuring in plane resolution, while a supplemental phantom was used for z-direction resolution measurement. The 3D MTF was acquired for the CT scanner (Fig. 2) and assumed to be isotropic in all three directions [29]. The in plane MTF is shown in Figure 4.



**Figure 2: Modulated Transfer Function (MTF)**

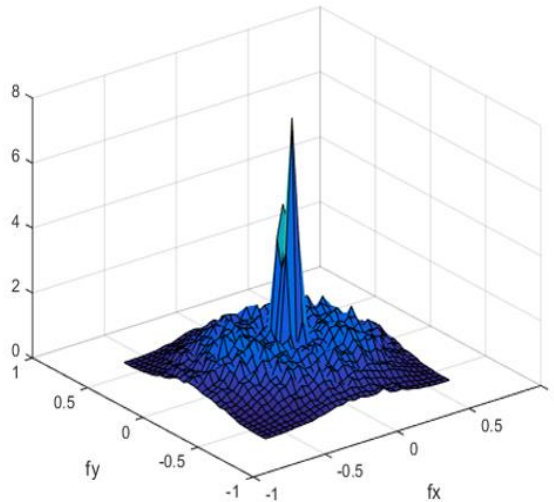
### 3.4.2 Noise Addition

Noise measurement was done in 3D to comply with the nature of the CT dataset based on the NPS. The NPS was determined from Mercury phantom images (v 3.0, Duke University) using a previously published technique by Chen et al [29]. A square ROI of size 64 was chosen in a uniform region in the Mercury phantom. A least squares polynomial fit of second order (interpolation) was applied to the ROI and from which the low frequency background was subtracted. In this case, the NPS was calculated as

$$NPS(u, v, w; N) = \langle \frac{1}{V} |F_{3D} \{I(x, y, z; N) - \overline{I(x, y, z; N)}\}|^2 \rangle$$

where  $I(x,y,z;N)$  is the intensity of the volume at the noise magnitude  $N$ ;  $u$ ,  $v$ , and  $w$  are the orthogonal spatial frequencies; and  $V$  is the volume of interest (VOI) (Figure 5). The acquired NPS was interpolated to match the frequencies in each direction of the image. White Gaussian noise with mean zero and standard deviation of 1 was generated and

filtered by the NPS. The resultant noise was then normalized by its standard deviation and shifted to obtain a mean of zero and added to the 3D-CLB model.

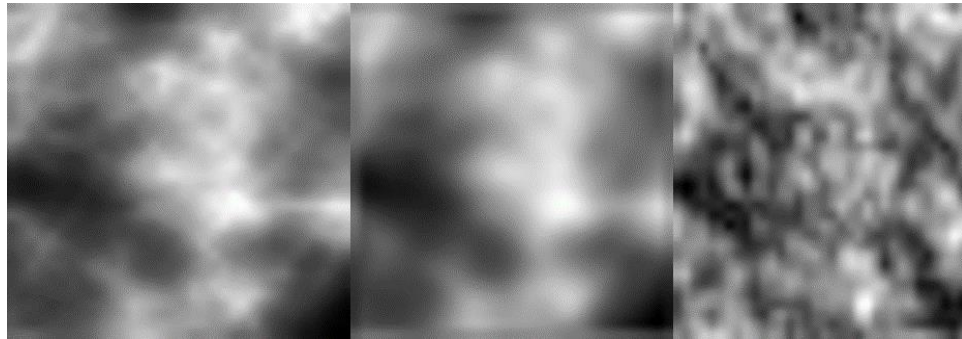


**Figure 3: Three dimensional Noise Power Spectrum**

### **3.4.3 Effects of MTF and Noise Addition**

The nodules simulated using this method represent realistic lung lesion internal heterogeneity both visually and statistically. Figure 6 illustrate the effects of MTF and NPS degradation on the initial input, with the left most image showing the final synthesized image. The original CLB before system degradation maintains fine structures, while the middle image (i.e. after MTF blurring) shows lower resolution and undefined edges. The right most image represents scanner specific noise that is added to the blurred image. From original image to MTF blurring and noise addition, image details are less apparent due to blur and noise while intensity information is preserved even with the addition of noise. Since the noise and blur simulation was performed

under the same imaging condition as that of the actual images, the image features are similar to that of the real image dataset with the same imaging system.



**Figure 4: A demonstration of how MTF and NPS affects initial CLB model in the axial direction. Left: initial CLB model generation using the best fit parameter from GA output. Middle: effects of MTF blurring. Right: noise addition onto blurred CLB model.**

### **3.4.4 Optimization of 3D-CLB parameters**

It is important to optimize the governing parameters of 3D-CLB to generate the most realistic texture similar to real lung lesion textures. Due to the stochastic nature of the 3D-CLB, traditional optimization methods tend to land in local minimums, and are highly dependent on the input bounds. The Genetic Algorithms tackle this problem. In this workm A Genetic Algorithm method was used in combination with the Mahalanobis distance as the objective function.

#### **3.4.4.1 Genetic Algorithm**

An initial set of 3D-CLB governing parameters were fed into the Genetic Algorithm (GA) [32], and bounds were provided for each of the parameters as seen in Table 2 and Table 3. This algorithm attempts to mimic the natural selection process by

(1) randomly generating a population of 3D-CLB parameters and corresponding images (blurred and with quantum noise added), (2) comparing those synthesized images in terms of their texture features to the texture features of real nodules, and (3) randomly mutating the best fitting parameters iteratively over a pre-defined number of generations until the global best-fit parameters are found.

**Table 2: Genetic Algorithm parameters**

Parameter	Definition	Value
L	Number of genes in a chromosome (number of parameters in a set)	10
S	Size of chromosome population (number of sets of parameters)	200
m	Number of times a chromosome was generated (number of times a 3D-CLB volume was generated from the same set)	10
$p_c$	Crossover probability (the probability of a parameter in a set surviving to the next generation)	0.8
N	Number of generations	100

**Table 3: 3D-CLB bounds for Genetic Algorithm**

Bounds	$K_{density}$	$N_{mean}$	$\sigma_x$	$\sigma_y$	$\sigma_z$	$\alpha$	$\beta$	$L_x$	$L_y$	$L_z$
Lower	0.01	0.1	0.1	0.1	0.1	1	1	1	1	1
Upper	0.2	20	5	5	5	3	3	5	5	5
Initial	0.1	10	2	2	2	1	1	2.5	2.5	2.5

Bounds for each parameter were applied to limit the randomness and repetition of the 3D-CLB generation as well as to save execution time for optimization for 100 generations. When no bounds were applied, the random update of parameters created 3D-CLB volumes with high variations between generations, affecting the efficiency of

finding the best fit set of parameters. When no lower bound was applied, the image volume created were generally blank, lacking initial and sub-seeds, not representing realistic texture, as shown in Figure 3 (left). Figure 3 (middle) represents the texture generated with upper bounds in the GA input. Image volumes of more uniform appearances were created due to lack of upper bounds for each parameter, as shown in Figure 3 (right). For example, a high  $K_{\text{density}}$  number means more initial seeds are spread on per  $\text{mm}^3$ . With the addition of sub-seeds ( $N_{\text{mean}}$ ), pixels are aggregated over the entire field of view, creating a denser volume. Based on visual appearances, such bounds were chosen to create realistic 3D image textures in a timely manner.

