December 2018

Change Descriptors for Determining Nodule Malignancy in Lung CT Screening Images

Benjamin Geiger
University of South Florida, benjamingeiger@gmail.com

Change Descriptors for Determining Nodule Malignancy in Lung CT Screening Images

by

Benjamin Geiger

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Computer Science
Department of Computer Science and Engineering
College of Engineering
University of South Florida

Major Professor: Lawrence Hall, Ph.D.
Dmitry Goldgof, Ph.D.
Sudeep Sarkar, Ph.D.

Date of Approval:
October 1, 2018

Keywords: Computer-Aided Diagnosis, Radiomics, Lung Cancer, Prognosis, Classification

DEDICATION

This work is dedicated to my grandfather, who taught me the importance of academic excellence; to my mother, who taught me the importance of following your dreams; and to my nephew, who I hope to teach the same.

ACKNOWLEDGMENTS

I would like to thank:

- Dr. Lawrence Hall, for guiding me and having confidence in me when I had none in myself.
- Dr. Dmitry Goldgof, for his assistance in all matters academic.
- Dr. Sudeep Sarkar, for his help in the clutch.
- Dr. Samuel Hawkins and Dmitry Cherezov, my teammates, without whom this research would have been for naught.
- The faculty and staff of the Computer Science and Engineering department of USF.
- The research staff at H. Lee Moffitt Cancer Center.
- My family and friends.

TABLE OF CONTENTS

LIST OF TABLES

ABSTRACT

Computed tomography (CT) imagery is an important weapon in the fight against lung cancer; various forms of lung cancer are routinely diagnosed from CT imagery. The growth of the suspect nodule is known to be a prognostic factor in the diagnosis of pulmonary cancer, but the change in other aspects of the nodule, such as its aspect ratio, density, spiculation, or other features usable for machine learning, may also provide prognostic information.

We hypothesized that adding combined feature information from multiple CT image sets separated in time could provide a more accurate determination of nodule malignancy. To this end, we combined data from multiple CT images for individual patients taken from the National Lung Screening Trial. The resulting dataset was compared to equivalent datasets featuring single CT images for each patient. Feature reduction and normalization was performed as is standard.

The highest accuracy achieved was 83.71% on a subset of features chosen by a combination of manual feature stability testing and the Correlation-based Feature Selection algorithm and classified by the Random Forests algorithm. The highest accuracy achieved with individual CT images was 81.00%, on a feature set consisting solely of the volume of the nodule in cubic centimeters.

CHAPTER 1

INTRODUCTION

By definition, cancer involves the uncontrolled division of cells, and by extension, the uncontrolled growth of tissues in the body. It stands to reason, then, that if we derive features from images of cancerous growths, the change in those features over time could be important. Changes in features over time could indicate malignancy more accurately than would the features at any single point in time.

Lung cancer is responsible for more deaths per year in the United States than any other form of cancer [1]. Typically, lung cancer is preceded by the development of pulmonary nodules, which can be detected with low-dose CT imagery. The ability of low-dose CT scans to detect pulmonary nodules was demonstrated via the National Lung Screening Trial (NLST) [24]. However, these nodules are frequently benign, and determining which nodules are malignant and which are benign is challenging at best.

There are many image features that have been developed by various researchers over the years with the intent of finding one or several that are indicative of malignancy. [2][27][18] These features can generally be grouped into three categories: *shape features*, or features derived from the shape of the nodule (as segmented either by a radiologist or by an automated process such as that developed by Dr. Yuhua Gu et al. [15]); *intensity features*, or features derived from the intensity of the pixels or voxels within the segmented nodule; and *texture features*, or features derived from the internal texture of the nodule or potentially the tissue immediately surrounding it.

Thus far, the overwhelming majority of features have been based on the images from a single CT scan. Studies have been performed on images from scans performed a matter of minutes apart, but these features are used primarily to determine the stability of the computed features, rather than reveal information about the nodules themselves [2][3].

An alternative is to combine features from multiple scans, taken months or years apart, to get a more informed idea of whether the nodule is changing (hence likely to be considered malignant) or static (more likely to be benign).

There are two primary approaches to be taken here. The first is to use the same features from two (or more) scans taken a significant time apart, and rely on the classifier to determine the connection between each pair of columns. Another is to create new features by taking the difference between the values in each feature for each pair of scans and using it as a new feature to be fed into the feature selection and classification pipeline.

The goal of this research is to determine whether this approach is feasible, and how much of an improvement (if any) the additional information provides. Some small but statistically significant improvement was found.

CHAPTER 2

RELATED WORK

2.1 Radiomics

According to Lambin *et al.*,

"The underlying hypothesis of Radiomics is that advanced image analysis on conventional and novel medical imaging [sic] could capture additional information not currently used, and more specifically, that genomic and proteomics patterns can be expressed in terms of macroscopic image-based features." (p. 3) [23]

The authors point out that medical imaging has improved in four distinct ways:

1. Improved hardware: from the early days of X-ray imaging, researchers have moved to single-slice computed tomography (CT) imagery, and from there to multiple-slice CT and combined CT and positron emission tomography (CT/PET). Magnetic resonance imaging (MRI) scans are also used. The resolution has also improved significantly, allowing researchers to see with more and more detail the contents of suspicious regions in the body.

2. Improved image agents: contrast agents can be used to highlight the regions of the body that are of interest, by tracing blood flow or substance absorption.

3. Improved standardization: "Historically, radiology has been a qualitiative science", writes Lambin, indicating previous reliance on human interpretation of radiological imagery (and the ad-hoc nature of much of the imagery collection process). Recently, there has been a trend toward standardized protocols, making radiology more reproducible and the collected imagery usable across individual cases (or locations).

4. Improved analysis: given the previous improvements, computer-aided detection (CAD) systems have improved as well, with advanced algorithms giving more details about detected regions.

A search of the literature has found the papers listed in Table 2.1.

2.1.1 Feature Extraction

Extraction of features from a CT image is a process with several steps.

The first step (after identification of a nodule, which must still, as of this writing, be performed by humans[1]) is nodule segmentation, or separating the nodule from its environment. This is often done automatically, but with a human verifying results. For instance, Dr. Yuhua Gu, et al., introduced an algorithm to segment a nodule [15], which combines a relatively straightforward region-growing algorithm with a clever iteration process.

Extraction of features from segmented lung nodule imagery is the realm of software packages such as Definiens Lung Tumor Analysis (LuTA) and 3D Slicer [13]. Features are continually being developed; Dhara *et al.* list several dozen features, some of which were unavailable at the time of the experiments discussed in Chapter 3 [9].

Generally, the segmented nodule is analyzed in isolation, but some work has been done on including other parts of the imagery. For instance, Dilger *et al.* include a portion of the parenchyma and obtain an AUC measure of 0.938 when distinguishing malignant from benign nodules, versus 0.918 using only the segmented nodule [10].

2.1.2 Feature Selection/Dimensionality Reduction

One issue that appears to be common in radiomics is the relatively small size of the data set on which classifiers are trained. Any given data set may only have a few dozen or a couple hundred patients, due to the difficulty of obtaining informed consent from patients and the relative rarity of certain forms of cancer. At the same time, image processing suites such as Definiens can generate hundreds of features based on Laws texture masks and wavelet decomposition; see Appendix A for

[1] Deep learning methods may be able to take over the nodule identification role, but as of the time of this writing, they are not able to do so with anything resembling reliability.

an example. Therefore, it is nearly mandatory that some form of feature selection or extraction be employed to reduce the dimensionality of the feature set.

Feature selection is an active area of research, with a goal of keeping the features that inform the classification while rejecting the ones that are irrelevant or redundant. Fundamentally, the better the correlation between the selected features and the class, the better the feature selection algorithm can be considered to be.

A classic example of a ranking feature selection algorithm is Relief, introduced in 1992 by Kenji Kira and Larry Rendell [20]. It was extended in 1994 by Igor Kononenko into six variants, labeled "Relief-A" through "Relief-F". The last of these variants is among the best-known feature selection algorithms; it extends the original Relief algorithm to support classification problems with more than two classes.

The LASSO operator, introduced in 1996 by Robert Tibshirani, performs both feature selection and data normalization on the input data [28]. LASSO is a refinement of the Nonnegative Garotte, introduced by Leo Breiman in 1995 [4].

Tools such as Principal Component Analysis, Independent Component Analysis, or Linear Discriminant Analysis have been used to combine the hundreds or thousands of features into a few that express most of the information in the original feature set.

2.1.2.1 Dimensionality Reduction of High-Dimensionality Data

Typically, feature selection is intended for data sets with high, but not extremely high, feature counts. As mentioned, a typical radiomic dataset may include only several dozen patients, but several hundred individual features. Genetic data sets are even more difficult, as the number of features may number in the millions (and, needless to say, even the largest data sets will only have a few thousand patients). The problem of sifting through these millions of features to find the ones that are most informative has grown alongside the preeminence of datasets that would require such algorithms.

For high-dimensional data, algorithms such as least-angle regression (LARS) show some promise. LARS, introduced by Efron *et al.* in 2004, is a refinement of the venerable stepwise regression algorithm; its authors assert that it is suitable for cases where $p >> n$. LARS is a generalization of both the LASSO and Forward Stagewise linear regression.

However, LARS suffers when features are correlated, which is increasingly likely with increasing dimensionality; this makes LARS' utility for high-dimensional data more limited without significant preprocessing.

2.1.3 Data Synthesis

Another approach to dealing with the paucity of data is to generate synthetic training data. For instance, the SMOTE algorithm [8] can be run on an existing training set to generate additional elements to train on, by taking existing elements pairwise and generating a new element somewhere "between" the two: taking the two elements as points in N-dimensional space, the algorithm creates a synthetic element somewhere on the line segment between those points.

