

## Appendix E1

### Section 1: CT Image Acquisition Parameters

#### Site 1:

Nonenhanced chest CT was acquired using a multidetector row CT system (in majority cases the manufacturer was either Siemens [Syngo, Siemens AG, Erlangen, Germany] or Philips [iCT, Philips Medical Systems, Cleveland, Ohio]). The acquisition parameters were as follows: 120–140 kVp exposure; 1.2 pitch and matrix,  $512 \times 512$ .

#### Site 2:

All patients underwent nonenhanced CT of the chest performed by using 16/64 slice CT system (Siemens [SOMATOM Sensation], General Electric [Lightspeed16, GE Medical Systems, Waukesha, Wisconsin], or Toshiba [Aquilion, Tochigi-ken, Japan]). The acquisition parameters were as follows: 120 kVp exposure; 1 pitch and image matrix,  $512 \times 512$ .

Precaution was taken to equally distribute cases in training and test set and account for variability in slice thickness and reconstruction kernel, as shown in Table E1.

### Section 2: Peri-nodular Mask

In the cohort of 290 CT scans, the pixel sizes ranged from  $0.42 \times 0.42$  mm to  $0.97 \times 0.97$  mm with an average size of  $0.73 \times 0.73$  mm (see Fig E1).

The peri-nodular mask was dilated a distance of 30 mm by taking into account the individual pixel size of each case. The number of pixels dilated in each peri-nodular mask were calculated as follows:

$$\# \text{ pixels dilated} = \frac{30}{\text{pixel size (in mm)}}$$

During this peri-nodular texture analysis of the lung parenchyma, care was taken to threshold the Hounsfield units (HU) of the CT scan, to remove air ( $< -900$  HU) and mediastinal muscle pixels ( $> -100$  HU). To avoid any edge artifacts that might arise during feature extraction, the thresholded ‘dead’ pixels of the CT scan were substituted by using an averaging filter across its  $9 \times 9$  neighborhood.

### Section 3: Texture Features Extracted

The following texture features were extracted from the intra-and peri-nodular regions:

- Haralick (13 descriptors): Quantifies the spatial gray-level co-occurrence matrix (GLCM) within local neighborhoods around each pixel in an image, stored in the form of matrices (36).

- Gabor (25 descriptors): Considers oriented textures via changes in direction and scale to capture microarchitectures (37). Each descriptor quantifies response to a given Gabor filter at a specific frequency ( $f = 0, 2, 4, 8, 16, 32$ ) and orientation ( $\theta = \pi/8, \pi/4, 3\pi/8, \pi/2, 3\pi/4$ ).

- Laws energy (25 descriptors):  $5 \times 5$  masks that are symmetric or antisymmetric to extract level (L), edge (E), spot (S), wave (W), and ripple (R) patterns on an image (38). Convolution of these masks with every image resulted in a total of 25 distinct Laws features for each image for every CT sequence.

## Section 4: Interreader Segmentation Variability

A randomly selected cohort of 60 cases from the training set (30 adenocarcinomas and 30 granulomas) was assessed for interreader variability of manual segmentation using the independent and separate measurements of reader 1 (P.R, a cardiothoracic radiologist with 10 years of experience) and reader 2 (K.B, a physician with 3 years of experience in cardiothoracic radiology research). These annotations were obtained in a single readout session.

The Dice Similarity Coefficient (DSC) score is a statistical measure that validates the performance of both readers and evaluates the reproducibility of their manual segmentations (39). The DSC score range is calculated between 0 to 1, where 0 indicates no spatial overlap between two sets of manual segmentation results and 1 represents a complete overlap. A DSC score of 0.81 was obtained between the two readers in our experiment, which represents a substantial agreement and reproducibility between the annotated regions of interest as previously demonstrated by Zou et al (40).

The impact of interreader variability of manual annotations on the most discriminating intranodular and peri-nodular radiomic features (listed in Table 3) was assessed using the absolute value of the percent coefficient of variation (%COV). %COV is a standardized metric to assess the extent of dispersion with respect to the mean of a given population (41). It is often expressed as a percentage, where a lower value indicates less variability in data and a higher percentage value indicates more variability in the given population.

$$\%COV = abs\left(\frac{\text{Standard Deviation}}{\text{Mean}} \times 100\right)$$

Figure E2 below shows the %COV across 12 radiomic features (as listed in experiment 1 and experiment2, Table 2 from the main paper) extracted from the largest slice on the nodule for Reader 1 and Reader 2 segmentations. Across the same population, we found that the average %COV for the top radiomic features from the intranodular and peri-nodular region for Reader 1 and Reader 2 was 9.18 and 13.56 respectively. These low %COV strongly suggest that our top performing radiomic features are largely resilient to variations as a result of reader segmentations.

## Section 5: Intranodular Radiomic Classifier

To determine intranodular imaging features that best discriminated adenocarcinomas from granulomas in CT,  $t$  test selection was implemented to pick the top 12 features with the lowest  $P$  value. In a supervised setting, these top features were used to train Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), Support Vector Machines (SVM-linear and RBF kernels) and Random Forest classifiers. The classifiers were then validated on an independent set ( $n = 145$ ) and the associated performance was assessed using the Area Under the Curve (AUC).