The bounds as well as the initial set for each 3D-CLB parameter were the inputs to the Genetic Algorithm. There were 10 governing parameters of the 3D-CLB, demonstrated as parameter L in Table 2. For the first Genetic Algorithm generation, a random 200 sets of L were generated and demonstrated as S, which is the number of sets of 3D-CLB parameters. The random sets were generated based on a Roulette selection principle. Each set of parameters went through the 3D-CLB process 10 times, demonstrated as m. Note that due to the stochastic nature of the texture generation algorithm, textures with different visual properties may be generated even with the same set of parameters. The m number of repetitions per set aimed to reduce the random nature of the texture algorithm. Each time the a 3D-CLB was generated with the repeating governing parameter, GLCM and GLRLM texture features were calculated,

creating a  $10 \times 21$  texture feature matrix, where there was 21 texture features. The comparison between the real patient nodule texture feature data and the synthesized texture feature data were compared using a Mahalanobis Distance, which will be discussed in detail in the next section. Once the distance was found, the algorithm continues to the next set of parameter in the first generation. Eventually, the Genetic Algorithm goes through all generations (100 number of generations in total, demonstrated as N), and the process of 3D-CLB generation and texture feature calculation was repeated.

The underlying principles of updating the 3D-CLB governing parameters between generations is based on biological selection, crossover and mutation. The crossover process is randomly switching chromosomes (one governing parameter) from one set to another, while the mutation process is changing a chromosome completely. The crossover and mutation probability was set to Matlab Genetic Algorithm Optimization default to reduce human interference. The fitness of children (which set of parameters provides the least Mahalanobis distance) were determined in each generation, and the best fit parameter survives to the next generation.

#### **3.4.4.2 Mahalanobis Distance**

To compensate for the randomness of the 3D-CLB model (as the same parameters may generate different texture feature values for repeated image synthesis runs), 10 samples of 3D-CLB were synthesized for each set of parameters in one

generation of the GA. The texture features from the CLB synthesized images were grouped into a p by q matrix,  $\vec{Y}$ , where p represents the number of samples, 10, and q represents the number of features, 21, in this case.

The synthesized and real textures features were compared using the Mahalanobis distance,  $d = (\vec{Y} - \vec{\mu}) \vec{S}^{-1}(\vec{Y} - \vec{\mu})'$ , between the texture features of the synthesized images,  $\vec{Y}$ , and the library of features from real nodules,  $\vec{X}$ . Here  $\vec{\mu}$  and  $\vec{S}$  are the means and covariances of  $\vec{X}$ .  $\vec{X}$  is an m by n matrix, where m is the number of real nodules and n is the number of features, in this case, 21. The number of nodules (m) were 5 times greater than the number of samples generated from 3D-CLB (p), ensuring the proper use of Mahalanobis distance. The GA aims to minimize the Mahalanobis distance by iteratively updating the 3D-CLB parameters until the best fits were found, as seen in the general flowchart in Figure 2. The stopping criteria for the GA was reaching 100 generations (for computational efficiency) or the Mahalanobis distance reaches an oscillating steady state, whichever comes first.

#### **3.4.4.3 Results of Optimization**

The genetic algorithm showed a decreasing trend of objective function value with increasing number of generations (Figure 7). The Mahalanobis distance was decreased by 30%, and 13% compared to the fitness function value of the first GA

generation for heterogeneity levels 2 and 3, respectively, after approximately 30 generations suggesting that the synthesized textures were objectively similar to those from real nodules. The best fit parameters at generation 99 are shown in Table 4. Overall, the population median expressed a steady decrease in Mahalanobis distance, while the best population first demonstrated rapid decline in the first 20 generations and then expressed oscillating behavior while remaining under the initial value. This signifies that the system has reached minimum and the texture features calculated from 3D-CLB generation governed by the input parameters are approaching the real lesion texture features. The oscillating behavior could be explained by the random nature of  $m$  realizations of 3D-CLB in each generation that were computed to evaluate the fitness function. Image texture and texture features are not of a one-to-one relationship, meaning images with same texture features could have different visual representations, and vice versa. The 21 texture features used for comparison between real and synthetic lesions can evaluate the variations more precisely, even when the images have similar first-order statistics.

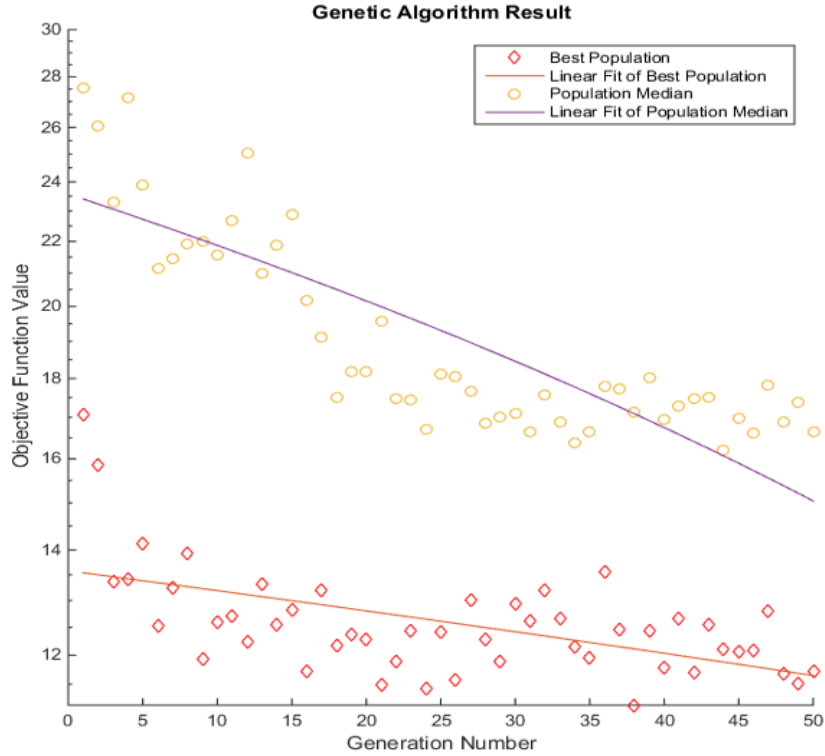


Figure 5: Genetic algorithm fitting result. The population median shows the average Mahalanobis distance for each generation, demonstrated by the data points on top. The best populations show the lowest Mahalanobis distance for each generation, demonstrated by the data points below. A log fit was done to each of the dataset separately, showing a linear decline in Mahalanobis distance for both the population median and best population.