2.1.4 Classification

Classic classification algorithms such as naïve Bayes, decision trees (*e.g.* C4.5), and support vector machines [7] continue to be the workhorses of the radiomics classification problem. However, new approaches are being developed.

Kuruvilla and Gunavathi applied neural networks to the classification problem in 2014 [22]. Using feed-forward and back-propagation networks along with statistical features derived from the entirety of the lung field, they claim to have a 93.3% accuracy in detecting lung cancer.

2.2 Time-Differential Features

As of the time of writing, the attempt to use time-differential features alongside radiomic techniques seems to be a fairly small niche. Few papers have discussed it, and none of those have been in the context of diagnosis. Instead, the only papers that have done so have been addressing prognosis of confirmed cancer cases based on changes due to treatment.

The idea of using multiple readings to extract additional information from a series of radiographic images appears to originate at least as far back as 2005, when Hoekstra *et al.* compared [^{18}F]-2-fluoro-2-deoxy-D-glucose positron emission tomography (^{18}FDG-PET) before and during induction chemotherapy (IC) [17]. Across 47 patients, the rate of change in the metabolic rate of glucose (rate of glucose consumption) showed prognostic value, with a hazard ratio of 1.95.

Van Elmpt *et al.* took ^{18}FDG-PET/CT scans of 34 non-small-cell lung cancer patients before radiotherapy began, and again during the second week of radiotherapy, measuring the CT volume and the standardized uptake value parameters. [29]

Carvalho *et al.* examine the use of delta radiomics with ^{18}FDG-PET exams, using it to predict the outcome of treatment. Specifically, the authors examined ^{18}FDG-PET scans taken before and during the second week of radiotherapy to predict survival rate. Model creation was performed on a set of 54 patients, and verification was performed on a set of 58 patients ranging from stage IIa to IIIb. Features included both shape and texture features, limited to those with sufficiently high intra-class correlation (ICC) on both test-retest and interobserver stability analysis (IOSA), and further reduced with the LASSO method. With this model, the authors obtained a C statistic (area under the ROC curve) of 0.58-0.61. [6]

Another group that has experimented with time-differential features is Fave *et al.*; their experiment is similar in nature to Carvalho's. They computed 62 features for each of 107 patients, including both shape and texture features, and then compared to dose fraction to determine correlation. Models were computed using leave-one-out verification, both for clinical-features-only and clinical-plus-delta-radiomics-features input sets, and for local-recurrence, distant-metastases, and overall-survival. For overall-survival, the addition of delta-radiomics features improved the C statistic from 0.52 to 0.62; for distant metastases it improved from 0.53 to 0.58, and did not improve at all for local-recurrence. Autocorrelation, kurtosis, and compactness were the features most commonly selected in the delta-radiomics form. [12]

Table 2.1 Existing literature.

Author	**Title**
Year	**Description**
National Lung Screening Research Team	Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening
2011	The National Lung Screening Trial was conducted to determine whether low-dose CT imagery was at least as effective as classic X-ray imagery at detecting lung cancer at an early stage, with screenings performed on 53,454 patients at high risk of lung cancer (each patient having a history of more than 30 pack-years of smoking). The patients undergoing screening CT exams had a mortality rate of 247 deaths per 100,000 person-years, compared to 309 deaths per 100,000 person-years for the X-ray patients, or a 20.0% reduction in mortality for the CT patients. Overall rate of death for the CT group was reduced by 6.7% compared to the X-ray group. This clearly shows that screening via low-dose CT reduces mortality due to lung cancer.

Table 2.1 (Continued)

Author	**Title**
Year	**Description**
Gu *et al.*	Automated delineation of lung tumors from CT images using a single click ensemble segmentation approach
2013	Dr. Gu *et al.* introduce a technique used in our analysis. Using an ensemble of automated segmentation algorithm runs with varying start points, the "single click ensemble segmentation" algorithm finds a stable and accurate segmentation for a given lung nodule, given only a single starting point (hence the name). This is beneficial for low-dose CT imagery, as borders are more difficult to ascertain automatically, and small variations in the border can result in disproportionate variations in the computed features, leading to varying results.
Lambin *et al.*	Radiomics: Extracting more information from medical images using advanced feature analysis
2012	This paper introduces the concept of radiomics. The authors cover the fundamental differences between radiomic approaches and others: primarily, the hypothesis that genomic and/or proteomic data can be inferred from imagery, and using the heterogeneity of the analyzed masses for additional predictive power.

Table 2.1 (Continued)

Author	**Title**
Year	**Description**
Fried *et al.*	Stage III Non–Small Cell Lung Cancer: Prognostic Value of FDG PET Quantitative Imaging Features Combined with Clinical Prognostic Factors
2016	The authors use ^{18}FDG-PET imagery to determine prognosis after treatment for NSCLC patients undergoing radiation treatment.
Kuruvilla and Gunavathi	Lung cancer classification using neural networks for CT images
2014	In this report, the authors use CT imagery passed into neural networks to determine malignancy of nodules. Using these networks, an accuracy of 93.3% was obtained. It stands to reason that these results can be further improved by introducing change descriptors into the analysis, but this will have to be a subject for further study.

Table 2.1 (Continued)

Author	**Title**
Year	**Description**
Hoekstra *et al.*	Prognostic Relevance of Response Evaluation Using [18F]-2-Fluoro-2-Deoxy-D-Glucose Positron Emission Tomography in Patients With Locally Advanced Non-Small-Cell Lung Cancer
2005	This is an early attempt at using change in tumors for additional data. Where our research was conducted with the goal of determining malignancy of nodules of unknown status, the authors of this paper focused their research on evaluation of radiation therapy on known-malignant tumors. Also, instead of CT imagery, the authors used ^{18}FDG-PET scans for source imagery. The only feature analyzed was the volume of each tumor.
Carvalho *et al.*	Early variation of FDG-PET radiomics features in NSCLC is related to overall survival - the "delta radiomics" concept
2016	Here we find an early mention of "delta radiomics". As with Hoekstra *et al.*, the authors use ^{18}FDG-PET imagery instead of CT images, and focus entirely on treatment efficacy rather than diagnosis. The concepts are present, though.

Table 2.1 (Continued)

Author	**Title**
Year	**Description**
Fave *et al.*	TU-D-207B-02: Delta-Radiomics: The Prognostic Value of Therapy-Induced Changes in Radiomics Features for Stage III Non-Small Cell Lung Cancer Patients
2016	In this study, the authors use delta-radiomics features in CT images to improve prognostic predictions for patients with Stage III non-small-cell lung cancer (NSCLC). This is, again, focused primarily on the evaluation of treatment efficacy rather than prediction of malignancy. The addition of delta-radiomics features improved the C-index of overall survival prediction from 0.52 to 0.62, with corresponding predictions of 0.53 to 0.58 for distant-metastases, but with no improvement for local recurrence.
Pearce *et al.*	Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study
2012	This paper highlights the risk of high-dose CT imagery for routine screening, primarily in children. While the absolute risk of leukemia or brain cancer is small (74 and 135, respectively, out of 176,000), a cumulative dose of 50-60 mSv triples the baseline risk, indicating a need for low-dose CT scans over riskier standard-dose scans. (For comparison, a typical chest CT scan has a dose of about 8 mSv.)

Table 2.1 (Continued)

Author	**Title**
Year	**Description**
Miglioretti *et al.*	The use of computed tomography in pediatrics and the associated radiation exposure and estimated cancer risk
2013	Pediatric CT imagery and its dosages and risks are studied here. According to the authors, between 3% and 8% of chest CT scans performed on patients between the ages of birth and 15 years had an effective dose of 20 mSv or higher. Pediatric CT scans are responsible for one solid cancer in girls for every 330-480 chest scans; the risk is lower for boys at 1,080-1,650 chest scans for each solid cancer. Reducing the highest quartile of doses to the median could prevent 43% of the 4,870 predicted cancers caused by pediatric CT scans each year.

Table 2.1 (Continued)

Author	**Title**
Year	**Description**
Smith-Bindman *et al.*	Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer
2009	The authors of this study examine the risk of cancer introduced from CT examinations. The CT scans studied are of more standard diagnostic dosages, ranging from 2 mSv for head scans to 31 mSv for multiphase abdominal and pelvic scans. Routine chest CTs had a median dosage of 8 mSv, equivalent to 117-119 chest X-ray scans. An estimated 1 in 270 women and 1 in 600 men who undergo a routine chest CT at age 40 will develop cancer due to that examination; this risk doubles at age 20 and is reduced by half at age 60. The authors suggest the use of lower and more standardized doses.
Dhara *et al.*	A Combination of Shape and Texture Features for Classification of Pulmonary Nodules in Lung CT Images
2016	The authors use a similar methodology to that used in this work, with a single time point. Specifically, they combine shape and texture features with feature selection and traditional classification (specifically, support vector machines, as opposed to a variety of algorithms used in this work). Their classification is more nuanced, with five "levels of malignancy" (classes) instead of the binary benign/malignant division used in our work.