The top 12 radiomic feature set obtained from the training set after via the *t* test based feature selection are listed in Table E5. Gabor features were most frequently represented, occupying ten of the top 12 features (Figure E3). Using these top twelve features, highest AUC on the test set was obtained using for the SVM classifier with a linear kernel. This classifier yielded an AUC of 0.75, accuracy of 0.65, sensitivity of 0.75 and specificity of 0.61 at the operating point on the ROC curve (see Figure E6). Table E7 lists the AUCs for the other machine learning classifiers.

## Section 6: Deep Learning Model

As seen in Figure E7, our convolutional neural network architecture, consisted of two sets of convolutional, activation (ReLU), and pooling layers, followed by a fully-connected layer, activation, another fully-connected, and finally a softmax classifier (Table E6). A simple patch-based classification approach was implemented where the softmax classifier returned a probability of each patch belonging to either of the two classes of interest (adenocarcinoma or granuloma). The model was trained over 100 epochs after which the weights were locked down. The performance metric across the training runs are shown in Figure E8. The learned weights were then used on the independent validation set of 145 studies, and the predicted probabilities were utilized to generate the receiver operating characteristic curve.

Using the same training and testing sets as in the previous experiments, with LeNet model accuracy approaches 1 as number of epochs increases, and plateaus over 80 epochs (see Figure E8). The weights after 100 epochs were locked down and used on the test set. The testing AUC was computed to be 0.76. At the operating point, the corresponding accuracy was computed to be 0.737.

## References

36. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Trans Syst Man Cybern 1973;SMC-3(6):610–621 .
37. Jain AK, Farrokhnia F. Unsupervised texture segmentation using Gabor filters. 1990 IEEE International Conference on Systems, Man, and Cybernetics Conference Proceedings. 1990; 14–19 .
38. Laws KI. Textured image segmentation. University of Southern California Los Angeles Image Processing Inst, 1980 Jan. Report No.: USCIP-940. <http://www.dtic.mil/docs/citations/ADA083283>. Accessed November 24, 2016.
39. Dice LR. Measures of the amount of ecologic association between species. Ecology 1945;26(3):297–302 .
40. Zou KH, Warfield SK, Bharatha A, et al. Statistical validation of image segmentation quality based on a spatial overlap index. Acad Radiol 2004;11(2):178–189 .
41. Shafiq-Ul-Hassan M, Zhang GG, Latifi K, et al. Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels. Med Phys 2017;44(3):1050–1062 .

**Table E1: CT Image Acquisition Parameter Distribution over Training Set and Independent Test Validation**

	Training Set			Independent Validation Set		
	Adenocarcinoma	Granuloma	Total	Adenocarcinoma	Granuloma	Total
Patients	73	72	145	72	73	145
<b>Scanner Type</b>						
GE Medical Systems	0	3	3	0	1	1
Siemens	44	49	93	53	32	85
Philips	28	18	46	19	39	58
Toshiba	1	2	3	0	1	1
<b>Slice Thickness (mm)</b>						
1	8	5	13	9	1	10
1.5	0	1	1	0	0	0
2	42	38	80	40	54	94
2.5	0	0	0	0	1	1
3	10	15	25	12	4	16
3.2	9	0	6	0	0	0
3.35	0	1	1	0	0	0
4	0	0	0	0	1	1
5	5	11	16	10	12	22
6	1	1	2	0	0	0
6.5	1	0	1	1	0	1
<b>Reconstruction Kernel</b>						
<i>GE Medical Systems</i>						
Standard	0	1	1	0	1	1
<i>SIEMENS</i>						
B20f	0	1	1	0	1	1
BZOs	0	1	1	0	0	1
B30f	10	8	18	5	6	11
B30S	0	5	5	1	2	3
B3Lf	9	4	13	13	2	15
B31s	0	1	1	1	1	2
B35f	7	8	15	12	0	12
B35s	2	2	4	0	0	0
B40f	2	2	4	5	0	5
B40s	0	1	1	0	0	0
B4Lf	5	11	16	8	8	16
B41s	0	0	0	0	1	1
B50f	2	3	5	2	8	10
B50S	1	0	1	0	0	0
B60f	3	1	4	1	2	3
B70f	1	3	4	0	0	0
B70s	1	0	1	0	1	1
T20S	1	0	1	5	0	5
<i>PHILIPS</i>						
A	1	0	1	0	2	2
B	8	4	12	5	16	21
C	6	5	11	5	8	13

D	9	2	11	2	0	2
E	0	1	1	1	0	1
L	3	6	9	5	11	16
YA	1	0	1	0	2	2
YB	0	0	0	1	0	1
<i>TOSHIBA</i>						
FC02	0	1	1	0	0	0
FC08	0	0	0	0	1	1
FC10	0	1	1	0	0	0
FC52	1	0	1	0	0	0