Table 4: Best fit 3D-CLB parameters at 99<sup>th</sup> generations in Genetic Algorithm

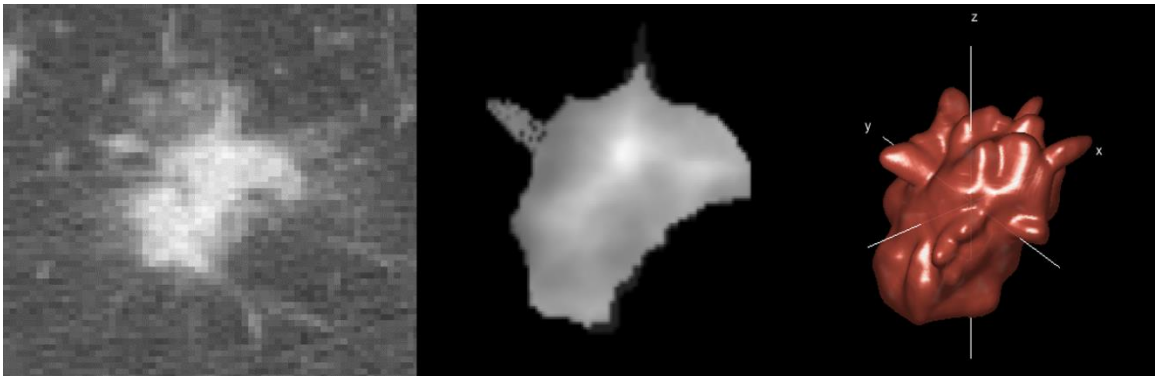
Lesion Category	$K_{density}$	$N_{mean}$	Sigma ( $\sigma_x$ )	Sigma ( $\sigma_y$ )	Sigma ( $\sigma_z$ )	Alpha ( $\alpha$ )	Beta ( $\beta$ )	Lx	Ly	Lz
Category 2	0.0091562	14	4	3	5	1.1668	1.068	5	4	1
Category 3	0.0099007	18	3	4	4	1.0471	1.0287	4	1	3

## 4. Fabrication of Virtual Realistic Lung Nodule

The successful design of lesion texture enables generating realistic lung lesions. Both shape and texture were taken into consideration during the design process to mimic the biological representation of lesions during imaging conditions. A voxel based 3D printing method was further utilized for fabricating phantom lesions.

### 4.1 Shape Generation

The shape of the nodules were modeled based on an in house program (Duke Lesion Tool) which was utilized to randomly generate lesions of various shapes, based on the shapes of real lobulated nodules [8]. Spiculated models were generated by creating polygon mesh models based on segmented real nodules. Lobulated models were designed with a spokes model, where the 3D radii following a random Poisson distribution.



**Figure 6: A demonstration of the nodule simulation. The image on the left is a real patient nodule. The image in the middle is the nodule simulated with textured background using the in house lesion modeling software. The image on the right is a 3D illustration of the nodule.**

## 4.2 Texture Generation

Texture was added based on the best fit 3D-CLB parameters from Genetic Algorithm result. An example of the nodule modeling process is shown in Figure 6. The governing parameters of 3D-CLB from the previous optimization process were inputted into the texture generation tab. The designed virtual nodule with shape and texture were inserted into real patient images with an alpha blending technique [14].

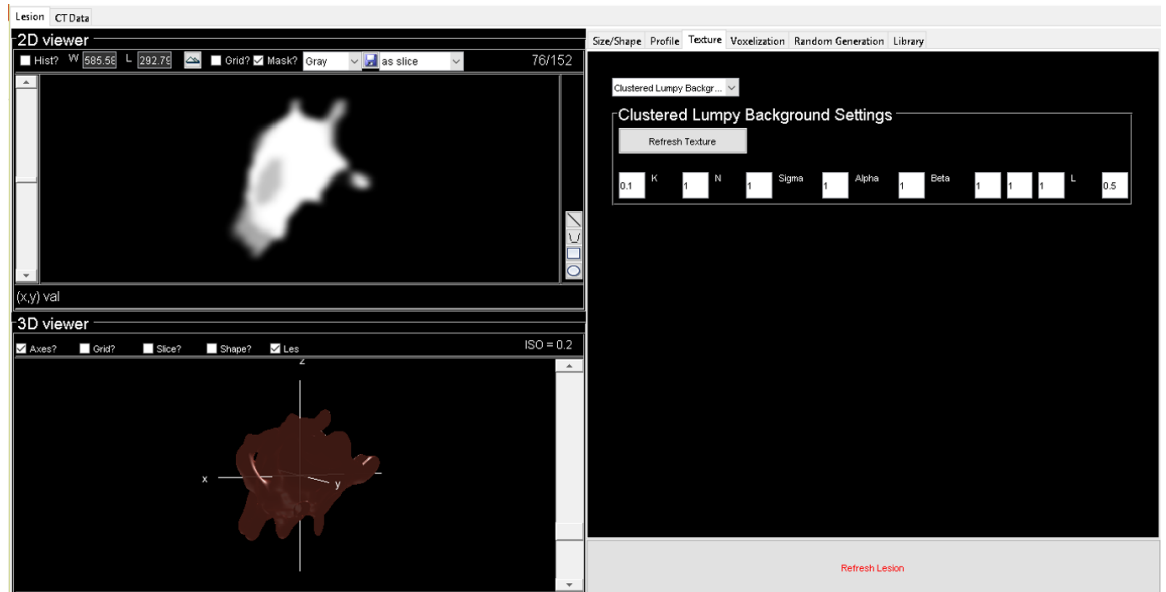


Figure 7: Graphic User Interface of the inhouse lesion generation program.



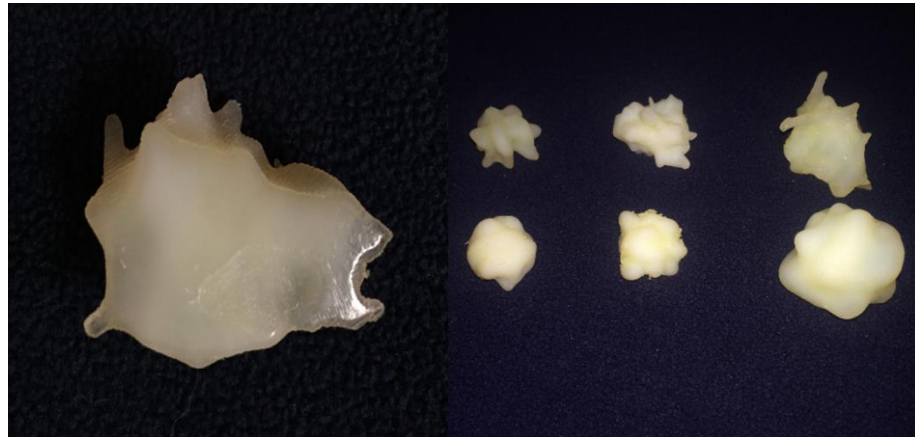
**Figure 8: A comparison between nodules simulated without texture and nodules simulated with texture. The left most image is a real patient nodule. The middle image is a lesion simulated without texture. The right most image is a lesion simulated with texture.**

### ***4.3 Voxel Based 3D Printing***

The finalized lesion model was processed and input into a multi-material 3D printer (Stratesys Connex 350). The lesion model was first voxelized to 0.042x0.084x0.03mm with double precision intensity values. Values of 0 (100% material A) to 1 (100% material B) were assigned to each voxel to represent the proportions of the two base materials used. This fine resolution is much smaller than the typical resolution of a clinical CT system ensuring the macroscopic resolution achieved, thus classic Floyd-Steinberg dithering method was applied to distribute two base materials spatially [33]. The two photopolymer base materials were TangoPlus (75 HU at 120 kVp) and VeroWhite (115 HU at 120kVp) representative of the attenuation of soft tissue and lesions. Figure 11 and Figure 12 represent the 3D printing machine used and the actual lesions printed.



