

Predicting the Occurrence of Radiation Induced Pneumonitis by Texture Analysis of CT Images from Lung Cancer Patients

Dean Montgomery^{*1}, Sorcha Campbell², Kun Cheng¹, Yang Feng¹, John Murchison³, Ai Wain Yong³, Gillian Ritchie³, Duncan B McLaren², Sara C Erridge², Stephen McLaughlin⁴, and William H Nailon¹

¹ Department of Oncology Physics, Western General Hospital, Edinburgh and the University of Edinburgh

² Department of Radiation Oncology, Western General Hospital, Edinburgh

³ Department of Radiology, Royal Infirmary of Edinburgh, Edinburgh

⁴ School of Engineering and Physical Sciences, Heriot Watt University

Abstract In 13-37% of cases, lung cancer patients treated with radiotherapy suffer from radiation induced lung disease, such as radiation induced pneumonitis. Three dimensional (3D) texture analysis, combined with patient-specific clinical parameters, were used to compute unique features ($n=2138$). Principal component analysis (PCA) was used to remove highly correlated features and a series of support vector machines (SVM) were used for classification in a leave one out scheme. On radiotherapy planning CT data of 57 patients, (14 symptomatic, 43 asymptomatic), the classifier obtained an area under the receiver operating curve of 0.873 with sensitivity, specificity and accuracy of 92%, 72% and 87% respectively. The combination of texture and clinical features demonstrates a statistically significant performance increase over the use of the clinical features alone. With further development the approach has the potential to be used to predict the likelihood of patients developing radiation induced pneumonitis in a clinical environment.

1 Introduction

Over the last two decades lung cancer has accounted for the majority of cancer attributable deaths and in Scotland the incidence of lung cancer is amongst the highest in the world [7,12]. Patients with stage I and stage II disease are treated with radical radiotherapy if they are inoperable or decline surgery. For stage III disease radical radiotherapy often combined with chemotherapy is the treatment of choice. The exposure of lung tissue to ionising radiation during radiotherapy can lead to physiological changes in the lung tissues and subsequently, radiation induced lung injury. Radiation induced pneumonitis develops in 13-37% of

^{*} This work is generously supported by NHS Lothian, Edinburgh and Lothian Health foundation (charity number SC007342), the James Clerk Maxwell Foundation, the Jamie King Uro-Oncology Endowment Fund and Darwin Awards from the University of Edinburgh.

patients after radiotherapy and causes reduced respiratory function in affected patients [17]. Months to years after radiotherapy, pulmonary fibrosis may develop - this is permanent scarring of the lung tissues and can lead to the serious impairment of oxygen transfer. It is currently difficult to reliably predict whether a patient is at a higher risk of radiation induced lung injury because the causal mechanism of this condition is not well understood [2,19].

To date there has been research into predicting the risk by looking at patient characteristics, chemotherapy factors and various dosimetric parameters. There have also been investigations into the use of image analysis techniques for the classification of existing lung disease, including fibrosis [1].

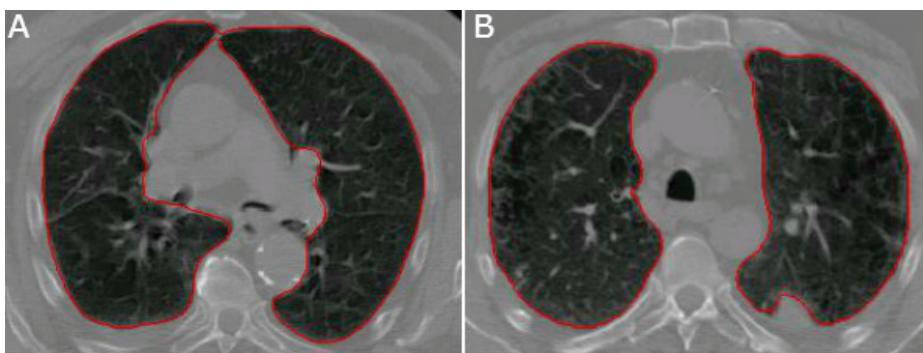


Figure 1. Single CT slice from a patient with lung cancer: A is a patient that did not sustain radiation induced injury after radiotherapy. B is a patient that developed radiation induced pneumonitis after radiotherapy. The red line is the lung volume as outlined during the radiotherapy planning stage. Visually the scans show no differentiating characteristics.

1.1 Pre-RT Dosimetric Parameters for Predicting the Risk of Pneumonitis

Significant correlation has been observed between properties of the dose-volume histogram (DVH) and the onset of radiation induced pneumonitis. A DVH is calculated from the patient radiotherapy plan and the percentage of lung tissue exposed to a particular dose is estimated - V_{Dose} . Several studies have shown correlation between V_{Dose} and radiation induced pneumonitis [4]. Mean lung dose (MLD), derived from the DVH, has been shown to correlate with pneumonitis [16]. Normal-tissue complication probability (NTCP) has also been used to determine the risk of pneumonitis. NTCP assumes that the probability of complications follows a sigmoidal dose-response relationship and uses mathematical models based on the tolerance dose for whole organ irradiation and the steepness of the dose-response curve to calculate the probability of lung damage [15]. Dosimetric parameters, clinical parameters and the location of the tumour

in the lung have been combined and used to classify patients with high and low risk better than when only a single parameter has been used [13].