**Table E2: Radiomic Features and Their Corresponding Pathophysiology**

Feature category	Descriptor	Intuitive Description	Relevance to SPN pathophysiology
Haralick features (Repeated occurrence of gray level configuration in the texture represented via the gray-level co-occurrence matrix (GLCM), which varies rapidly with distance in fine textures and slowly in large textures)	Inverse Difference Moment (IDM)	IDM is a reflection of the presence or absence of uniformity, and hence is a measure of local regions of homogeneity High IDM: Higher presence of locally uniform windows in GLCM Low IDM: Higher presence of locally heterogeneous windows in GLCM	Captures underlying heterogeneity within the lesion.
	Correlation	Quantifies the linear patterns in an image based on the distance parameter.	Increased presence of linear patterns yield higher correlation values, lack of image linearity yield lower correlation values
	Sum Entropy	Measure of GLCM relationship to distribution of intensity with respect to entropy. Entropy is the measure of disorder.	Higher entropy is indicative of more chaotic arrangement in areas of high viable cell population
	Sum Variance	Measure of GLCM relationship to distribution of intensity with respect to variance. High sum variance: greater standard deviation of sum average. Low sum variance: low standard deviation of sum average	Possibly accounting for greater variation of scattered atypia and local accumulation of mitotic processes as observed on histopathology.
Laws features	E5, L5, S5, R5 (combination in both X and Y directions)	E-Edges L-Level S-Spots R-Ripples	Accounting for characteristic qualitative appearance of ROIs (for e.g.: ring enhancement patterns, diffused tumor calcification or patchy appearances)
Gabor	Quantifies response to a given Gabor filter at a specific frequency and orientation	These filters comprise of various scales and orientations to locally characterize intensity variations.	Possibly capturing changes in tumor microarchitecture. And quantifying tumor lymphangiogenesis in adenocarcinomas and lymphohistiocytic inflammatory response in granulomas.

**Table E3: List of Shape Features**

Features	Description
Size	Including Width, Height, Depth of bounding box
Area	from 2D slices of each nodule
Perimeter	from 2D slices of each nodule
Eccentricity	foci of the ellipse and to major axis length
Extend	ratio of pixels in the region to pixels in the total bounding box
Compactness	ratio of the perimeter squared to the product of $4\pi$ and area
Radial distance	distances from center of each slice to contour points
Roughness	perimeter of slices divided by convex perimeter

Elongation	from major and minor axis
Convexity	from convex hull
Equivalent Diameter	Diameter of circle with same area of slices
Sphericity	3D compactness

**Table E4: Human Machine Comparison**

Patient ID	Ground Truth	Reader 1 Decision	Reader 2 Decision	Machine Probability to be granuloma	Machine Probability to be adenocarcinoma	Machine Decision
1	Adenocarcinoma	Adenocarcinoma	Granuloma	0.15698702	0.84301298	Adenocarcinoma
2	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.451797943	0.548202057	Adenocarcinoma
3	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.428191871	0.571808129	Adenocarcinoma
4	Adenocarcinoma	Adenocarcinoma	Granuloma	0.217179742	0.782820258	Adenocarcinoma
5	Adenocarcinoma	Granuloma	Granuloma	0.170487314	0.829512686	Adenocarcinoma
6	Adenocarcinoma	Granuloma	Granuloma	0.310319347	0.689680653	Adenocarcinoma
7	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.594666818	0.405333182	Granuloma
8	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.398721827	0.601278173	Adenocarcinoma
9	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.182128699	0.817871301	Adenocarcinoma
10	Adenocarcinoma	Adenocarcinoma	Granuloma	0.457522726	0.542477274	Adenocarcinoma
11	Adenocarcinoma	Adenocarcinoma	Granuloma	0.225932417	0.774067583	Adenocarcinoma
12	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.546704053	0.453295947	Granuloma
13	Adenocarcinoma	Granuloma	Granuloma	0.627371556	0.372628444	Granuloma
14	Adenocarcinoma	Granuloma	Granuloma	0.44143471	0.55856529	Adenocarcinoma
15	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.487733583	0.512266417	Adenocarcinoma
16	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.258908979	0.741091021	Adenocarcinoma
17	Adenocarcinoma	Granuloma	Granuloma	0.343601119	0.656398881	Adenocarcinoma
18	Adenocarcinoma	Granuloma	Granuloma	0.388671284	0.611328716	Adenocarcinoma
19	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.459183649	0.540816351	Adenocarcinoma
20	Adenocarcinoma	Granuloma	Granuloma	0.339556798	0.660443202	Adenocarcinoma
21	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.697577049	0.302422951	Granuloma
22	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.592326747	0.407673253	Granuloma
23	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.467184248	0.532815752	Adenocarcinoma
24	Adenocarcinoma	Adenocarcinoma	Granuloma	0.327260692	0.672739308	Adenocarcinoma
25	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.421689026	0.578310974	Adenocarcinoma
26	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.508677261	0.491322739	Granuloma
27	Adenocarcinoma	Adenocarcinoma	Granuloma	0.279868522	0.720131478	Adenocarcinoma
28	Adenocarcinoma	Granuloma	Adenocarcinoma	0.708279217	0.291720783	Granuloma
29	Adenocarcinoma	Granuloma	Granuloma	0.412411045	0.587588955	Adenocarcinoma
30	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.435403834	0.564596166	Adenocarcinoma
31	Adenocarcinoma	Adenocarcinoma	Granuloma	0.500569409	0.499430591	Granuloma
32	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.262578169	0.737421831	Adenocarcinoma
33	Adenocarcinoma	Granuloma	Granuloma	0.533491209	0.466508791	Granuloma
34	Adenocarcinoma	Adenocarcinoma	Granuloma	0.66020064	0.33979936	Granuloma
35	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.529480295	0.470519705	Granuloma
36	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.46269746	0.53730254	Adenocarcinoma
37	Adenocarcinoma	Granuloma	Adenocarcinoma	0.22140399	0.77859601	Adenocarcinoma
38	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.488864055	0.511135945	Adenocarcinoma
39	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.641295672	0.358704328	Granuloma
40	Adenocarcinoma	Granuloma	Granuloma	0.277193319	0.722806681	Adenocarcinoma
41	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.219110513	0.780889487	Adenocarcinoma