**Figure 9: Connex Stratasys 350 3D printing machine**



**Figure 10: Demonstration of 3D printed lesions. Left: Cross section of a spiculated lesion. Right: lesions of various shapes and texture.**

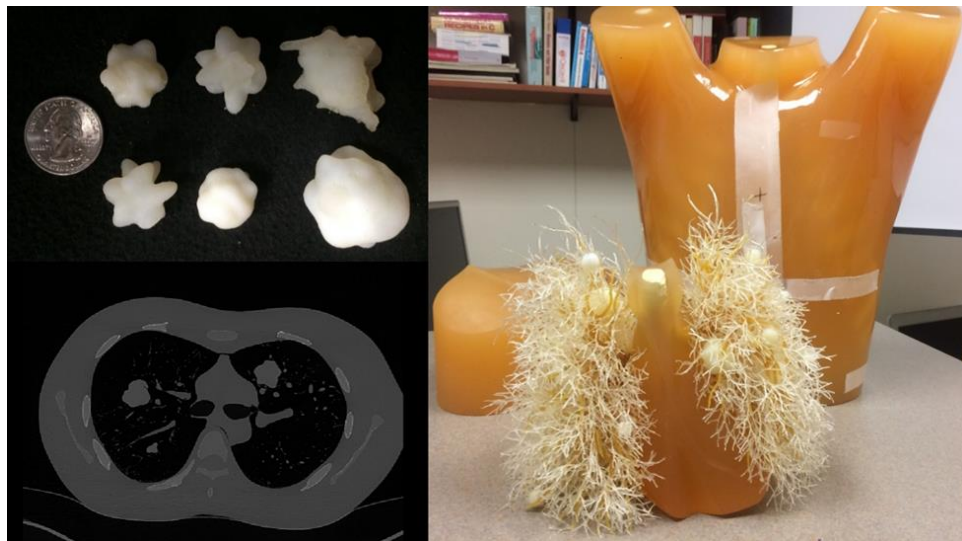
## **5. Implementation of Virtual Realistic Lung Nodule**

The realistic textured lung phantoms with known ground truth (i.e., the lesion texture known based on the initial virtual design) enables the assessment of the impact of imaging condition on texture quantification. This impact can be characterized in terms of variability, but also bias, enabling calibration of measured values to reference conditions. In our methodology, the phantoms were inserted into realistic lung environment of a commercially available lung phantom. A comprehensive set of imaging conditions were applied to assess lesion texture quantification variability and accuracy.

### ***5.1 Imaging Protocol***

The virtual lung lesions were physically fabricated using a multi-material 3D printer (Objet Connex3, Stratasys) programmed to spatially distribute different materials on a voxel-by-voxel basis ( $42 \times 84 \times 30 \mu\text{m}^3$  nominal voxel size). Lung lesions varied in their shape (spherical, lobulated, and spiculated), texture (homogeneous, heterogeneous), and size (in plane diameter  $< 1.5\text{cm}$ ,  $1.5\text{cm} < \text{in plane diameter} < 3\text{cm}$ ). They were then embedded into an anthropomorphic chest phantom (Multipurpose Chest Phantom N1, Kyoto Kagaku). The phantom was scanned on a clinical CT system (Definition Flash, Siemens) at three radiation dose levels ( $\text{CTDI}_{\text{Vol}} = 0.67, 1.42, 5.80 \text{ mGy}$ ). Images were reconstructed at 0.6 and 5.0 mm section thickness using FBP and a range of clinically-applicable iterative reconstruction (IR low strength, IR high strength)

algorithms and kernels (standard, edge, soft). The GE setting for IR was AsirV 40 and 60, and the Siemens setting for IR was Asir 2 and 4. The correspondent Siemens kernels that were comparable to GE kernels were B31, B26, and B70, for standard, edge, and soft respectively. The Siemens IR kernels were I31, I26, and I70, for standard, edge, and soft respectively.



**Figure 13: A demonstration of between and after scanning. Top left: lesion sizes when compared to a coin. Right: lesions embedded into the lung phantom. Bottom left: CT images after phantom scanning.**

## **5.2 Texture Feature Quantification**

Segmentation of quantitative imaging features was performed using a semi-automated lesion segmentation program (Seg3D). The segmented lesion masks were eroded and dilated to avoid the undefined edge of lesions, ensuring the texture features calculated were a good representation of the actual lesions themselves. Texture features derived from gray level co-occurrence matrix were measured (energy, contrast,

correlation, homogeneity, variance, entropy, and dissimilarity). Corresponding texture features were extracted from the virtual representations of the lesions to serve as ground-truth.

### **5.3 Statistical Methods**

Texture features Energy, Contrast, Homogeneity, Texture Entropy, Correlation, Variance, and Dissimilarity error were compared homogeneous and heterogeneous lesions. The absolute error (imaged lesion features – ground truth features) was used to determine the bias in feature measurements. Percent relative bias  $((\text{imaged lesion features} - \text{ground truth features}) / \text{ground truth features})$  and standard deviation was also performed in order to compare all the features on the same scale.

Significance of Bias and variance determine whether the imaging conditions and lesions conditions have significant effects on feature measurements. The baseline of the model is Siemens Somatom Flash scanner, slice thickness of 5mm, standard kernel, reconstruction with FBP, CTDI of 5.8, lesion size of large, and lesion shape of spherical. Other conditions were compared to the baseline, and the effects, standard error, and p-value for significance was measured for bias. For the analysis of variance, the variations within each imaging and lesion condition were compared.

## **6. Effects of Imaging Protocol on Texture Quantification**

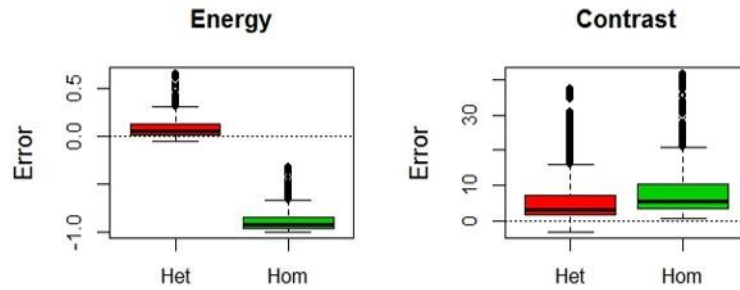
Bias was observed between homogeneous lesions features absolute errors and that of heterogeneous features, especially for 4 features (Energy, Homogeneity, Correlation, Dissimilarity). Most features have high percent relative bias and standard deviation. In the bias and variance between imaging and lesion conditions, a significant difference was observed.