Table 2.1 (Continued)

Author	**Title**
Year	**Description**
Dilger *et al.*	Improved pulmonary nodule classification utilizing quantitative lung parenchyma features
2015	In this paper, the authors examine the use of features derived from the CT imagery of the parenchyma (functional tissue) immediately surrounding a specified lung nodule to determine the malignancy of that nodule. The CT imagery was derived from high-resolution (and therefore high-dose) scans of 50 nodules; including parenchymal features improved the classification AUC from 0.918 (nodule features only) to 0.938 (including parenchymal features).
Van Elmpt *et al.*	Response Assessment Using 18F-FDG PET Early in the Course of Radiotherapy Correlates with Survival in Advanced-Stage Non–Small Cell Lung Cancer
2012	The authors of this study examine the ^{18}FDG-PET/CT scans of 34 non-small-cell lung cancer patients before and after two weeks of radiotherapy, using CT volume and standardized uptake value parameters to determine prognosis.

CHAPTER 3

EXPERIMENTS

Several experiments were performed to determine the efficacy of change descriptors for lung nodule classification.

3.1 National Lung Screening Trial, Delta Features

For this experiment, we used a subset of the low-dose CT imagery from the National Lung Screening Trial (NLST)[24].

The NLST compared the effectiveness of low-dose CT imagery to that of traditional X-ray chest radiography. Patients were at high risk of lung cancer; all patients were between the ages of 55 and 74 at the time they began participating in the study, and had been a smoker for at least 30 pack-years. If they were no longer smokers, then they were required to have quit within the 15 years prior to the beginning of the study to be eligible to participate. 53,454 patients were enrolled, with 26,722 of them undergoing low-dose CT screening (where "low-dose" is an average effective dose of 1.5 mSv, versus the approximately 8 mSv average effective dose delivered in a typical chest CT screening) and the remainder undergoing posteroanterior chest radiography, either via screen-film methods or digital methods.

Each patient was required to undergo up to three screenings. The first reading ('time 0'), was given shortly after the patient was randomly assigned to a group. Subsequent screenings ('time 1' and 'time 2') were at approximately one-year intervals afterward. Patients with a diagnosis of lung cancer were not screened again after the diagnosis.

Of the group receiving low-dose CT scans, we focused on two pairs of cohorts. Cohort IC1 ("Incident Cohort 1") were patients who had pulmonary nodules detected during the first screening and in whom lung cancer was diagnosed from the second screening ('time 1') of the same nodule. Cohort NC1 ("Non-incident Cohort 1") were patients who had pulmonary nodules detected during

all three screenings, but who were matched based on their attributes with the patients in cohort IC1. Cohort IC2 were patients for whom nodules were detected during the first two screenings, and cancer was diagnosed due to the third screening ('time 2'). Cohort NC2 were patients, also with pulmonary nodules in all three screenings, who were paired with the patients in cohort IC2 in the same way as with cohorts NC1 and IC1. We considered patients in cohorts IC1 and IC2 to be 'positive', and patients in cohorts NC1 and NC2 'negative', for the purpose of classification. This experiment was performed on a subset of these cohorts, patients for whom the imagery was available and for whom the same nodules had been segmented across each screening. Due to the labor-intensive nature of that process, there were 476 patients in these subsets: 77 in cohort IC1, 83 in cohort IC2, 172 in cohort NC1, and 144 in cohort NC2. Cohorts IC1 and NC1 were combined to form "cohort 1" and cohorts IC2 and NC2 were likewise combined to form "cohort 2".

Image features were extracted from the CT images with the Lung Tumor Analysis application on the Definiens Developer XD © platform. This yielded 219 features, listed in appendix A.

In addition to the full set of 219 features, classification was performed on three subsets:

1. *NC1Stable:* features which, on the readings of patients in cohort NC1, had a Concordance Correlation Coefficient (CCC) of greater than 0.6; this indicates that the features are stable across multiple readings and therefore less likely to be noise. This subset contained 37 features.

2. *RiderC95:* features which, when generated from the Reference Imaging Database to Evaluate Response (RIDER) image set, had a CCC of greater than 0.95 under both manual segmentation and a one-click ensemble segmentation algorithm[15]. The RIDER data set contains pairs of images taken fifteen minutes apart, a "test-retest" strategy, to make the detection of noise in features more likely. This subset contains 23 features.

3. *Volume:* the "Volume [cm^3]" feature considered by itself. This was intended to rule out the possibility that just the change in volume was responsible for the improved results.

Each data set was split into three screenings: *Diagnosis*, the scan taken at the time of diagnosis (time 1 for cohort 1 and time 2 for cohort 2); *Prior*, the scan taken approximately one year prior to diagnosis (time 0 for cohort 1 and time 1 for cohort 2); and *Delta*, which is a straightforward arithmetic difference of the Prior features from the Diagnosis features.

The values in each feature were normalized by linearly scaling and offsetting the feature values such that the range for each feature was $[-1, 1]$ in cohort 1. The same scale and offset were made to the corresponding feature in cohort 2, despite some of the resulting values ending up outside the $[-1, 1]$ range.

All classification was performed via the Weka[16] data mining suite.

The following classifiers were tried:

- Naïve Bayes[19]
- J48 (C4.5 decision trees)[26]
- Random Forests[5] with 200 trees, selecting from $log_2(f)$ features each tree (where f is the number of features)
- Support vector machines (LibSVM)[7], with C and γ tuned via grid search
  - Linear kernel
  - Radial basis function kernel

3.1.1 Results

For these datasets, guessing the majority class gives an accuracy of 65.16% (77 'positive', 144 'negative'). With only two exceptions, the listed classifiers gave an accuracy higher than guessing the majority.

The highest accuracy, 83.71%, was from the NC1Stable feature subset, using the Random Forests classifier with 15 features selected by CFS. In general, classification on the Delta screenings gave higher accuracy than Diagnosis alone; this improvement was statistically significant as determined by the Wilcoxon signed-rank test ($p < 0.05$). The exception is the volume-only data set, as there were insufficient data for the test to be accurate.

In most cases, Random Forests outperformed the other classifiers; the improvement in performance tends to increase as the number of features increases. Conversely, Random Forests gave only mediocre performance on the Volume subset, as it only had one feature to work with.

Due to the need for a hyperparameter optimization step (executed via grid search on cost and γ), the SVM classifiers were slower than the others by orders of magnitude. While they did

Table 3.1 Accuracy and AUC for experiments on all features at single time points.

Delta					
Feat. Sel.	**Naïve Bayes**	**J48**	**Random Forests**	**SVM Linear**	**SVM RBF**
All Features	71.95% (0.704)	73.76% (0.622)	81.00% (0.825)	75.57% (0.731)	77.38% (0.764)
CFS	71.95% (0.705)	74.21% (0.634)	81.00% (0.826)	75.57% (0.731)	77.38% (0.764)
CFS (5)	78.73% (0.797)	80.54% (0.742)	77.38% (0.772)	76.02% (0.749)	77.83% (0.768)
CFS (10)	78.28% (0.804)	80.54% (0.742)	80.54% (0.807)	76.02% (0.751)	77.83% (0.807)
CFS (15)	78.28% (0.807)	80.54% (0.749)	81.45% (0.824)	78.28% (0.759)	80.09% (0.830)
CFS (20)	78.28% (0.788)	80.54% (0.749)	**81.90% (0.840)**	76.02% (0.725)	81.00% (0.810)
Relief-F (5)	75.11% (0.770)	76.47% (0.712)	73.30% (0.772)	74.66% (0.701)	74.66% (0.771)
Relief-F (10)	77.83% (0.782)	75.11% (0.666)	79.64% (0.787)	75.57% (0.704)	75.57% (0.795)
Relief-F (15)	78.73% (0.788)	76.92% (0.716)	79.19% (0.805)	76.92% (0.702)	76.92% (0.794)
Relief-F (20)	78.28% (0.787)	76.47% (0.710)	80.54% (0.806)	76.47% (0.712)	78.28% (0.810)
Diagnosis					
Feat. Sel.	**Naïve Bayes**	**J48**	**Random Forests**	**SVM Linear**	**SVM RBF**
All Features	77.83% (0.708)	76.47% (0.693)	**79.19% (0.808)**	77.38% (0.792)	75.57% (0.763)
CFS	77.83% (0.708)	76.02% (0.682)	77.38% (0.813)	77.38% (0.792)	75.57% (0.763)
CFS (5)	76.02% (0.759)	73.76% (0.669)	76.47% (0.788)	76.02% (0.759)	74.66% (0.752)
CFS (10)	76.02% (0.790)	71.04% (0.630)	78.28% (0.810)	75.11% (0.819)	75.57% (0.783)
CFS (15)	74.66% (0.746)	75.11% (0.704)	78.28% (0.815)	75.57% (0.818)	76.92% (0.769)
CFS (20)	74.66% (0.732)	74.21% (0.666)	78.73% (0.801)	75.57% (0.806)	74.66% (0.766)
Relief-F (5)	71.04% (0.747)	76.02% (0.715)	72.40% (0.776)	71.95% (0.741)	75.11% (0.789)
Relief-F (10)	73.30% (0.747)	71.04% (0.728)	75.11% (0.801)	76.47% (0.768)	77.83% (0.810)
Relief-F (15)	75.11% (0.714)	71.95% (0.681)	76.92% (0.790)	77.83% (0.780)	75.57% (0.758)
Relief-F (20)	76.02% (0.768)	71.95% (0.681)	75.57% (0.795)	77.38% (0.785)	78.28% (0.798)
Prior					
Feat. Sel.	**Naïve Bayes**	**J48**	**Random Forests**	**SVM Linear**	**SVM RBF**
All Features	69.23% (0.680)	66.97% (0.621)	73.30% (0.777)	71.49% (0.758)	69.68% (0.770)
CFS	69.23% (0.680)	66.52% (0.611)	72.85% (0.782)	71.49% (0.758)	69.68% (0.770)
CFS (5)	71.49% (0.649)	65.61% (0.638)	69.68% (0.732)	71.49% (0.660)	68.33% (0.686)
CFS (10)	73.30% (0.665)	67.42% (0.626)	73.30% (0.775)	71.49% (0.712)	71.49% (0.723)
CFS (15)	71.95% (0.690)	67.87% (0.611)	73.30% (0.767)	**75.11% (0.757)**	72.85% (0.752)
CFS (20)	72.40% (0.679)	69.23% (0.632)	72.40% (0.767)	74.21% (0.743)	72.40% (0.724)
Relief-F (5)	70.59% (0.723)	68.33% (0.645)	67.87% (0.739)	69.23% (0.677)	66.06% (0.707)
Relief-F (10)	67.87% (0.688)	69.23% (0.587)	70.59% (0.766)	70.59% (0.735)	71.95% (0.754)
Relief-F (15)	72.85% (0.722)	68.78% (0.561)	71.49% (0.803)	73.76% (0.687)	71.04% (0.718)
Relief-F (20)	69.68% (0.721)	69.23% (0.561)	73.76% (0.794)	74.21% (0.714)	73.76% (0.762)

Table 3.2 Accuracy and AUC for experiments on the NC1Stable feature subset at single time points.