42	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.13796198	0.86203802	Adenocarcinoma
43	Adenocarcinoma	Granuloma	Granuloma	0.068405524	0.931594476	Adenocarcinoma
44	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.336761786	0.663238214	Adenocarcinoma
45	Adenocarcinoma	Granuloma	Granuloma	0.279784671	0.720215329	Adenocarcinoma
46	Adenocarcinoma	Adenocarcinoma	Granuloma	0.507869671	0.492130329	Granuloma
47	Adenocarcinoma	Granuloma	Granuloma	0.512467548	0.487532452	Granuloma
48	Adenocarcinoma	Granuloma	Granuloma	0.170699255	0.829300745	Adenocarcinoma
49	Adenocarcinoma	Granuloma	Granuloma	0.168254547	0.831745453	Adenocarcinoma
50	Adenocarcinoma	Granuloma	Granuloma	0.199324536	0.800675464	Adenocarcinoma
51	Adenocarcinoma	Adenocarcinoma	Granuloma	0.220089569	0.779910431	Adenocarcinoma
52	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.187993409	0.812006591	Adenocarcinoma
53	Adenocarcinoma	Granuloma	Granuloma	0.699486118	0.300513882	Granuloma
54	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.168752822	0.831247178	Adenocarcinoma
55	Adenocarcinoma	Granuloma	Granuloma	0.490589132	0.509410868	Adenocarcinoma
56	Adenocarcinoma	Granuloma	Granuloma	0.200721347	0.799278653	Adenocarcinoma
57	Adenocarcinoma	Granuloma	Granuloma	0.175873853	0.824126147	Adenocarcinoma
58	Adenocarcinoma	Adenocarcinoma	Granuloma	0.564827451	0.435172549	Granuloma
59	Adenocarcinoma	Adenocarcinoma	Granuloma	0.127602442	0.872397558	Adenocarcinoma
60	Adenocarcinoma	Granuloma	Granuloma	0.129927656	0.870072344	Adenocarcinoma
61	Adenocarcinoma	Granuloma	Granuloma	0.264220845	0.735779155	Adenocarcinoma
62	Adenocarcinoma	Adenocarcinoma	Granuloma	0.217930978	0.782069022	Adenocarcinoma
63	Adenocarcinoma	Granuloma	Granuloma	0.212184059	0.787815941	Adenocarcinoma
64	Adenocarcinoma	Adenocarcinoma	Adenocarcinoma	0.133448459	0.866551541	Adenocarcinoma
65	Adenocarcinoma	Granuloma	Granuloma	0.160772986	0.839227014	Adenocarcinoma
66	Adenocarcinoma	Granuloma	Granuloma	0.115126289	0.884873711	Adenocarcinoma
67	Adenocarcinoma	Granuloma	Granuloma	0.214203904	0.785796096	Adenocarcinoma
68	Adenocarcinoma	Granuloma	Granuloma	0.144694627	0.855305373	Adenocarcinoma
69	Adenocarcinoma	Granuloma	Adenocarcinoma	0.159683611	0.840316389	Adenocarcinoma
70	Adenocarcinoma	Granuloma	Granuloma	0.374095144	0.625904856	Adenocarcinoma
71	Adenocarcinoma	Granuloma	Granuloma	0.126530107	0.873469893	Adenocarcinoma
72	Adenocarcinoma	Granuloma	Granuloma	0.148974261	0.851025739	Adenocarcinoma
73	Granuloma	Granuloma	Adenocarcinoma	0.322456663	0.677543337	Adenocarcinoma
74	Granuloma	Granuloma	Granuloma	0.647035946	0.352964054	Granuloma
75	Granuloma	Granuloma	Adenocarcinoma	0.485036359	0.514963641	Adenocarcinoma
76	Granuloma	Granuloma	Granuloma	0.517373973	0.482626027	Granuloma
77	Granuloma	Adenocarcinoma	Adenocarcinoma	0.433816398	0.566183602	Adenocarcinoma
78	Granuloma	Adenocarcinoma	Adenocarcinoma	0.495053649	0.504946351	Adenocarcinoma
79	Granuloma	Adenocarcinoma	Granuloma	0.467139728	0.532860272	Adenocarcinoma
80	Granuloma	Granuloma	Granuloma	0.609387618	0.390612382	Granuloma
81	Granuloma	Granuloma	Granuloma	0.247117744	0.752882256	Adenocarcinoma
82	Granuloma	Adenocarcinoma	Adenocarcinoma	0.503335025	0.496664975	Granuloma
83	Granuloma	Adenocarcinoma	Granuloma	0.675145581	0.324854419	Granuloma
84	Granuloma	Adenocarcinoma	Adenocarcinoma	0.479232204	0.520767796	Adenocarcinoma
85	Granuloma	Granuloma	Granuloma	0.229430508	0.770569492	Adenocarcinoma
86	Granuloma	Adenocarcinoma	Adenocarcinoma	0.521294709	0.478705291	Granuloma
87	Granuloma	Granuloma	Granuloma	0.695016389	0.304983611	Granuloma
88	Granuloma	Adenocarcinoma	Adenocarcinoma	0.740984791	0.259015209	Granuloma
89	Granuloma	Adenocarcinoma	Granuloma	0.799722283	0.200277717	Granuloma
90	Granuloma	Granuloma	Granuloma	0.167484876	0.832515124	Adenocarcinoma