### **6.1 Feature measurements Results**

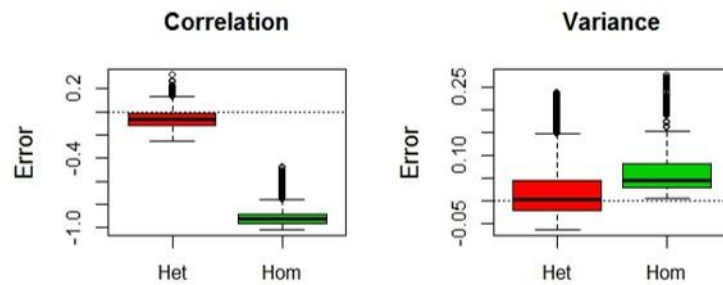
Correlation between ground truth lesions were demonstrated in Table 5. The heterogeneous lesion errors were plotted in red and the homogeneous lesion errors were plotted in green from Figure 11 – 14. Error bars were indicated for each dataset. Percent relative bias and standard deviation for each feature between heterogeneous lesion error and homogeneous lesion error was shown in Table 5. Note that relative bias and standard deviation cannot be calculated for some features in homogeneous lesions because the truth value is 0. The range of data for each boxplot is represented by the horizontal distance from the lowest value to the highest value, including outliers. The median of the data was represented by the solid horizontal line in the box. The lower bound of the box is the first quantile while the upper bound of the box is the third quantile, with the median in between.

**Table 5: Correlation between ground truth lesion features**

	Energy	Contrast	Correlation	Homogeneity	Variance	Texture Entropy	Dissimilarity
Energy	1	-0.7	1	0.94	-0.93	-1	-0.99
Contrast	-0.7	1	-0.74	-0.88	0.87	0.71	0.77
Correlation	1	-0.74	1	0.96	-0.95	-1	-1
Homogeneity	0.94	-0.88	0.96	1	-0.99	-0.95	-0.97
Variance	-0.93	0.87	-0.95	-0.99	1	0.94	0.97
SumAverage	-1	0.71	-1	-0.95	0.94	1	0.99
Texture Entropy	-0.99	0.77	-1	-0.97	0.97	0.99	1



**Figure 11: Absolute error of heterogeneous lesion error and homogeneous lesion error for Energy and Contrast.**



**Figure 12: Absolute error of heterogeneous lesion error and homogeneous lesion error for Correlation and Variance.**

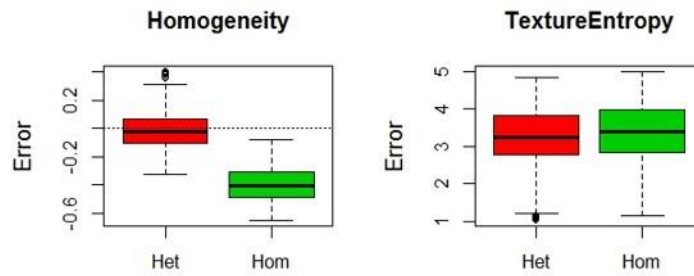


Figure 13: Absolute error of heterogeneous lesion error and homogeneous lesion error for Homogeneity and Texture Entropy.

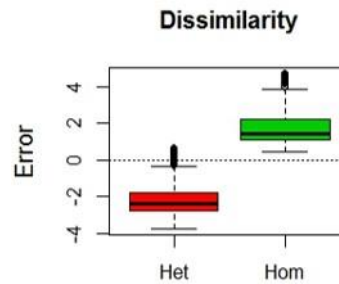


Figure 14: Absolute error of heterogeneous lesion error and homogeneous lesion error for Dissimilarity.

Table 6: Percent relative bias and standard deviation for each feature between heterogeneous lesion error and homogeneous lesion error

Feature	Heterogeneous		Homogeneous	
	Rel. Bias	Rel. SD	Rel. Bias	Rel. SD
Energy	295.85	374.14	-88.37	10.88
Contrast	472.04	508.25	Inf	NA
Correlation	-37.25	48.81	-91.24	6.74
Homogeneity	-1.86	19.41	-39.96	11.87
Variance	45.04	119.09	Inf	NA
Texture Entropy	5143	1108.03	Inf	NA
Dissimilarity	-57.8	21	Inf	NA

## 6.2 Bias

Table 6 – 12 shows the significance against baseline condition for heterogeneous lesions. The baseline of the model is Siemens Somatom Flash scanner, slice thickness of 5mm, standard kernel, reconstruction with FBP, CTDI of 5.8, lesion size of large, and lesion shape of spherical. In the table, kernels were labeled as the GE kernel for standardization, and each kernel has a Siemens Somatom Flash kernel correspondant, as stated previously in the imaging protocol section. The iterative reconstruction algorithm was also standardized to with low and high strength, while the setting was different but comparable between GE and Siemens scanners. Lesion sizes were labeled as 1 (small) and 2 (large). Lesion Shape was labeled as 1 (spherical), 2 (lobulated), and 3 (spiculated). It is imperative to explore the effects of imaging conditions as well as lesions conditions of textured lesion phantoms, thus the homogeneous lesions were ignored for this analysis. The significance level was indicated by the  $\Pr(>|t|)$  column, which is the p-value. Values of 0.0000 means very high significance where the p value is close to 0. The number of significant digits allowed in the table omitted the non-zero digits for the p-value. In Figure15, bias significance for heterogeneous lesions were shown, with red circle indicating very highly significant ( $p < 0.001$ ), yellow circle indicating highly significant ( $p < 0.01$ ), green circle indicating significant ( $p < 0.05$ ), and gray circle indicating not significant ( $p > 0.05$ ).