Delta					
Feat. Sel.	**Naïve Bayes**	**J48**	**Random Forests**	**SVM Linear**	**SVM RBF**
All Features	78.73% (0.758)	75.57% (0.671)	82.81% (0.818)	78.28% (0.736)	77.83% (0.730)
CFS	78.73% (0.758)	76.92% (0.719)	82.81% (0.811)	78.28% (0.736)	77.83% (0.730)
CFS (5)	77.38% (0.762)	80.54% (0.742)	79.19% (0.781)	74.66% (0.743)	77.38% (0.742)
CFS (10)	78.73% (0.788)	80.54% (0.749)	81.45% (0.809)	79.19% (0.770)	79.19% (0.816)
CFS (15)	78.73% (0.779)	78.73% (0.697)	**83.71% (0.814)**	78.73% (0.749)	78.73% (0.805)
CFS (20)	78.73% (0.795)	78.73% (0.697)	82.81% (0.828)	77.83% (0.750)	79.64% (0.784)
Relief-F (5)	76.92% (0.775)	76.02% (0.699)	76.92% (0.747)	79.19% (0.755)	75.11% (0.752)
Relief-F (10)	76.47% (0.776)	76.47% (0.694)	77.83% (0.769)	77.83% (0.737)	77.83% (0.723)
Relief-F (15)	78.28% (0.782)	76.02% (0.681)	78.28% (0.796)	77.83% (0.739)	78.28% (0.745)
Relief-F (20)	78.73% (0.792)	75.57% (0.655)	82.35% (0.822)	77.38% (0.743)	78.28% (0.774)
Diagnosis					
Feat. Sel.	**Naïve Bayes**	**J48**	**Random Forests**	**SVM Linear**	**SVM RBF**
All Features	78.73% (0.778)	75.11% (0.718)	79.19% (0.822)	76.47% (0.795)	76.47% (0.813)
CFS	78.73% (0.778)	75.57% (0.719)	78.28% (0.810)	76.47% (0.795)	76.47% (0.813)
CFS (5)	76.02% (0.825)	79.64% (0.741)	78.28% (0.765)	75.57% (0.814)	77.83% (0.806)
CFS (10)	76.47% (0.782)	**80.09% (0.746)**	78.28% (0.762)	76.47% (0.814)	79.64% (0.782)
CFS (15)	76.92% (0.802)	78.28% (0.734)	79.64% (0.785)	77.38% (0.811)	77.38% (0.788)
CFS (20)	78.73% (0.802)	77.38% (0.740)	79.19% (0.781)	77.83% (0.802)	78.73% (0.790)
Relief-F (5)	73.30% (0.771)	71.95% (0.691)	76.92% (0.774)	73.30% (0.782)	73.76% (0.769)
Relief-F (10)	76.02% (0.821)	70.59% (0.683)	74.66% (0.777)	72.85% (0.787)	74.21% (0.765)
Relief-F (15)	72.40% (0.823)	71.04% (0.690)	76.02% (0.809)	74.66% (0.822)	75.11% (0.828)
Relief-F (20)	78.73% (0.817)	78.73% (0.727)	79.19% (0.809)	78.28% (0.825)	73.76% (0.818)
Prior					
Feat. Sel.	**Naïve Bayes**	**J48**	**Random Forests**	**SVM Linear**	**SVM RBF**
All Features	73.76% (0.718)	67.87% (0.635)	74.21% (0.766)	69.23% (0.693)	71.95% (0.684)
CFS	73.76% (0.718)	70.14% (0.647)	73.76% (0.756)	69.23% (0.693)	71.95% (0.684)
CFS (5)	73.76% (0.658)	65.61% (0.645)	69.68% (0.706)	72.40% (0.633)	71.95% (0.632)
CFS (10)	71.49% (0.708)	65.61% (0.613)	73.30% (0.719)	70.14% (0.615)	70.14% (0.669)
CFS (15)	73.30% (0.691)	67.42% (0.601)	72.85% (0.744)	68.78% (0.605)	67.42% (0.619)
CFS (20)	72.85% (0.688)	67.42% (0.601)	**74.21% (0.751)**	70.14% (0.631)	67.87% (0.686)
Relief-F (5)	69.68% (0.647)	66.97% (0.550)	66.97% (0.681)	65.16% (0.578)	67.42% (0.694)
Relief-F (10)	68.33% (0.698)	66.97% (0.537)	73.76% (0.771)	67.42% (0.627)	72.85% (0.714)
Relief-F (15)	71.49% (0.769)	67.87% (0.678)	71.49% (0.762)	71.04% (0.648)	66.97% (0.716)
Relief-F (20)	72.85% (0.732)	68.33% (0.668)	73.76% (0.759)	71.49% (0.676)	71.95% (0.671)

Table 3.3 Accuracy and AUC for experiments on the RiderC95 feature subset at single time points.

Delta					
Feat. Sel.	**Naïve Bayes**	**J48**	**Random Forests**	**SVM Linear**	**SVM RBF**
All Features	77.83% (0.777)	77.38% (0.733)	81.00% (0.820)	77.83% (0.810)	81.00% (0.809)
CFS	77.83% (0.777)	77.38% (0.733)	81.90% (0.836)	77.83% (0.810)	81.00% (0.809)
CFS (5)	77.38% (0.776)	77.83% (0.703)	78.73% (0.790)	74.66% (0.770)	76.92% (0.756)
CFS (10)	77.83% (0.786)	76.92% (0.709)	81.45% (0.808)	80.09% (0.798)	81.00% (0.808)
CFS (15)	77.83% (0.800)	77.38% (0.716)	**82.35% (0.825)**	78.28% (0.785)	80.09% (0.809)
CFS (20)	77.83% (0.797)	77.38% (0.716)	81.45% (0.815)	78.28% (0.793)	79.64% (0.786)
Relief-F (5)	76.47% (0.791)	77.38% (0.700)	81.00% (0.781)	77.38% (0.777)	81.00% (0.836)
Relief-F (10)	76.92% (0.798)	79.19% (0.737)	**82.35% (0.811)**	79.19% (0.807)	79.19% (0.815)
Relief-F (15)	77.38% (0.798)	77.38% (0.728)	81.90% (0.828)	78.28% (0.806)	79.19% (0.805)
Relief-F (20)	77.83% (0.790)	77.38% (0.728)	81.90% (0.808)	78.73% (0.794)	79.19% (0.795)
Diagnosis					
Feat. Sel.	**Naïve Bayes**	**J48**	**Random Forests**	**SVM Linear**	**SVM RBF**
All Features	77.38% (0.777)	75.11% (0.685)	**80.54% (0.801)**	80.09% (0.855)	79.19% (0.797)
CFS	77.38% (0.777)	74.21% (0.666)	**80.54% (0.800)**	80.09% (0.855)	79.19% (0.797)
CFS (5)	75.11% (0.829)	73.76% (0.740)	76.02% (0.765)	76.02% (0.828)	77.38% (0.810)
CFS (10)	75.57% (0.812)	75.11% (0.738)	78.28% (0.777)	77.38% (0.820)	77.83% (0.794)
CFS (15)	77.38% (0.767)	74.66% (0.739)	79.19% (0.788)	77.83% (0.824)	76.92% (0.786)
CFS (20)	77.38% (0.777)	75.57% (0.712)	79.19% (0.793)	79.19% (0.858)	79.64% (0.797)
Relief-F (5)	73.76% (0.798)	73.76% (0.725)	75.11% (0.781)	77.83% (0.817)	77.38% (0.800)
Relief-F (10)	77.83% (0.797)	73.76% (0.661)	77.83% (0.820)	78.73% (0.842)	80.09% (0.802)
Relief-F (15)	79.19% (0.789)	73.76% (0.636)	77.83% (0.814)	80.09% (0.847)	**80.54% (0.802)**
Relief-F (20)	77.38% (0.781)	75.57% (0.682)	78.28% (0.805)	80.09% (0.854)	80.09% (0.797)
Prior					
Feat. Sel.	**Naïve Bayes**	**J48**	**Random Forests**	**SVM Linear**	**SVM RBF**
All Features	75.57% (0.716)	66.52% (0.622)	72.85% (0.783)	**76.92% (0.769)**	**76.92% (0.761)**
CFS	75.57% (0.716)	68.33% (0.634)	73.76% (0.787)	**76.92% (0.769)**	**76.92% (0.761)**
CFS (5)	74.21% (0.708)	62.90% (0.599)	68.78% (0.722)	73.76% (0.677)	71.49% (0.669)
CFS (10)	76.47% (0.735)	66.97% (0.643)	73.30% (0.758)	74.21% (0.686)	72.85% (0.657)
CFS (15)	76.02% (0.763)	66.97% (0.660)	75.57% (0.781)	76.02% (0.760)	72.85% (0.798)
CFS (20)	75.57% (0.744)	69.23% (0.648)	76.47% (0.778)	76.92% (0.764)	74.66% (0.783)
Relief-F (5)	73.30% (0.736)	67.42% (0.688)	70.14% (0.746)	71.95% (0.733)	72.85% (0.728)
Relief-F (10)	75.11% (0.708)	68.33% (0.683)	73.76% (0.769)	76.02% (0.730)	74.66% (0.723)
Relief-F (15)	74.66% (0.718)	66.06% (0.650)	73.76% (0.789)	76.47% (0.768)	73.76% (0.792)
Relief-F (20)	74.66% (0.716)	66.52% (0.612)	74.21% (0.781)	76.02% (0.763)	76.02% (0.753)

Table 3.4 Accuracy and AUC for experiments on feature subset consisting solely of volume in cubic centimeters at single time points.