91	Granuloma	Granuloma	Granuloma	0.604168881	0.395831119	Granuloma
92	Granuloma	Granuloma	Granuloma	0.793995055	0.206004945	Granuloma
93	Granuloma	Granuloma	Granuloma	0.466947239	0.533052761	Adenocarcinoma
94	Granuloma	Granuloma	Granuloma	0.820598836	0.179401164	Granuloma
95	Granuloma	Granuloma	Granuloma	0.614592672	0.385407328	Granuloma
96	Granuloma	Adenocarcinoma	Adenocarcinoma	0.565909981	0.434090019	Granuloma
97	Granuloma	Granuloma	Adenocarcinoma	0.508951349	0.491048651	Granuloma
98	Granuloma	Granuloma	Granuloma	0.592362637	0.407637363	Granuloma
99	Granuloma	Granuloma	Granuloma	0.679019111	0.320980889	Granuloma
100	Granuloma	Granuloma	Granuloma	0.114896247	0.885103753	Adenocarcinoma
101	Granuloma	Granuloma	Granuloma	0.2334818	0.7665182	Adenocarcinoma
102	Granuloma	Granuloma	Granuloma	0.458155947	0.541844053	Adenocarcinoma
103	Granuloma	Granuloma	Granuloma	0.348893928	0.651106072	Adenocarcinoma
104	Granuloma	Granuloma	Granuloma	0.368355109	0.631644891	Adenocarcinoma
105	Granuloma	Adenocarcinoma	Granuloma	0.44925938	0.55074062	Adenocarcinoma
106	Granuloma	Granuloma	Granuloma	0.868440244	0.131559756	Granuloma
107	Granuloma	Adenocarcinoma	Adenocarcinoma	0.339678087	0.660321913	Adenocarcinoma
108	Granuloma	Adenocarcinoma	Granuloma	0.492012287	0.507987713	Adenocarcinoma
109	Granuloma	Adenocarcinoma	Adenocarcinoma	0.550445047	0.449554953	Granuloma
110	Granuloma	Granuloma	Granuloma	0.516011103	0.483988897	Granuloma
111	Granuloma	Adenocarcinoma	Granuloma	0.636684478	0.363315522	Granuloma
112	Granuloma	Adenocarcinoma	Granuloma	0.802220259	0.197779741	Granuloma
113	Granuloma	Granuloma	Granuloma	0.398955824	0.601044176	Adenocarcinoma
114	Granuloma	Granuloma	Granuloma	0.700085546	0.299914454	Granuloma
115	Granuloma	Granuloma	Granuloma	0.402380168	0.597619832	Adenocarcinoma
116	Granuloma	Granuloma	Adenocarcinoma	0.55477403	0.44522597	Granuloma
117	Granuloma	Granuloma	Granuloma	0.510849378	0.489150622	Granuloma
118	Granuloma	Granuloma	Granuloma	0.389949797	0.610050203	Adenocarcinoma
119	Granuloma	Granuloma	Granuloma	0.752794434	0.247205566	Granuloma
120	Granuloma	Granuloma	Granuloma	0.706200334	0.293799666	Granuloma
121	Granuloma	Adenocarcinoma	Adenocarcinoma	0.560087419	0.439912581	Granuloma
122	Granuloma	Adenocarcinoma	Granuloma	0.503834788	0.496165212	Granuloma
123	Granuloma	Granuloma	Granuloma	0.577140752	0.422859248	Granuloma
124	Granuloma	Adenocarcinoma	Adenocarcinoma	0.535145577	0.464854423	Granuloma
125	Granuloma	Adenocarcinoma	Adenocarcinoma	0.521104111	0.478895889	Granuloma
126	Granuloma	Granuloma	Granuloma	0.707044276	0.292955724	Granuloma
127	Granuloma	Granuloma	Granuloma	0.714973642	0.285026358	Granuloma
128	Granuloma	Granuloma	Granuloma	0.583077546	0.416922454	Granuloma
129	Granuloma	Adenocarcinoma	Granuloma	0.595917392	0.404082608	Granuloma
130	Granuloma	Granuloma	Granuloma	0.578412376	0.421587624	Granuloma
131	Granuloma	Granuloma	Granuloma	0.691840653	0.308159347	Granuloma
132	Granuloma	Adenocarcinoma	Adenocarcinoma	0.420612843	0.579387157	Adenocarcinoma
133	Granuloma	Adenocarcinoma	Granuloma	0.858248592	0.141751408	Granuloma
134	Granuloma	Granuloma	Adenocarcinoma	0.455447785	0.544552215	Adenocarcinoma
135	Granuloma	Granuloma	Granuloma	0.571245827	0.428754173	Granuloma
136	Granuloma	Granuloma	Granuloma	0.722669039	0.277330961	Granuloma
137	Granuloma	Adenocarcinoma	Granuloma	0.461698041	0.538301959	Adenocarcinoma
138	Granuloma	Adenocarcinoma	Adenocarcinoma	0.621469142	0.378530858	Granuloma
139	Granuloma	Adenocarcinoma	Adenocarcinoma	0.905624642	0.094375358	Granuloma



140	Granuloma	Adenocarcinoma	Adenocarcinoma	0.919283893	0.080716107	Granuloma
141	Granuloma	Adenocarcinoma	Adenocarcinoma	0.810264724	0.189735276	Granuloma
142	Granuloma	Granuloma	Granuloma	0.637635392	0.362364608	Granuloma
143	Granuloma	Adenocarcinoma	Granuloma	0.455177001	0.544822999	Adenocarcinoma
144	Granuloma	Granuloma	Granuloma	0.261016412	0.738983588	Adenocarcinoma
145	Granuloma	Granuloma	Granuloma	0.661045577	0.338954423	Granuloma