**Table 7: Contrast significance against baseline condition for heterogeneous lesions**

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	10.8592	0.2696	40.28	0.0000
scannerRevolution CT	-1.9760	0.1537	-12.86	0.0000
slicethickness0.6	-7.8497	0.1531	-51.27	0.0000
kernelEDGE	-1.3840	0.1844	-7.50	0.0000
kernelSOFT	0.0603	0.1846	0.33	0.7438
ReconAlgorithmIR_High	-0.3464	0.1798	-1.93	0.0541
ReconAlgorithmIR_Low	-0.5277	0.1925	-2.74	0.0061
ctdi	0.0627	0.0331	1.90	0.0580
lesionsize1	1.5332	0.1523	10.06	0.0000
factor(LesionShape)2	0.5925	0.1834	3.23	0.0012
factor(LesionShape)3	0.0454	0.1834	0.25	0.8046

**Table 8: Correlation significance against baseline condition for heterogeneous lesions**

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-0.1047	0.0031	-33.52	0.0000
scannerRevolution CT	-0.0090	0.0018	-5.05	0.0000
slicethickness0.6	0.0471	0.0018	26.56	0.0000
kernelEDGE	-0.0210	0.0021	-9.82	0.0000
kernelSOFT	-0.0040	0.0021	-1.86	0.0634
ReconAlgorithmIR_High	0.0110	0.0021	5.30	0.0000
ReconAlgorithmIR_Low	0.0035	0.0022	1.58	0.1152
ctdi	0.0008	0.0004	2.01	0.0450
lesionsize1	0.0689	0.0018	39.05	0.0000
factor(LesionShape)2	0.0070	0.0021	3.31	0.0009
factor(LesionShape)3	-0.0386	0.0021	-18.15	0.0000

**Table 9: Dissimilarity significance against baseline condition for heterogeneous lesions**

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-1.6093	0.0313	-51.43	0.0000
scannerRevolution CT	-0.2190	0.0178	-12.27	0.0000
slicethickness0.6	-0.8530	0.0178	-47.99	0.0000
kernelEDGE	0.1314	0.0214	6.14	0.0000
kernelSOFT	-0.0049	0.0214	-0.23	0.8175
ReconAlgorithmIR_High	-0.1483	0.0209	-7.10	0.0000
ReconAlgorithmIR_Low	-0.1050	0.0223	-4.70	0.0000
ctdi	-0.0137	0.0038	-3.57	0.0004
lesionsize1	-0.2695	0.0177	-15.24	0.0000
factor(LesionShape)2	0.0197	0.0213	0.93	0.3544
factor(LesionShape)3	0.3675	0.0213	17.26	0.0000

**Table 10: Energy significance against baseline condition for heterogeneous lesions**

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.1508	0.0039	38.62	0.0000
scannerRevolution CT	0.0232	0.0022	10.41	0.0000
slicethickness0.6	0.0143	0.0022	6.46	0.0000
kernelEDGE	-0.0868	0.0027	-32.47	0.0000
kernelSOFT	-0.0044	0.0027	-1.63	0.1038
ReconAlgorithmIR_High	0.0462	0.0026	17.72	0.0000
ReconAlgorithmIR_Low	0.0232	0.0028	8.33	0.0000
ctdi	0.0105	0.0005	21.89	0.0000
lesionsize1	-0.0274	0.0022	-12.40	0.0000
factor(LesionShape)2	-0.0256	0.0027	-9.62	0.0000
factor(LesionShape)3	-0.0627	0.0027	-23.61	0.0000

**Table 11: Homogeneity significance against baseline condition for heterogeneous lesions**

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-0.1023	0.0040	-25.33	0.0000
scannerRevolution CT	0.0298	0.0023	12.95	0.0000
slicethickness0.6	0.0639	0.0023	27.85	0.0000
kernelEDGE	-0.0937	0.0028	-33.91	0.0000
kernelSOFT	0.0056	0.0028	2.03	0.0421
ReconAlgorithmIR_High	0.0461	0.0027	17.12	0.0000
ReconAlgorithmIR_Low	0.0240	0.0029	8.32	0.0000
ctdi	0.0081	0.0005	16.35	0.0000
lesionsize1	0.1102	0.0023	48.31	0.0000
factor(LesionShape)2	0.0238	0.0027	8.65	0.0000
factor(LesionShape)3	0.0010	0.0027	0.38	0.7044

**Table 12: Texture Entropy significance against baseline condition for heterogeneous lesions**

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	2.8312	0.0249	113.53	0.0000
scannerRevolution CT	-0.0282	0.0142	-1.98	0.0475
slicethickness0.6	-0.4245	0.0142	-29.97	0.0000
kernelEDGE	0.4346	0.0171	25.47	0.0000
kernelSOFT	-0.0017	0.0171	-0.10	0.9196
ReconAlgorithmIR_High	-0.3075	0.0166	-18.49	0.0000
ReconAlgorithmIR_Low	-0.1715	0.0178	-9.63	0.0000
ctdi	-0.0527	0.0031	-17.22	0.0000
lesionsize1	0.4688	0.0141	33.27	0.0000
factor(LesionShape)2	0.3071	0.0170	18.09	0.0000
factor(LesionShape)3	0.5362	0.0170	31.60	0.0000

**Table 13: Variance significance against baseline condition for heterogeneous lesions**

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.0337	0.0023	14.72	0.0000
scannerRevolution CT	0.0399	0.0013	30.56	0.0000
slicethickness0.6	-0.0272	0.0013	-20.96	0.0000
kernelEDGE	-0.0248	0.0016	-15.86	0.0000
kernelSOFT	-0.0002	0.0016	-0.16	0.8736
ReconAlgorithmIR_High	-0.0008	0.0015	-0.54	0.5924
ReconAlgorithmIR_Low	-0.0027	0.0016	-1.64	0.1004
ctdi	0.0012	0.0003	4.25	0.0000
lesionsize1	-0.0117	0.0013	-9.02	0.0000
factor(LesionShape)2	-0.0000	0.0016	-0.02	0.9818
factor(LesionShape)3	-0.0060	0.0016	-3.85	0.0001



**Figure 15: Colored circled plot of bias significance for heterogeneous lesions**

### 6.3 Analysis of Variance

Table 13 – 19 shows the variance for each feature error for heterogeneous lesions. The significance value is indicated by Pr(>F) column, and values of 0.0000 means very high significance. The non-zero digits were again omitted due to the number of significant digits. The various combinations of imaging conditions and lesion conditions were compared within themselves. In Figure 16, bias significance for heterogeneous lesions were shown, with red circle indicating very highly significant ( $p < 0.001$ ), yellow circle indicating highly significant ( $p < 0.01$ ), green circle indicating significant ( $p < 0.05$ ), and gray circle indicating not significant ( $p > 0.05$ ).

**Table 14: Contrast variance significance for heterogeneous lesions**

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
scanner	1	6260.09	6260.09	193.64	0.0000
slicethickness	1	86833.26	86833.26	2686.00	0.0000
kernel	2	2507.36	1253.68	38.78	0.0000
ReconAlgorithm	2	271.11	135.55	4.19	0.0151
ctdi	1	118.53	118.53	3.67	0.0556
lesionsize	1	17418.71	17418.71	538.81	0.0000
factor(LesionShape)	2	2801.14	1400.57	43.32	0.0000
Residuals	5761	186241.82	32.33		

**Table 15: Correlation significance for heterogeneous lesions**

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
scanner	1	0.09	0.09	25.70	0.0000
slicethickness	1	3.18	3.18	893.19	0.0000
kernel	2	0.47	0.24	66.27	0.0000
ReconAlgorithm	2	0.13	0.06	17.89	0.0000
ctdi	1	0.02	0.02	5.00	0.0254
lesionsize	1	1.08	1.08	302.56	0.0000
factor(LesionShape)	2	0.18	0.09	24.76	0.0000
Residuals	5761	20.51	0.00		