1.2 Post-RT Image Analysis and Existing Lung Disease

There has been a significant amount of research carried out on identifying regions of the lung that have already changed due to lung disease. Texture analysis has been used to classify unhealthy regions of lung tissue [14,10]. Volumetric texture features have been used to classify interstitial lung diseases on CT images with some success [20,21,22].

Here image analysis techniques were applied to the radiotherapy planning CT scans, see Figure 1, to identify patients at a higher risk of developing radiation induced injuries. Three dimensional texture analysis was used to create a series of features for a number of patients. These features, together with patient-specific clinical features, were used to define a feature space that was used in a supervised classification scheme to predict the risk of pneumonitis from the planning CT scan. We are aware of no previous work that has applied such texture analysis to radiotherapy planning CT scans in an effort to predict the risk of induced pneumonitis.

2 Materials and Methods

Radiotherapy planning CT scans ($n=57$) of patients with lung cancer treated at the Edinburgh Cancer Centre in 2009 were selected for this study. All patients were scanned using a 3 mm CT slice thickness, (IGE HiSpeed Fx/i, GE Medical Systems, Milwaukee, WI, USA) resulting in a resolution of approximately 1 mm in the axial plane with a 2048 grey-level range. Of the 57 patients, 14 developed pneumonitis and 43 showed no symptoms of lung injury after treatment. The presence of radiation induced lung injury was categorised by expert clinicians involved in the radiotherapy and follow up treatment of the patients with lung cancer. The regions of the CT scans which were labelled as lung were extracted from the DICOM images. Thresholding was applied to remove the bronchial tree and a threshold implemented to replace blood vessels with the average pixel value of surrounding tissue.

2.1 Clinical Features

A number of clinical and dosimetric features, which are measured in the normal course of radical radiotherapy, were selected for use in the classification feature space. These features were age, smoker/non-smoker, if the patient suffers from asthma or similar conditions, T-stage, N-stage, if the patient was treated with chemotherapy, the radiotherapy dose and number of fractions, the V_{20} , V_{10} , V_5 , MLD and the size of the planning treatment volume (PTV).

2.2 Texture Analysis

First Order Statistics (FOS) are derived from the grey-level histogram. This is a simple approach which makes use of standard descriptors to characterise a region. The features, however do not take into account the spatial relationships and the correlation between pixels. The seven common features calculated are mean, variance, coarseness, skew, kurtosis, energy and entropy.

Grey Level Co-occurrence Matrices (GLCMs) are used to generate second order statistics. An entry (i, j) in a GLCM is the probability of finding a pixel with grey-level i at a distance d_s and angle α from a pixel with grey-level j . In 3D, 13 GLCMs are required to describe the texture in all directions for each d_s . In order to extract information from the GLCMs Haralick et al. [11] proposed a set of 14 local features. Many of these features can be highly correlated, therefore feature reduction or selection used prior to classification.

Grey Level Run Length Matrices (GLRLMs) contain information on the run of a grey-level in a particular direction [9]. The run-length is the number of pixels in a run. Fine textures will be dominated by short run-lengths, while coarse textures will be dominated by longer run-lengths. GLRLM methods are analogous to the GLCM method: 13 GLRLMs are used to describe texture in the same manner as the GLCMs. GLRLMs were first introduced by Galloway [9]: “The [GLRLM] matrix element (i, j) specifies the number of times that the picture contains a run of length j , in the given direction, consisting of points having grey level i (or lying in grey level range i).” Eleven features that describe the texture of the image are calculated from the GLRLMs.

Grey Level Size Zone Matrices (GLSZMs) are analogous to GLRLMs, but instead of measuring the length of a run in a particular direction, the size of a connected region of the same grey level is measured [18]. An entry (i, j) in a GLSZM is the probability of finding a connected region with grey-level i and of size j within an image. This method can easily be used in 2D or 3D and the same features are extracted as from GLRLMs, for example, long run emphasis becomes large zone emphasis. This leads to the calculation of 11 features for each GLSZM.