**Table E5: Intranodular Radiomic Features**

Feature number	Feature Family	Descriptor	Statistic	Location	P value*
1	Gabor	$f = 32, \theta = 3\pi/4$	Kurtosis	Intranodular	0.0007
2	Gabor	$f = 4, \theta = 3\pi/4$	Skewness	Intranodular	0.007
3	Gabor	$f = 4, \theta = \pi/2$	Median	Intranodular	0.0008
4	Laws energy	W5E5†	Median	Intranodular	0.001
5	Laws energy	S5E5	Median	Intranodular	0.0005
6	Gabor	$f = 32, \theta = 3\pi/2$	Median	Intranodular	0.0004
7	Gabor	$f = 32, \theta = \pi/2$	Skewness	Intranodular	0.0004
8	Gabor	$f = 4, \theta = \pi/8$	Kurtosis	Intranodular	0.0004
9	Gabor	$f = 4, \theta = \pi/4$	Median	Intranodular	0.0004
10	Gabor	$f = 4, \theta = \pi/8$	Skewness	Intranodular	0.0004
11	Gabor	$f = 4, \theta = \pi/8$	Standard Deviation	Intranodular	0.0003
12	Gabor	$f = 4, \theta = \pi/8$	Median	Intranodular	0.0003

Column 1 reflects the feature number, column 2 indicates the feature family of the feature, column 3 provides details about the feature (such as  $f$  = frequency of the filter,  $\theta$  = orientation of the filter with respect to the normal axis). Column 4 reflects the statistic that was derived for the feature, column 5 reflects the nodule region of feature extraction, column 6 provides the  $P$  values of the features to distinguish adenocarcinomas from granulomas.

\* The  $P$  values were computed using the paired Student  $t$  test for continuous variables.

† Feature descriptors for Laws Energy-Descriptors include all combinations of five 1D filters: level (L), edge (E), spot (S), wave (W), and ripple (R). For example, in this experiment, W5E5 (feature number 4) implies that Laws energy based textural patterns of waves (W) in the horizontal direction and edges (E) in the vertical direction using 5-pixel by 5 pixel 2-D convolution filter was statistically different ( $P < .001$ ) between adenocarcinomas and granulomas. Further information on Laws energy features can be found in Table E2.

**Table E6: Hyper Parameter Optimization Details:**

Parameter	Details
Convolution 1	Kernel size: 5, stride: 2, padding = 'same'
Pool 1 (MAX pooling)	Kernel size: 2, stride: 2, padding = 'same'
Convolution 2	Kernel size: 5, stride: 2, padding = 'same'
Pool 2 (MAX pooling)	Kernel size: 2, stride: 2, padding = 'same'
Convolution 3	Kernel size: 5, stride: 2, padding = 'same'
Pool 3 (MAX pooling)	Kernel size: 2, stride: 2, padding = 'same'

**Table E7: AUC Values Obtained on the Training and Independent Test Set Using Different Classifiers**

	Intranodular classifier	
	Training AUC [95% CI]	Testing AUC [95% CI]
LDA	0.72 [0.55, 0.89]	0.72 [0.56, 0.89]
SVM-linear	<b>0.73 [0.56, 0.89]</b>	<b>0.75 [0.59, 0.91]</b>
SVM-RBF	0.63 [0.46, 0.82]	0.66 [0.48, 0.84]
RF	0.69 [0.52, 0.86]	0.72 [0.55, 0.88]
QDA	0.74 [0.58, 0.90]	0.72 [0.55, 0.88]

**Table E8: List of Textural Features Extracted**

252 features extracted from the intranodular region and each of the 6 perinodular regions				
#	Feature	frequency	theta	Statistic
1	Gabor Wavelet 1	2	$\pi/8$	Median
2	Gabor Wavelet 1			Standard Deviation
3	Gabor Wavelet 1			Skewness
4	Gabor Wavelet 1			Kurtosis
5	Gabor Wavelet 2		$\pi/4$	Median
6	Gabor Wavelet 2			Standard Deviation
7	Gabor Wavelet 2			Skewness
8	Gabor Wavelet 2			Kurtosis
9	Gabor Wavelet 3		$3^* \pi/8$	Median
10	Gabor Wavelet 3			Standard Deviation
11	Gabor Wavelet 3			Skewness
12	Gabor Wavelet 3			Kurtosis
13	Gabor Wavelet 4		$\pi/2$	Median
14	Gabor Wavelet 4			Standard Deviation
15	Gabor Wavelet 4			Skewness
16	Gabor Wavelet 4			Kurtosis
17	Gabor Wavelet 5		$3^* \pi/4$	Median
18	Gabor Wavelet 5			Standard Deviation
19	Gabor Wavelet 5			Skewness
20	Gabor Wavelet 5			Kurtosis
21	Gabor Wavelet 6	4	$\pi/8$	Median
22	Gabor Wavelet 6			Standard Deviation
23	Gabor Wavelet 6			Skewness
24	Gabor Wavelet 6			Kurtosis
25	Gabor Wavelet 7		$\pi/4$	Median
26	Gabor Wavelet 7			Standard Deviation
27	Gabor Wavelet 7			Skewness
28	Gabor Wavelet 7			Kurtosis
29	Gabor Wavelet 8		$3^* \pi/8$	Median
30	Gabor Wavelet 8			Standard Deviation
31	Gabor Wavelet 8			Skewness
32	Gabor Wavelet 8			Kurtosis
33	Gabor Wavelet 9		$\pi/2$	Median
34	Gabor Wavelet 9			Standard Deviation
35	Gabor Wavelet 9			Skewness
36	Gabor Wavelet 9			Kurtosis