Delta				
Naïve Bayes	**J48**	**Random Forests**	**SVM Linear**	**SVM RBF**
74.66% (0.768)	**80.54% (0.763)**	77.38% (0.798)	71.95% (0.738)	75.57% (0.775)
Diagnosis				
Naïve Bayes	**J48**	**Random Forests**	**SVM Linear**	**SVM RBF**
72.40% (0.779)	**81.00% (0.758)**	76.47% (0.725)	75.11% (0.751)	78.73% (0.750)
Prior				
Naïve Bayes	**J48**	**Random Forests**	**SVM Linear**	**SVM RBF**
71.04% (0.676)	67.42% (0.632)	69.68% (0.656)	**73.76% (0.659)**	72.85% (0.657)

provide the best performance in some cases (specifically, RiderC95 features with Prior data), the improvement is unlikely to justify the runtime. This is offset, however, by the ease with which the grid search process could be parallelized given a framework that supported it. As each point in the grid search (at least without grid extension) is an independent classification, they could be performed on different cores or different nodes with minimal difficulty.

3.2 National Lung Screening Trial, Delta Features With Single-Time Features

This experiment was an extension of the experiment described in section 3.1. The Prior, Delta, and Diagnosis features were the same as in that experiment.

For this run, in addition to considering each set separately, we evaluated various classifiers on two combinations of sets: Delta+Diagnosis and Prior+Delta+Diagnosis. The goal is to determine whether combining single-time features with delta features could lead to improved performance over delta features alone.

Feature subsets for each data set were combined. Hence, each subset had twice as many features for Delta+Diagnosis and three times as many features for Prior+Delta+Diagnosis. Feature selection was performed with the same number of features; a feature selection algorithm that retrieves 10 features retrieved 10 features no matter the input feature count. Again, the exception is Volume, as no feature selection was performed on it. However, feature selection *could* have been performed, as there were two and three features, respectively, for Delta+Diagnosis and Prior+Delta+Diagnosis. This was not explored as the feature selectors generally returned many

more than three features. Also, other than the volume features from Prior and Delta, any other combination of features has already been analyzed, so there is little point to investigating them again.

3.2.1 Results

In general, combining both delta features and single-time features typically does not yield significant benefits beyond simply using delta features. In fact, comparing Table 3.5 to Table 3.1, Table 3.6 to Table 3.2, and Table 3.7 to Table 3.3 reveals that the addition of single-time features to delta features tends to result in degraded performance relative to delta features on their own with most classifiers and feature selectors tested.

For all features (Table 3.5), the combination of Delta and Diagnosis features resulted in an accuracy of 81.00%, versus 81.90% for Delta alone. NC1Stable (Table 3.6) saw a smaller drop, from 83.71% with Delta alone to 83.26% with Delta and Diagnosis. The best result for the RiderC95 subset (Table 3.7) remains the same, at 82.35%.

For the volume subset, the addition of both Prior and Diagnosis features yield improved performance: 81.90% versus 81.00% for Diagnosis alone or Delta with Diagnosis or 80.54% for Delta alone. See Table 3.8.

Table 3.5 Accuracy and AUC for experiments on all features at combinations of time points.

Delta+Diagnosis					
Feat. Sel.	**Naïve Bayes**	**J48**	**Random Forests**	**SVM Linear**	**SVM RBF**
All Features	75.57% (0.716)	69.68% (0.586)	80.09% (0.834)	75.57% (0.801)	76.47% (0.821)
CFS	75.57% (0.716)	73.76% (0.614)	81.00% (0.837)	75.57% (0.801)	76.47% (0.821)
CFS (5)	77.38% (0.777)	79.64% (0.756)	79.19% (0.803)	74.66% (0.763)	76.02% (0.737)
CFS (10)	78.28% (0.798)	78.73% (0.725)	79.19% (0.835)	76.47% (0.772)	76.92% (0.826)
CFS (15)	78.73% (0.781)	77.38% (0.679)	79.19% (0.830)	75.57% (0.782)	77.38% (0.797)
CFS (20)	77.38% (0.778)	75.11% (0.692)	78.73% (0.850)	77.83% (0.819)	
Relief-F (5)	72.40% (0.661)	72.40% (0.698)	76.02% (0.784)	71.95% (0.687)	76.47% (0.758)
Relief-F (10)	72.40% (0.669)	70.59% (0.714)	77.38% (0.823)	71.95% (0.716)	71.95% (0.682)
Relief-F (15)	71.95% (0.671)	73.30% (0.673)	76.47% (0.813)	71.95% (0.752)	71.95% (0.683)
Relief-F (20)	71.95% (0.679)	74.21% (0.739)	76.02% (0.803)	72.40% (0.770)	71.95% (0.697)
Prior+Delta+Diagnosis					
Feat. Sel.	**Naïve Bayes**	**J48**	**Random Forests**	**SVM Linear**	**SVM RBF**
All Features	76.02% (0.717)	77.83% (0.745)	81.00% (0.840)	77.83% (0.838)	73.30% (0.757)
CFS	75.57% (0.712)	76.02% (0.753)	80.54% (0.836)	77.83% (0.838)	73.30% (0.757)
CFS (5)	77.38% (0.777)	79.64% (0.756)	79.19% (0.803)	74.66% (0.763)	76.02% (0.737)
CFS (10)	78.73% (0.801)	78.28% (0.715)	77.38% (0.836)	73.76% (0.764)	77.38% (0.837)
CFS (15)	79.19% (0.763)	76.02% (0.633)	79.19% (0.832)	75.11% (0.773)	76.92% (0.810)
CFS (20)	79.64% (0.774)	81.00% (0.741)	80.09% (0.844)	77.38% (0.806)	76.02% (0.791)
Relief-F (5)	71.95% (0.739)	71.04% (0.695)	76.02% (0.808)	71.95% (0.764)	71.95% (0.723)
Relief-F (10)	71.95% (0.703)	71.04% (0.648)	77.38% (0.797)	71.04% (0.762)	75.57% (0.766)
Relief-F (15)	71.95% (0.680)	70.14% (0.631)	77.83% (0.807)	71.95% (0.749)	71.95% (0.684)
Relief-F (20)	71.95% (0.699)	72.85% (0.695)	76.92% (0.808)	71.95% (0.782)	71.95% (0.772)

Table 3.6 Accuracy and AUC for experiments on the NC1Stable feature subsets at combinations of time points.

Delta+Diagnosis					
Feat. Sel.	**Naïve Bayes**	**J48**	**Random Forests**	**SVM Linear**	**SVM RBF**
All Features	78.28% (0.774)	77.83% (0.719)	81.45% (0.832)	80.09% (0.790)	81.00% (0.849)
CFS	78.28% (0.774)	78.28% (0.721)	81.90% (0.833)	80.09% (0.790)	81.00% (0.849)
CFS (5)	77.83% (0.757)	79.64% (0.756)	77.83% (0.804)	77.83% (0.791)	77.83% (0.774)
CFS (10)	79.19% (0.790)	81.00% (0.762)	79.64% (0.812)	77.38% (0.828)	78.28% (0.806)
CFS (15)	79.19% (0.795)	81.00% (0.740)	83.26% (0.832)	80.09% (0.786)	80.54% (0.790)
CFS (20)	78.28% (0.786)	79.64% (0.742)	82.35% (0.820)	81.00% (0.783)	82.35% (0.814)
Relief-F (5)	72.85% (0.720)	69.68% (0.583)	69.68% (0.715)	70.59% (0.724)	67.87% (0.666)
Relief-F (10)	73.30% (0.773)	71.95% (0.717)	76.02% (0.819)	74.66% (0.777)	77.83% (0.816)
Relief-F (15)	71.49% (0.748)	74.21% (0.729)	77.38% (0.802)	74.66% (0.750)	81.00% (0.806)
Relief-F (20)	74.21% (0.813)	68.78% (0.642)	79.19% (0.797)	77.38% (0.771)	76.92% (0.786)
Prior+Delta+Diagnosis					
Feat. Sel.	**Naïve Bayes**	**J48**	**Random Forests**	**SVM Linear**	**SVM RBF**
All Features	78.73% (0.768)	78.73% (0.672)	82.35% (0.832)	79.64% (0.791)	79.64% (0.856)
CFS	78.73% (0.768)	81.45% (0.715)	82.35% (0.849)	79.64% (0.791)	79.64% (0.856)
CFS (5)	76.92% (0.782)	78.73% (0.723)	78.73% (0.781)	76.92% (0.764)	76.92% (0.777)
CFS (10)	78.73% (0.787)	80.54% (0.744)	79.64% (0.821)	78.73% (0.813)	80.09% (0.816)
CFS (15)	79.19% (0.804)	82.35% (0.775)	82.81% (0.839)	81.45% (0.803)	81.90% (0.818)
CFS (20)	79.19% (0.812)	83.26% (0.798)	81.45% (0.833)	80.54% (0.804)	80.54% (0.820)
Relief-F (5)	69.68% (0.687)	67.42% (0.578)	69.68% (0.723)	69.23% (0.681)	67.42% (0.699)
Relief-F (10)	72.40% (0.803)	71.95% (0.687)	79.19% (0.798)	71.95% (0.793)	71.95% (0.794)
Relief-F (15)	72.40% (0.809)	71.95% (0.687)	76.92% (0.799)	72.40% (0.778)	71.95% (0.773)
Relief-F (20)	71.95% (0.809)	68.78% (0.653)	77.38% (0.806)	73.30% (0.712)	71.95% (0.755)

Table 3.7 Accuracy and AUC for experiments on the RiderC95 feature subset at combinations of time points.