**Table 16: Dissimilarity significance for heterogeneous lesions**

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
scanner	1	74.41	74.41	172.83	0.0000
slicethickness	1	1019.26	1019.26	2367.36	0.0000
kernel	2	22.73	11.37	26.40	0.0000
ReconAlgorithm	2	23.70	11.85	27.52	0.0000
ctdi	1	5.65	5.65	13.12	0.0003
lesionsize	1	298.16	298.16	692.52	0.0000
factor(LesionShape)	2	109.77	54.89	127.48	0.0000
Residuals	5761	2480.37	0.43		

**Table 17: Energy significance for heterogeneous lesions**

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
scanner	1	0.65	0.65	95.69	0.0000
slicethickness	1	0.23	0.23	33.10	0.0000
kernel	2	9.07	4.53	664.47	0.0000
ReconAlgorithm	2	2.16	1.08	158.39	0.0000
ctdi	1	3.32	3.32	486.07	0.0000
lesionsize	1	3.52	3.52	515.88	0.0000
factor(LesionShape)	2	3.25	1.62	238.02	0.0000
Residuals	5761	39.30	0.01		

**Table 18: Homogeneity significance for heterogeneous lesions**

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
scanner	1	1.22	1.22	181.58	0.0000
slicethickness	1	5.53	5.53	824.66	0.0000
kernel	2	11.75	5.87	876.46	0.0000
ReconAlgorithm	2	2.16	1.08	161.30	0.0000
ctdi	1	1.98	1.98	294.86	0.0000
lesionsize	1	5.58	5.58	833.03	0.0000
factor(LesionShape)	2	4.65	2.32	346.65	0.0000
Residuals	5761	38.61	0.01		

**Table 19: Texture Entropy significance for heterogeneous lesions**

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
scanner	1	0.52	0.52	1.83	0.1763
slicethickness	1	239.49	239.49	849.09	0.0000
kernel	2	239.27	119.63	424.16	0.0000
ReconAlgorithm	2	96.65	48.33	171.34	0.0000
ctdi	1	83.66	83.66	296.60	0.0000
lesionsize	1	269.05	269.05	953.93	0.0000
factor(LesionShape)	2	284.15	142.08	503.72	0.0000
Residuals	5761	1624.89	0.28		

**Table 20: Variance significance for heterogeneous lesions**

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
scanner	1	2.30	2.30	959.76	0.0000
slicethickness	1	1.01	1.01	421.66	0.0000
kernel	2	0.77	0.38	160.12	0.0000
ReconAlgorithm	2	0.01	0.00	1.40	0.2467
ctdi	1	0.04	0.04	17.89	0.0000
lesionsize	1	0.34	0.34	143.69	0.0000
factor(LesionShape)	2	0.09	0.04	18.67	0.0000
Residuals	5761	13.82	0.00		

	Contrast	Correlation	Dissimilarity	Energy	Homogeneity	Texture Entropy	Variance
Scanner Type	●	●	●	●	●	●	●
Slice Thickness	●	●	●	●	●	●	●
Kernel	●	●	●	●	●	●	●
Recon Algorithm	●	●	●	●	●	●	●
CTDI	●	●	●	●	●	●	●
Lesion Size	●	●	●	●	●	●	●
Lesion Shape	●	●	●	●	●	●	●

**Figure 16: Colored circled plot of variance significance for heterogeneous lesions**

## **6.4 Discussion**

Table 5 shows the correlation of the ground truth lesion features and high correlation was observed between most of the features. Although 24 physical lesions were present, the physical lesions were a repeat of the virtually designed ones. There were 12 distinct ground truth lesions (6 heterogeneous, 6 homogeneous) designed, each of which obtaining sizes of large or small, shapes of spherical, lobulated, or speculated. It is indicated that for the lesions designed, the features were highly correlated.

From Figure 12 - 15, bias was observed for all features, especially for Energy, Homogeneity, Correlation, and Dissimilarity. Normally, the median of the boxplot goes through the zero line. The four features mentioned above shifted greatly from the zero line, indicating strong bias. The GLCM texture features are not derived from normal

gray level pixel intensity value, but rather the frequency of occurrence in the GLCM, making the understanding less intuitive. Energy, also known as uniformity or angular second moment, is the sum of squared elements in the GLCM. Homogeneity measures how close the distribution of elements are in the GLCM diagonal. Correlation measures the joint probability of pixel pairs. Dissimilarity measures the local difference. High bias for those four features between homogeneous lesion errors and heterogeneous lesion errors indicates that the texture feature calculation of lesions of uniform background and textured background were affected by imaging conditions differently.

Table 5 also demonstrated high relative bias and high relative standard deviation. Due to the fact that each feature was on a difference scale, normalization was done in order to compare the features with each other. Strong relative bias was seen in Contrast, Energy, and Texture Entropy for heterogeneous lesions, each of which has high relative standard deviation. It is expected to observe high variations for each feature for both homogeneous and heterogeneous lesions when imaged due to the change in imaging conditions. Note that values of Inf or NA for the homogeneous lesions were due to the fact that some homogeneous lesion truth has value of zero. For the values that were present, we observed higher Energy relative bias and relative standard deviation between the two textured lesion categories. For Correlation, homogeneous lesions had higher relative bias than heterogeneous lesions, while the relative standard deviation for homogeneous lesions were smaller. The same trend was

observed for Homogeneity. The high relative bias in Energy, Correlation, and Homogeneity between heterogeneous and homogeneous lesion categories again indicated that lesion texture features with various background (heterogeneous or homogeneous) were affected by imaging conditions differently.

Table 6-12 showed the results for bias. The baseline of the model is Siemens Somatom Flash scanner, slice thickness of 5mm, standard kernel, reconstruction with FBP, CTDI of 5.8, lesion size of large, and lesion shape of spherical. Other conditions were compared to the baseline, and the effects, standard error, and p-value for significance was measured for bias. For the analysis of variance, the variations within each imaging and lesion condition were compared. To provide a clearer view of which conditions' bias are strongly significant (i.e., the feature result of this condition is significantly different from the baseline condition), a circled plot with color indications were constructed, as seen in Figure 15.

Most of the conditions were significantly biased from the baseline condition, except for some conditions. For kernel of Soft, the only feature that was significant was Homogeneity, while the other features were not significantly different from the baseline, which is kernel of Standard. For reconstruction algorithms, IR with high strength is not significantly biased from FBP for features Contrast and Variance. For CTDI of 5.8, the feature Contrast was not significantly different from baseline condition of CTDI 0.6. For lesions shapes other than the baseline, features Contrast, Dissimilarity, Homogeneity,

and Variance were not significantly different from baseline, which has shape spherical. The result for bias significance demonstrated that the features in a texture feature category should be considered separately when exploring the effects of imaging conditions, and that each feature was affected greatly by most imaging conditions. Sizes of lesions should also be taken into consideration when calculating texture features under various imaging conditions.