37	Gabor Wavelet 10		$3*\pi/4$	Median	
38	Gabor Wavelet 10			Standard Deviation	
39	Gabor Wavelet 10			Skewness	
40	Gabor Wavelet 10			Kurtosis	
41	Gabor Wavelet 11	8	$\pi/8$	Median	
42	Gabor Wavelet 11			Standard Deviation	
43	Gabor Wavelet 11			Skewness	
44	Gabor Wavelet 11			Kurtosis	
45	Gabor Wavelet 12		$\pi/4$	Median	
46	Gabor Wavelet 12			Standard Deviation	
47	Gabor Wavelet 12			Skewness	
48	Gabor Wavelet 12			Kurtosis	
49	Gabor Wavelet 13		$3*\pi/8$	Median	
50	Gabor Wavelet 13			Standard Deviation	
51	Gabor Wavelet 13			Skewness	
52	Gabor Wavelet 13			Kurtosis	
53	Gabor Wavelet 14		$\pi/2$	Median	
54	Gabor Wavelet 14			Standard Deviation	
55	Gabor Wavelet 14			Skewness	
56	Gabor Wavelet 14			Kurtosis	
57	Gabor Wavelet 15		$3*\pi/4$	Median	
58	Gabor Wavelet 15			Standard Deviation	
59	Gabor Wavelet 15			Skewness	
60	Gabor Wavelet 15			Kurtosis	
61	Gabor Wavelet 16	16	$\pi/8$	Median	
62	Gabor Wavelet 16			Standard Deviation	
63	Gabor Wavelet 16			Skewness	
64	Gabor Wavelet 16			Kurtosis	
65	Gabor Wavelet 17			$\pi/4$	Median
66	Gabor Wavelet 17				Standard Deviation
67	Gabor Wavelet 17				Skewness
68	Gabor Wavelet 17				Kurtosis
69	Gabor Wavelet 18			$3*\pi/8$	Median
70	Gabor Wavelet 18				Standard Deviation
71	Gabor Wavelet 18				Skewness
72	Gabor Wavelet 18				Kurtosis
73	Gabor Wavelet 19			$\pi/2$	Median
74	Gabor Wavelet 19				Standard Deviation
75	Gabor Wavelet 19				Skewness
76	Gabor Wavelet 19				Kurtosis
77	Gabor Wavelet 20		$3*\pi/4$	Median	
78	Gabor Wavelet 20			Standard Deviation	
79	Gabor Wavelet 20			Skewness	
80	Gabor Wavelet 20			Kurtosis	
81	Gabor Wavelet 21	32	$\pi/8$	Median	
82	Gabor Wavelet 21			Standard Deviation	
83	Gabor Wavelet 21			Skewness	
84	Gabor Wavelet 21			Kurtosis	
85	Gabor Wavelet 22		$\pi/4$	Median	

86	Gabor Wavelet 22			Standard Deviation
87	Gabor Wavelet 22			Skewness
88	Gabor Wavelet 22			Kurtosis
89	Gabor Wavelet 23			Median
90	Gabor Wavelet 23			Standard Deviation
91	Gabor Wavelet 23			Skewness
92	Gabor Wavelet 23		Kurtosis	
93	Gabor Wavelet 24		Median	
94	Gabor Wavelet 24		Standard Deviation	
95	Gabor Wavelet 24		Skewness	
96	Gabor Wavelet 24		Kurtosis	
97	Gabor Wavelet 25		Median	
98	Gabor Wavelet 25	Standard Deviation		
99	Gabor Wavelet 25	Skewness		
100	Gabor Wavelet 25	Kurtosis		
101	Laws Energy 1	L5 L5	Median	
102	Laws Energy 1		Standard Deviation	
103	Laws Energy 1		Skewness	
104	Laws Energy 1		Kurtosis	
105	Laws Energy 2	L5E5	Median	
106	Laws Energy 2		Standard Deviation	
107	Laws Energy 2		Skewness	
108	Laws Energy 2		Kurtosis	
109	Laws Energy 3	L5S5	Median	
110	Laws Energy 3		Standard Deviation	
111	Laws Energy 3		Skewness	
112	Laws Energy 3		Kurtosis	
113	Laws Energy 4	L5 W5	Median	
114	Laws Energy 4		Standard Deviation	
115	Laws Energy 4		Skewness	
116	Laws Energy 4		Kurtosis	
117	Laws Energy 5	L5R5	Median	
118	Laws Energy 5		Standard Deviation	
119	Laws Energy 5		Skewness	
120	Laws Energy 5		Kurtosis	
121	Laws Energy 6	E5 L5	Median	
122	Laws Energy 6		Standard Deviation	
123	Laws Energy 6		Skewness	
124	Laws Energy 6		Kurtosis	
125	Laws Energy 7	E5E5	Median	
126	Laws Energy 7		Standard Deviation	
127	Laws Energy 7		Skewness	
128	Laws Energy 7		Kurtosis	
129	Laws Energy 8	E5S5	Median	
130	Laws Energy 8		Standard Deviation	
131	Laws Energy 8		Skewness	
132	Laws Energy 8		Kurtosis	
133	Laws Energy 9	E5 W5	Median	
134	Laws Energy 9		Standard Deviation	