Delta+Diagnosis					
Feat. Sel.	**Naïve Bayes**	**J48**	**Random Forests**	**SVM Linear**	**SVM RBF**
All Features	78.28% (0.778)	76.47% (0.711)	80.54% (0.842)	81.90% (0.834)	81.00% (0.805)
CFS	78.28% (0.778)	76.47% (0.711)	81.00% (0.847)	81.90% (0.834)	81.00% (0.805)
CFS (5)	78.73% (0.750)	81.00% (0.752)	78.73% (0.774)	76.47% (0.758)	76.47% (0.759)
CFS (10)	78.73% (0.822)	77.83% (0.721)	80.54% (0.816)	80.09% (0.829)	79.64% (0.821)
CFS (15)	78.28% (0.814)	76.92% (0.759)	80.54% (0.824)	79.64% (0.828)	80.09% (0.828)
CFS (20)	78.28% (0.809)	75.57% (0.732)	81.90% (0.830)	80.54% (0.815)	81.00% (0.821)
Relief-F (5)	77.38% (0.804)	74.21% (0.718)	77.38% (0.794)	76.92% (0.797)	78.28% (0.791)
Relief-F (10)	78.73% (0.820)	77.38% (0.672)	81.00% (0.834)	80.09% (0.859)	80.54% (0.821)
Relief-F (15)	78.73% (0.827)	76.47% (0.658)	80.09% (0.835)	81.45% (0.825)	81.00% (0.847)
Relief-F (20)	78.73% (0.812)	75.57% (0.642)	79.64% (0.818)	81.00% (0.825)	81.45% (0.866)
Prior+Delta+Diagnosis					
Feat. Sel.	**Naïve Bayes**	**J48**	**Random Forests**	**SVM Linear**	**SVM RBF**
All Features	78.73% (0.769)	78.28% (0.700)	80.54% (0.845)	80.54% (0.827)	80.09% (0.814)
CFS	78.73% (0.769)	79.19% (0.732)	81.45% (0.843)	80.54% (0.827)	80.09% (0.814)
CFS (5)	78.73% (0.750)	81.00% (0.752)	78.73% (0.774)	76.47% (0.758)	76.47% (0.759)
CFS (10)	77.83% (0.812)	81.45% (0.759)	81.00% (0.823)	77.83% (0.818)	80.09% (0.824)
CFS (15)	78.73% (0.824)	76.92% (0.697)	81.00% (0.838)	81.00% (0.822)	81.90% (0.842)
CFS (20)	78.28% (0.821)	76.47% (0.728)	80.54% (0.835)	81.45% (0.819)	81.90% (0.870)
Relief-F (5)	79.19% (0.847)	74.66% (0.721)	76.02% (0.807)	78.73% (0.785)	78.73% (0.788)
Relief-F (10)	79.64% (0.824)	74.21% (0.727)	76.02% (0.820)	79.64% (0.808)	79.64% (0.822)
Relief-F (15)	79.19% (0.824)	76.47% (0.698)	76.92% (0.830)	82.35% (0.831)	80.54% (0.837)
Relief-F (20)	79.19% (0.825)	76.47% (0.724)	79.64% (0.832)	81.45% (0.832)	81.45% (0.805)

Table 3.8 Accuracy and AUC for experiments on feature subset consisting solely of volume in cubic centimeters at combinations of time points.

Delta + Diagnosis				
Naïve Bayes	**J48**	**Random Forests**	**SVM Linear**	**SVM RBF**
75.11% (0.796)	81.00% (0.761)	77.83% (0.749)	76.02% (0.756)	80.09% (0.764)
Prior + Delta + Diagnosis				
Naïve Bayes	**J48**	**Random Forests**	**SVM Linear**	**SVM RBF**
76.02% (0.793)	81.00% (0.761)	80.09% (0.761)	74.66% (0.755)	81.90% (0.792)

CHAPTER 4

CONCLUSIONS

4.1 Results

With the data set and classifiers available, the highest accuracy increased from 81.00% to 83.71%, and the highest AUC increased from 0.858 to 0.859. The results are statistically significant, though small, via the Wilcoxon signed-rank test ($p < 0.05$).

These results show that the introduction of change descriptors can, under certain circumstances, improve classification accuracy. While we would certainly have preferred a greater improvement, the technique can still be improved.

4.2 Future Directions

Several potential research directions can be obtained from this work.

The most obvious is to introduce more sophisticated descriptors, such as using a weighted combination instead of a linear difference. These weights could be determined per-feature or a single set of weights could be determined for the entire feature set, or could even be produced from other features, to encode the relationships between them.

If more than two scans are available, the additional information could be used to determine the acceleration of changes. For lung cancer, however, this is unlikely for at least the foreseeable future as CT scans are both expensive and, ironically, increase the risk of cancer [25]. The National Lung Screening Trial has shown that low-dose CT screening improves outcomes relative to chest X-rays, but even low-dose CT screening introduces some risk; physicians are advised to limit the number and dosage of CT screenings to the minimum necessary [11].

LIST OF REFERENCES

[1] American Cancer Society. Cancer facts & figures 2013. Atlanta, 2013.

[2] Yoganand Balagurunathan, Yuhua Gu, Hua Wang, Virendra Kumar, Olya Grove, Sam Hawkins, Jongphil Kim, Dmitry B. Goldgof, Lawrence O. Hall, Robert A. Gatenby, and Robert J. Gillies. Reproducibility and prognosis of quantitative features extracted from CT images. *Translational Oncology*, 7(1):72 – 87, 2014. The Quantitative Imaging Network.

[3] Yoganand Balagurunathan, Virendra Kumar, Yuhua Gu, Jongphil Kim, Hua Wang, Ying Liu, DmitryB. Goldgof, LawrenceO. Hall, Rene Korn, Binsheng Zhao, LawrenceH. Schwartz, Satrajit Basu, Steven Eschrich, RobertA. Gatenby, and RobertJ. Gillies. Test–retest reproducibility analysis of lung CT image features. *Journal of Digital Imaging*, 27(6):805–823, 2014.

[4] Leo Breiman. Better subset regression using the nonnegative garrote. *Technometrics*, 37(4):373–384, 1995.

[5] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

[6] S. Carvalho, R. T. H. Leijenaar, E. G. C. Troost, W. van Elmpt, J. P. Muratet, F. Denis, D. De Ruysscher, H. J. W. L. Aerts, and P. Lambin. Early variation of fdg-pet radiomics features in nsclc is related to overall survival - the “delta radiomics” concept. *Radiotherapy and Oncology*, 118:S20–S21, 2016/10/30 2016.

[7] Chih-Chung Chang and Chih-Jen Lin. Libsvm: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, 2(3):27:1–27:27, May 2011.

[8] N.V. Chawla, K.W. Bowyer, L.O. Hall, and W.P. Kegelmeyer. SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16:321–357, 2002.

[9] Ashis Kumar Dhara, Sudipta Mukhopadhyay, Anirvan Dutta, Mandeep Garg, and Niranjan Khandelwal. A combination of shape and texture features for classification of pulmonary nodules in lung ct images. *Journal of Digital Imaging*, 29(4):466–475, 2016.

[10] Samantha K N Dilger, Johanna Uthoff, Alexandra Judisch, Emily Hammond, Sarah L Mott, Brian J Smith, John D Newell, Eric A Hoffman, and Jessica C Sieren. Improved pulmonary nodule classification utilizing quantitative lung parenchyma features. *Journal of Medical Imaging*, 2(4):041004, 10 2015.

[11] Miglioretti DL, Johnson E, Williams A, and et al. The use of computed tomography in pediatrics and the associated radiation exposure and estimated cancer risk. *JAMA Pediatrics*, 167(8):700–707, 2013.

[12] X Fave, L Zhang, J Yang, D Mackin, F Stingo, D Followill, P Balter, A Jones, D Gomez, and L Court. Tu-d-207b-02: Delta-radiomics: The prognostic value of therapy-induced changes in radiomics features for stage iii non-small cell lung cancer patients. *Medical Physics*, 43(6):3750–3750, 2016.

[13] Andriy Fedorov, Reinhard Beichel, Jayashree Kalpathy-Cramer, Julien Finet, Jean-CristopheC. Fillion-Robin, Sonia Pujol, Christian Bauer, Dominique Jennings, FionaM Fennessy, Milan Sonka, John Buatti, StephenR Aylward, JamesV Miller, Steve Pieper, and Ron Kikinis. 3d slicer as an image computing platform for the quantitative imaging network. *Magnetic Resonance Imaging*, 30(9):1323–41, 11 2012.