Table 13 – 18 showed the variance significance for heterogeneous lesions. Again, a clearer visual representation was shown in Figure 16. High and significant variance was seen almost all features except for some. For Contrast, there was not a significant difference between CTDI variations of 0.6 and 5.8. For Texture Entropy, there was not a significant difference between scanner types, GE Revolution and Siemens Somatom Flash. For Variance, there was not a significant difference between Reconstruction Algorithms, FBP, IR – low strength, and IR – high strength. The highly significant variance for most features were again an indication of the importance of imaging and lesion conditions on texture quantification.

## ***6.4 Conclusion and Future Work***

Variability related to CT imaging acquisition and reconstruction techniques is a clinically important source of bias and variance during lesion heterogeneity quantification. Lesion size and shape should also be taken into consideration. The features within a feature category should be considered separately since they were

affected by the imaging and lesion conditions differently. Future work include exploring materials for 3D printing with higher contrast, so the textures would be more visible after being imaged. During the segmentation process, possible human errors could occur due to the fact that the lesions were segmented by one individual at different times. A more robust automatic segmentation algorithm would be beneficial in eliminating these errors.

## Works Cited

1. Jemal, A., et al., Cancer statistics, 2010. *CA Cancer J Clin*, 2010. 60(5): p. 277-300.
2. National Lung Screening Trial Research, T., et al., Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med*, 2011. 365(5): p. 395-409.
3. Lazarus, D.R. and D.E. Ost, The solitary pulmonary nodule-deciding when to act? *Semin Respir Crit Care Med*, 2013. 34(6): p. 748-61.
4. Solomon, J. and E. Samei, A generic framework to simulate realistic lung, liver and renal pathologies in CT imaging. *Phys Med Biol*, 2014. 59(21): p. 6637-57.
5. Mendonca, P.R., et al., Model-based analysis of local shape for lesion detection in CT scans. *Med Image Comput Comput Assist Interv*, 2005. 8(Pt 1): p. 688-95.
6. Farag, A., et al., Modeling of the lung nodules for detection in LDCT scans. *Conf Proc IEEE Eng Med Biol Soc*, 2010. 2010: p. 3618-21.
7. Pastor, L., et al., A pulmonary nodule modeling tool as a diagnostic aid for lung HRCT images. *Comput Med Imaging Graph*, 2005. 29(8): p. 631-7.
8. Farag, A.A., et al., Quantitative nodule detection in low dose chest CT scans: new template modeling and evaluation for CAD system design. *Med Image Comput Comput Assist Interv*, 2005. 8(Pt 1): p. 720-8.
9. Davnall, F., et al., Assessment of tumor heterogeneity: an emerging imaging tool for clinical practice? *Insights Imaging*, 2012. 3(6): p. 573-89.
10. Gerlinger, M., et al., Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med*, 2012. 366(10): p. 883-92.
11. Kumar, V., et al., Radiomics: the process and the challenges. *Magn Reson Imaging*, 2012. 30(9): p. 1234-48.
12. Win, T., et al., Tumor heterogeneity and permeability as measured on the CT component of PET/CT predict survival in patients with non-small cell lung cancer. *Clin Cancer Res*, 2013. 19(13): p. 3591-9.

13. Ganeshan, B., et al., Texture analysis of non-small cell lung cancer on unenhanced computed tomography: initial evidence for a relationship with tumour glucose metabolism and stage. *Cancer Imaging*, 2010. 10: p. 137-43.
14. Solomon, J. and E. Samei, Quantum noise properties of CT images with anatomical textured backgrounds across reconstruction algorithms: FBP and SAFIRE. *Med Phys*, 2014. 41(9): p. 091908.
15. Solomon, J.B., X. Li, and E. Samei, Relating noise to image quality indicators in CT examinations with tube current modulation. *AJR Am J Roentgenol*, 2013. 200(3): p. 592-600.
16. Sigal-Cinqualbre, A.B., et al., Low-kilovoltage multi-detector row chest CT in adults: feasibility and effect on image quality and iodine dose. *Radiology*, 2004. 231(1): p. 169-74.
17. Huda, W., E.M. Scalzetti, and G. Levin, Technique factors and image quality as functions of patient weight at abdominal CT. *Radiology*, 2000. 217(2): p. 430-5.
18. Li, X., et al., Lung nodule detection in pediatric chest CT: quantitative relationship between image quality and radiologist performance. *Med Phys*, 2011. 38(5): p. 2609-18.
19. Zhao, B., et al., Exploring Variability in CT Characterization of Tumors: A Preliminary Phantom Study. *Transl Oncol*, 2014. 7(1): p. 88-93.
20. Mackin, D., et al., Measuring Computed Tomography Scanner Variability of Radiomics Features. *Invest Radiol*, 2015. 50(11): p. 757-65.
21. Baraldi, A. and F. Parmiggiani, An Investigation of the Textural Characteristics Associated with Gray-Level Cooccurrence Matrix Statistical Parameters. *Ieee Transactions on Geoscience and Remote Sensing*, 1995. 33(2): p. 293-304.
22. Wei, Q. and Y. Hu, A Study on Using Texture Analysis Methods for Identifying Lobar Fissure Regions in Isotropic CT Images. 2009 Annual International Conference of the Ieee Engineering in Medicine and Biology Society, Vols 1-20, 2009: p. 3537-3540.

23. Vallieres, M., et al., A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. *Phys Med Biol*, 2015. 60(14): p. 5471-96.
24. Bochud, F., C. Abbey, and M. Eckstein, Statistical texture synthesis of mammographic images with super-blob lumpy backgrounds. *Opt Express*, 1999. 4(1): p. 33-42.
25. Castella, C., et al., Mammographic texture synthesis: second-generation clustered lumpy backgrounds using a genetic algorithm. *Opt Express*, 2008. 16(11): p. 7595-607.
26. Li, J., et al., Tumor heterogeneity in neoplasms of breast, colon, and skin. *BMC Res Notes*, 2010. 3: p. 321.
27. Portilla, J. and E.P. Simoncelli, A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision*, 2000. 40(1): p. 49-71.
28. Kupinski, M.A., et al., Experimental determination of object statistics from noisy images. *J Opt Soc Am A Opt Image Sci Vis*, 2003. 20(3): p. 421-9.
29. Chen, B.Y., et al., Assessment of volumetric noise and resolution performance for linear and nonlinear CT reconstruction methods. *Medical Physics*, 2014. 41(7).
30. Friedman, S.N., et al., A simple approach to measure computed tomography (CT) modulation transfer function (MTF) and noise-power spectrum (NPS) using the American College of Radiology (ACR) accreditation phantom. *Medical Physics*, 2013. 40(5).
31. Brullmann, D.D. and B. d'Hoedt, The modulation transfer function and signal-to-noise ratio of different digital filters: a technical approach. *Dentomaxillofac Radiol*, 2011. 40(4): p. 222-9.
32. Whitley, D., A Genetic Algorithm Tutorial. *Statistics and Computing*, 1994. 4(2): p. 65-85.
33. R.W. Floyd, L. Steinberg. An Adaptive Algorithm for Spatial Gray Scale. In *Proceedings of Society for Information Displays (SID). Digest of technical papers*, pp. 36-37, 1975.