135	Laws Energy 9		Skewness
136	Laws Energy 9		Kurtosis
137	Laws Energy 10	S5 L5	Median
138	Laws Energy 10		Standard Deviation
139	Laws Energy 10		Skewness
140	Laws Energy 10		Kurtosis
141	Laws Energy 11	S5E5	Median
142	Laws Energy 11		Standard Deviation
143	Laws Energy 11		Skewness
144	Laws Energy 11		Kurtosis
145	Laws Energy 12	S5S5	Median
146	Laws Energy 12		Standard Deviation
147	Laws Energy 12		Skewness
148	Laws Energy 12		Kurtosis
149	Laws Energy 13	S5 W5	Median
150	Laws Energy 13		Standard Deviation
151	Laws Energy 13		Skewness
152	Laws Energy 13		Kurtosis
153	Laws Energy 14	S5E5	Median
154	Laws Energy 14		Standard Deviation
155	Laws Energy 14		Skewness
156	Laws Energy 14		Kurtosis
157	Laws Energy 15	S5R5	Median
158	Laws Energy 15		Standard Deviation
159	Laws Energy 15		Skewness
160	Laws Energy 15		Kurtosis
161	Laws Energy 16	W5 L5	Median
162	Laws Energy 16		Standard Deviation
163	Laws Energy 16		Skewness
164	Laws Energy 16		Kurtosis
165	Laws Energy 17	W5E5	Median
166	Laws Energy 17		Standard Deviation
167	Laws Energy 17		Skewness
168	Laws Energy 17		Kurtosis
169	Laws Energy 18	W5S5	Median
170	Laws Energy 18		Standard Deviation
171	Laws Energy 18		Skewness
172	Laws Energy 18		Kurtosis
173	Laws Energy 19	W5 W5	Median
174	Laws Energy 19		Standard Deviation
175	Laws Energy 19		Skewness
176	Laws Energy 19		Kurtosis
177	Laws Energy 20	W5R5	Median
178	Laws Energy 20		Standard Deviation
179	Laws Energy 20		Skewness
180	Laws Energy 20		Kurtosis
181	Laws Energy 21	R5 L5	Median
182	Laws Energy 21		Standard Deviation
183	Laws Energy 21		Skewness

184	Laws Energy 21		Kurtosis
185	Laws Energy 22	R5E5	Median
186	Laws Energy 22		Standard Deviation
187	Laws Energy 22		Skewness
188	Laws Energy 22		Kurtosis
189	Laws Energy 23	R5S5	Median
190	Laws Energy 23		Standard Deviation
191	Laws Energy 23		Skewness
192	Laws Energy 23		Kurtosis
193	Laws Energy 24	R5 W5	Median
194	Laws Energy 24		Standard Deviation
195	Laws Energy 24		Skewness
196	Laws Energy 24		Kurtosis
197	Laws Energy 25	R5R5	Median
198	Laws Energy 25		Standard Deviation
199	Laws Energy 25		Skewness
200	Laws Energy 25		Kurtosis
201	Haralick 1	Entropy	Median
202	Haralick 1		Standard Deviation
203	Haralick 1		Skewness
204	Haralick 1		Kurtosis
205	Haralick 2	Energy	Median
206	Haralick 2		Standard Deviation
207	Haralick 2		Skewness
208	Haralick 2		Kurtosis
209	Haralick 3	Inertia	Median
210	Haralick 3		Standard Deviation
211	Haralick 3		Skewness
212	Haralick 3		Kurtosis
213	Haralick 4	Inverse Difference Moment (IDM)	Median
214	Haralick 4		Standard Deviation
215	Haralick 4		Skewness
216	Haralick 4		Kurtosis
217	Haralick 5	Correlation	Median
218	Haralick 5		Standard Deviation
219	Haralick 5		Skewness
220	Haralick 5		Kurtosis
221	Haralick 6	Correlation Information 1	Median
222	Haralick 6		Standard Deviation
223	Haralick 6		Skewness
224	Haralick 6		Kurtosis
225	Haralick 7	Correlation Information 2	Median
226	Haralick 7		Standard Deviation
227	Haralick 7		Skewness
228	Haralick 7		Kurtosis
229	Haralick 8	Sum Average	Median
230	Haralick 8		Standard Deviation
231	Haralick 8		Skewness
232	Haralick 8		Kurtosis

233	Haralick 9	Sum Variance	Median
234	Haralick 9		Standard Deviation
235	Haralick 9		Skewness
236	Haralick 9		Kurtosis
237	Haralick 10	Sum Entropy	Median
238	Haralick 10		Standard Deviation
239	Haralick 10		Skewness
240	Haralick 10		Kurtosis
241	Haralick 11	Difference Average	Median
242	Haralick 11		Standard Deviation
243	Haralick 11		Skewness
244	Haralick 11		Kurtosis
245	Haralick 12	Difference Variance	Median
246	Haralick 12		Standard Deviation
247	Haralick 12		Skewness
248	Haralick 12		Kurtosis
249	Haralick 13	Difference Entropy	Median
250	Haralick 13		Standard Deviation
251	Haralick 13		Skewness
252	Haralick 13		Kurtosis