[14] Benjamin Geiger, Samuel Hawkins, Lawrence O. Hall, Dmitry B. Goldgof, Yoganand Balagurunathan, Robert A. Gatenby, and Robert J. Gillies. Change descriptors for determining nodule malignancy in national lung screening trial ct screening images, 2016.

[15] Yuhua Gu, Virendra Kumar, Lawrence O. Hall, Dmitry B. Goldgof, Ching-Yen Li, René Korn, Claus Bendtsen, Emmanuel Rios Velazquez, Andre Dekker, Hugo Aerts, Philippe Lambin, Xiuli Li, Jie Tian, Robert A. Gatenby, and Robert J. Gillies. Automated delineation of lung tumors from CT images using a single click ensemble segmentation approach. *Pattern Recognition*, 46(3):692 – 702, 2013.

[16] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. The WEKA data mining software: an update. *SIGKDD Explor. Newsl.*, 11(1):10–18, November 2009.

[17] Corneline J. Hoekstra, Sigrid G. Stroobants, Egbert F. Smit, Johan Vansteenkiste, Harm van Tinteren, Pieter E. Postmus, Richard P. Golding, Bonne Biesma, Frans J.H.M. Schramel, Nico van Zandwijk, Adriaan A. Lammertsma, and Otto S. Hoekstra. Prognostic relevance of response evaluation using [18f]-2-fluoro-2-deoxy-d-glucose positron emission tomography in patients with locally advanced nonäìsmall-cell lung cancer. *Journal of Clinical Oncology*, 23(33):8362–8370, 2005. PMID: 16293866.

[18] Artit C. Jirapatnakul, Anthony P. Reeves, Tatiyana V. Apanasovich, Matthew D. Cham, David F. Yankelevitz, and Claudia I. Henschke. Characterization of solid pulmonary nodules using three-dimensional features. In *Proc. SPIE 6514, Medical Imaging 2007: Computer-Aided Diagnosis, 65143E*, pages 65143E–65143E–8, 2007.

[19] George H. John and Pat Langley. Estimating continuous distributions in bayesian classifiers. In *Eleventh Conference on Uncertainty in Artificial Intelligence*, pages 338–345, San Mateo, 1995. Morgan Kaufmann.

[20] Kenji Kira and Larry A Rendell. The feature selection problem: Traditional methods and a new algorithm. In *AAAI*, pages 129–134, 1992.

[21] H. Krewer, B. Geiger, L. O. Hall, D. B. Goldgof, Y. Gu, M. Tockman, and R. J. Gillies. Effect of texture features in computer aided diagnosis of pulmonary nodules in low-dose computed tomography. In *2013 IEEE International Conference on Systems, Man, and Cybernetics*, pages 3887–3891, Oct 2013.

[22] Jinsa Kuruvilla and K Gunavathi. Lung cancer classification using neural networks for ct images. *Computer methods and programs in biomedicine*, 113(1):202–209, 2014.

[23] Philippe Lambin, Emmanuel Rios-Velazquez, Ralph Leijenaar, Sara Carvalho, Ruud GPM van Stiphout, Patrick Granton, Catharina ML Zegers, Robert Gillies, Ronald Boellard, André Dekker, and Hugo JWL Aerts. Radiomics: Extracting more information from medical images using advanced feature analysis. *European journal of cancer (Oxford, England : 1990)*, 48(4):441–446, 03 2012.

[24] National Lung Screening Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. *New England Journal of Medicine*, 365(5):395–409, 2011. PMID: 21714641.

[25] Mark S Pearce, Jane A Salotti, Mark P Little, Kieran McHugh, Choonsik Lee, Kwang Pyo Kim, Nicola L Howe, Cecile M Ronckers, Preetha Rajaraman, Alan W Craft, Louise Parker, and Amy Berrington de González. Radiation exposure from ct scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study. *The Lancet*, 380(9840):499–505, August 2012.

[26] Ross Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers, San Mateo, CA, 1993.

[27] Wei-Chih Shen, Yang-Hao Yu, and Cheng-Hung Chuang. Computer aided diagnosis for pulmonary nodule on low-dose computed tomography (LDCT) using density features. In *Computer Graphics, Imaging and Visualization (CGIV), 2011 Eighth International Conference on*, pages 166–169, 2011.

[28] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.

[29] Wouter van Elmpt, Michel Öllers, Anne-Marie C. Dingemans, Philippe Lambin, and Dirk De Ruysscher. Response assessment using 18f-fdg pet early in the course of radiotherapy correlates with survival in advanced-stage non–small cell lung cancer. *Journal of Nuclear Medicine*, 53(10):1514–1520, 2012.

APPENDIX A: DEFINIENS FEATURE LIST

Table A.1 All features and their subset membership.

Name	**NC1Stable?**	**RiderC95?**	**Volume?**
Longest Diameter [mm]	Yes	Yes	
Short Axis * Longest Diameter [mm_]	Yes	Yes	
Short Axis [mm]	Yes	Yes	
Mean [HU]	Yes	Yes	
StdDev [HU]		Yes	
Volume [cm_]	Yes	Yes	Yes
5a_3D_MacSpic_NumberOf		Yes	
8a_3D_Is_Attached_To_Pleural_Wall	Yes	Yes	
8b_3D_Relative_Border_To_Lung	Yes	Yes	
8c_3D_Relative_Border_To_PleuralWall	Yes	Yes	
8d_3D_Ratio_Free_To_Attached			
9a_3D_FractionalAnisotropy			
9b_3D_Circularity		Yes	
9c_3D_Compactness	Yes		
9d_3D_AV_Dist_COG_To_Border_[mm]	Yes		
9e_3D_SD_Dist_COG_To_Border_[mm]	Yes	Yes	
9f_3D_MIN_Dist_COG_To_Border_[mm]	Yes		
9g_3D_MAX_Dist_COG_To_Border_[mm]	Yes	Yes	
10a_3D_Relative_Volume_AirSpaces	Yes		
10b_3D_Number_AirSpaces	Yes		

Table A.1 (Continued)

Name	**NC1Stable?**	**RiderC95?**	**Volume?**
10c_3D_Av_Volume_AirSpaces_[mm_]	Yes		
10d_3D_SD_Volume_AirSpaces_[mm_]			
Asymmetry		Yes	
Compactness	Yes		
Density			
Elliptic Fit			
Main direction			
Radius of largest enclosed ellipse			
Radius of smallest enclosing ellipse			
Shape index	Yes		
Roundness		Yes	
Rectangular Fit			
Area (Pxl)	Yes		
Volume (Pxl)	Yes	Yes	
Number of pixels	Yes	Yes	
Width (Pxl)	Yes		
Thickness (Pxl)	Yes		
Length (Pxl)	Yes		
Length/Thickness			
Length/Width		Yes	
Border length (Pxl)	Yes		
avgCoocurrence-Homo			
avgCoocurrence-MP			
avgCoocurrence-contrast			

Table A.1 (Continued)

Name	**NC1Stable?**	**RiderC95?**	**Volume?**
avgCoocurrence-energy			
avgCoocurrence-entropy			
avgCoocurrence-mean			
avgGLN	Yes		
avgHGRE	Yes		
avgLGRE			
avgLRE			
avgLRHGE	Yes		
avgLRLGE			
avgRLN	Yes		
avgRP	Yes		
avgSRE			
avgSRHGE	Yes		
avgSRLGE			
3D Laws features E5 E5 E5 Layer 1			
3D Laws features E5 E5 L5 Layer 1		Yes	
3D Laws features E5 E5 R5 Layer 1		Yes	
3D Laws features E5 E5 S5 Layer 1			
3D Laws features E5 E5 W5 Layer 1			
3D Laws features E5 L5 E5 Layer 1			
3D Laws features E5 L5 L5 Layer 1			
3D Laws features E5 L5 R5 Layer 1			
3D Laws features E5 L5 S5 Layer 1			
3D Laws features E5 L5 W5 Layer 1			

Table A.1 (Continued)

Name	**NC1Stable?**	**RiderC95?**	**Volume?**
3D Laws features E5 R5 E5 Layer 1			
3D Laws features E5 R5 L5 Layer 1			
3D Laws features E5 R5 R5 Layer 1			
3D Laws features E5 R5 S5 Layer 1			
3D Laws features E5 R5 W5 Layer 1			
3D Laws features E5 S5 E5 Layer 1			
3D Laws features E5 S5 L5 Layer 1			
3D Laws features E5 S5 R5 Layer 1			
3D Laws features E5 S5 S5 Layer 1			
3D Laws features E5 S5 W5 Layer 1			
3D Laws features E5 W5 E5 Layer 1			
3D Laws features E5 W5 L5 Layer 1		Yes	
3D Laws features E5 W5 R5 Layer 1			
3D Laws features E5 W5 S5 Layer 1			
3D Laws features E5 W5 W5 Layer 1			
3D Laws features L5 E5 E5 Layer 1			
3D Laws features L5 E5 L5 Layer 1			
3D Laws features L5 E5 R5 Layer 1			
3D Laws features L5 E5 S5 Layer 1			
3D Laws features L5 E5 W5 Layer 1			
3D Laws features L5 L5 E5 Layer 1			
3D Laws features L5 L5 L5 Layer 1	Yes		
3D Laws features L5 L5 R5 Layer 1			
3D Laws features L5 L5 S5 Layer 1			

Table A.1 (Continued)

Name	**NC1Stable?**	**RiderC95?**	**Volume?**
3D Laws features L5 L5 W5 Layer 1			
3D Laws features L5 R5 E5 Layer 1			
3D Laws features L5 R5 L5 Layer 1			
3D Laws features L5 R5 R5 Layer 1			
3D Laws features L5 R5 S5 Layer 1			
3D Laws features L5 R5 W5 Layer 1			
3D Laws features L5 S5 E5 Layer 1			
3D Laws features L5 S5 L5 Layer 1			
3D Laws features L5 S5 R5 Layer 1			
3D Laws features L5 S5 S5 Layer 1			
3D Laws features L5 S5 W5 Layer 1			
3D Laws features L5 W5 E5 Layer 1			
3D Laws features L5 W5 L5 Layer 1		Yes	
3D Laws features L5 W5 R5 Layer 1			
3D Laws features L5 W5 S5 Layer 1			
3D Laws features L5 W5 W5 Layer 1			
3D Laws features R5 E5 E5 Layer 1			
3D Laws features R5 E5 L5 Layer 1			
3D Laws features R5 E5 R5 Layer 1			
3D Laws features R5 E5 S5 Layer 1			
3D Laws features R5 E5 W5 Layer 1			
3D Laws features R5 L5 E5 Layer 1			
3D Laws features R5 L5 L5 Layer 1			
3D Laws features R5 L5 R5 Layer 1			

Table A.1 (Continued)

Name	**NC1Stable?**	**RiderC95?**	**Volume?**
3D Laws features R5 L5 S5 Layer 1			
3D Laws features R5 L5 W5 Layer 1			
3D Laws features R5 R5 E5 Layer 1			
3D Laws features R5 R5 L5 Layer 1			
3D Laws features R5 R5 R5 Layer 1			
3D Laws features R5 R5 S5 Layer 1			
3D Laws features R5 R5 W5 Layer 1			
3D Laws features R5 S5 E5 Layer 1			
3D Laws features R5 S5 L5 Layer 1			
3D Laws features R5 S5 R5 Layer 1			
3D Laws features R5 S5 S5 Layer 1			
3D Laws features R5 S5 W5 Layer 1			
3D Laws features R5 W5 E5 Layer 1			
3D Laws features R5 W5 L5 Layer 1			
3D Laws features R5 W5 R5 Layer 1			
3D Laws features R5 W5 S5 Layer 1			
3D Laws features R5 W5 W5 Layer 1			
3D Laws features S5 E5 E5 Layer 1			
3D Laws features S5 E5 L5 Layer 1			
3D Laws features S5 E5 R5 Layer 1			
3D Laws features S5 E5 S5 Layer 1			
3D Laws features S5 E5 W5 Layer 1			
3D Laws features S5 L5 E5 Layer 1			
3D Laws features S5 L5 L5 Layer 1			

Table A.1 (Continued)

Name	**NC1Stable?**	**RiderC95?**	**Volume?**
3D Laws features S5 L5 R5 Layer 1			
3D Laws features S5 L5 S5 Layer 1			
3D Laws features S5 L5 W5 Layer 1			
3D Laws features S5 R5 E5 Layer 1			
3D Laws features S5 R5 L5 Layer 1			
3D Laws features S5 R5 R5 Layer 1			
3D Laws features S5 R5 S5 Layer 1			
3D Laws features S5 R5 W5 Layer 1			
3D Laws features S5 S5 E5 Layer 1			
3D Laws features S5 S5 L5 Layer 1			
3D Laws features S5 S5 R5 Layer 1			
3D Laws features S5 S5 S5 Layer 1			
3D Laws features S5 S5 W5 Layer 1			
3D Laws features S5 W5 E5 Layer 1			
3D Laws features S5 W5 L5 Layer 1			
3D Laws features S5 W5 R5 Layer 1			
3D Laws features S5 W5 S5 Layer 1			
3D Laws features S5 W5 W5 Layer 1			
3D Laws features W5 E5 E5 Layer 1			
3D Laws features W5 E5 L5 Layer 1			
3D Laws features W5 E5 R5 Layer 1			
3D Laws features W5 E5 S5 Layer 1			
3D Laws features W5 E5 W5 Layer 1			
3D Laws features W5 L5 E5 Layer 1			

Table A.1 (Continued)

Name	**NC1Stable?**	**RiderC95?**	**Volume?**
3D Laws features W5 L5 L5 Layer 1			
3D Laws features W5 L5 R5 Layer 1			
3D Laws features W5 L5 S5 Layer 1			
3D Laws features W5 L5 W5 Layer 1			
3D Laws features W5 R5 E5 Layer 1			
3D Laws features W5 R5 L5 Layer 1			
3D Laws features W5 R5 R5 Layer 1			
3D Laws features W5 R5 S5 Layer 1			
3D Laws features W5 S5 E5 Layer 1			
3D Laws features W5 S5 L5 Layer 1			
3D Laws features W5 R5 W5 Layer 1			
3D Laws features W5 S5 R5 Layer 1			
3D Laws features W5 S5 S5 Layer 1			
3D Laws features W5 S5 W5 Layer 1			
3D Laws features W5 W5 E5 Layer 1			
3D Laws features W5 W5 L5 Layer 1			
3D Laws features W5 W5 R5 Layer 1			
3D Laws features W5 W5 S5 Layer 1			
3D Laws features W5 W5 W5 Layer 1			
Histogram Mean Layer 1			
Histogram SD Layer 1			
Histogram ENERGY Layer 1			
Histogram ENTROPY Layer 1	Yes		
Histogram KUR Layer 1			

Table A.1 (Continued)

Name	**NC1Stable?**	**RiderC95?**	**Volume?**
Histogram SKEW Layer 1	Yes		
3D Wavelet decomposition. P2 L2 C9 Layer 1			
3D Wavelet decomposition. P1 L2 C9 Layer 1			
3D Wavelet decomposition. P2 L2 C10 Layer 1			
3D Wavelet decomposition. P2 L2 C11 Layer 1			
3D Wavelet decomposition. P2 L2 C12 Layer 1			
3D Wavelet decomposition. P2 L2 C13 Layer 1	Yes		
3D Wavelet decomposition. P2 L2 C14 Layer 1	Yes		
3D Wavelet decomposition. P2 L2 C15 Layer 1	Yes		
3D Wavelet decomposition. P2 L2 C1 Layer 1			
3D Wavelet decomposition. P2 L2 C2 Layer 1			
3D Wavelet decomposition. P2 L2 C3 Layer 1			
3D Wavelet decomposition. P2 L2 C4 Layer 1			
3D Wavelet decomposition. P2 L2 C5 Layer 1			
3D Wavelet decomposition. P2 L2 C6 Layer 1			
3D Wavelet decomposition. P2 L2 C7 Layer 1			
3D Wavelet decomposition. P2 L2 C8 Layer 1			
3D Wavelet decomposition. P1 L2 C11 Layer 1			
3D Wavelet decomposition. P1 L2 C10 Layer 1			
3D Wavelet decomposition. P1 L2 C12 Layer 1			
3D Wavelet decomposition. P1 L2 C13 Layer 1			
3D Wavelet decomposition. P1 L2 C14 Layer 1			
3D Wavelet decomposition. P1 L2 C15 Layer 1			
3D Wavelet decomposition. P1 L2 C1 Layer 1			

Table A.1 (Continued)

Name	**NC1Stable?**	**RiderC95?**	**Volume?**
3D Wavelet decomposition. P1 L2 C2 Layer 1			
3D Wavelet decomposition. P1 L2 C3 Layer 1			
3D Wavelet decomposition. P1 L2 C4 Layer 1			
3D Wavelet decomposition. P1 L2 C5 Layer 1			
3D Wavelet decomposition. P1 L2 C6 Layer 1			
3D Wavelet decomposition. P1 L2 C7 Layer 1			
3D Wavelet decomposition. P1 L2 C8 Layer 1			

APPENDIX B: COPYRIGHT PERMISSIONS

The permission below is for the use of material in chapters 1, 2, and 3, and appendix A.

Material appears from "Change descriptors for determining nodule malignancy in national lung screening trial CT screening images", by Benjamin Geiger, Samuel Hawkins, Lawrence O. Hall, Dmitry B. Goldgof, Yoganand Balagurunathan, Robert A. Gatenby, and Robert J. Gillies, published in SPIE Proceedings Vol. 9785: "Medical Imaging 2016: Computer-Aided Diagnosis".

Nicole Harris <nicoleh@spie.org> 11:37 AM (0 minutes ago)

to me

Dear Mr. Geiger,

Thank you for seeking permission from SPIE to reprint material from our publications. As author, SPIE shares the copyright with you, so you retain the right to reproduce your paper in part or in whole.

Publisher's permission is hereby granted under the following conditions:

(1) the material to be used has appeared in our publication without credit or acknowledgment to another source; and

(2) you credit the original SPIE publication. Include the authors' names, title of paper, volume title, SPIE volume number, and year of publication in your credit statement.

Sincerely,

Nicole Harris

Proceedings Coordinator, SPIE

The permission below is for the use of material in Appendix A.

Material appears from "Effect of Texture Features in Computer Aided Diagnosis of Pulmonary Nodules in Low-Dose Computed Tomography", by Henry Krewer, Benjamin Geiger, Lawrence O. Hall, Dmitry B. Goldgof, Yuhua Gu, Melvyn Tockman, and Robert J. Gillies, published in the proceedings of the 2013 IEEE International Conference on Systems, Man, and Cybernetics, © 2013 IEEE.

- **Does IEEE require individuals working on a thesis or dissertation to obtain formal permission for reuse?**

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you must follow the requirements listed below:

Textual Material

Using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.

In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.

If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.