Gabor Filters were first defined by Gabor in one dimension [8] and later extended by Daugman to 2D [6], which have the ability to model the orientation and frequency sensitivity of the human visual system. Here 3D Gabor filters were used where the filter was a Gaussian kernel function modulated by a sinusoidal plane wave defined as:

$$\varphi_{f,\theta,\phi} = S \times \exp \left(- \left(\left(\frac{x'}{\sigma_x} \right)^2 + \left(\frac{y'}{\sigma_y} \right)^2 + \left(\frac{z'}{\sigma_z} \right)^2 \right) \right) \times \exp(j2\pi(xu + yv + zw)),$$

where $u = F \sin \phi \cos \theta$, $v = F \sin \phi \sin \theta$ $w = F \cos \phi$, $[x' y' z']^T = R \times [xyz]^T$, S is

a normalisation scale and $F = \sqrt{u^2 + v^2 + w^2}$ is the amplitude of the complex sinusoid wave with frequency (u, v, w) . $\theta \in (0 \leq \theta < \pi)$ and $\phi \in (0 \leq \phi < \pi)$ are the orientations of the wave vector in the 3D frequency domain and $\sigma_x, \sigma_y, \sigma_z$ define the width of the Gaussian envelope in the x, y and z axis respectively. R defines the rotation matrix for transforming the Gaussian envelope to coincide with the orientation of the sinusoid. Features are calculated by generating different Gabor filters using a range of ϕ, θ and F . For each filter the energy after convolution is calculated and this results in one feature being generated per filter. Using $F = [25 : 25 : 100]$, $\theta = [-\pi/3 : \pi/6 : \pi/2]$ and $\phi = [-\pi/3 : \pi/6 : \pi/2]$, 144 features are generated per image.

2.3 Feature Reduction

To remove highly correlated features and to reduce the size of the feature set before training PCA was used. PCA is a well demonstrated method for mapping features into a linear sub-space while retaining maximal variance. A normalised feature matrix, \mathbf{X} , is constructed: where each row corresponds to features calculated from a patient and each column corresponds to a particular feature. The reduced feature matrix, \mathbf{Z} , in the new sub-space is, $\mathbf{Z} = \mathbf{X}\mathbf{U}'_{red}$, where \mathbf{U}'_{red} is constructed via the singular value decomposition of the covariance matrix, Σ , which is given by $\Sigma = \frac{1}{m}\mathbf{X}'\mathbf{X}$, where m is the number of rows in matrix \mathbf{X} .

The singular value decomposition of Σ provides two matrices of interest; \mathbf{U} : the eigenvectors of \mathbf{X} , and \mathbf{S} : the eigenvalues (on the diagonal) of \mathbf{X} . \mathbf{U}'_{red} is then constructed from \mathbf{U} , by retaining the first k eigenvectors which retain $V\%$ of the variance in the data. That is the smallest value of k that satisfies $(\sum_{i=1}^k \mathbf{S}_{ii} / \sum_{i=1}^m \mathbf{S}_{ii}) \geq V$. The drawback of PCA is that it obfuscates the relative importance of the original features in \mathbf{X} as each new feature in \mathbf{Z} is a linear combination of the original features in \mathbf{X} . This, unfortunately, means that the identity of the features that are of greatest importance to the classification scheme are unknown.

2.4 Classification

A Support Vector Machine (SVM) [5] is a supervised learning model. The model is constructed by first mapping our feature space into a high order space using a kernel function. This allows for the generation of a non-linear decision boundary. The decision boundary is then constructed by finding a hyperplane which separates the data in the new feature space. The hyperplane which is furthest from the data points (the one that maximises the margin) is chosen. This decision boundary can then be used to classify new data points.

A Gaussian kernel SVM was used in this work. Both the γ parameter in the kernel and the soft margin parameter C , were optimised by a grid search on a cross validation set. Once these parameters had been tuned, the SVM was trained and a model generated for classification. In this work libsvm [3], an open source SVM library, was used.

3 Results

A total of 2125 texture features, in addition to 13 clinical features, were calculated on each of the 57 cases: 7 first order features, 11 GLSZM features, 143 GLRLM features (11 features \times 13 directions), 1820 GLCM features (14 features \times 13 directions \times 10 distances) and 144 Gabor features. In order to calculate the features, each image was scaled to 8, 16, 32 and 64 grey-levels. PCA was performed on the feature matrices prior to SVM, leave one out, classification. The training set used during leave one out classification used either all the cases or consisted of a balanced training set with equal numbers of patients in each class to attempt to overcome potential training bias.

3.1 Classification Performance

The maximum classification performance was: Area Under Reciever Operator Curves (ROC) = 0.873, Sensitivity = 92%, Specificity = 72%, and Accuracy = 87%. This was achieved using an SVM trained with a balanced training set of all the texture features combined with the clinical features and PCA reduction applied to the data with 95% of the variance retained.

The results of the best performing classifiers are presented in Table 1.

Table 1. Best performing classifiers with performance metrics. Each classifier was implemented in a leave one out scheme using CT scans from 57 patients, of which 14 went on to develop pneumonitis and the rest remained healthy. The table shows whether PCA was used, and if so, how much variance was retained in the feature set.

Classifier	Grey Levels	Features Used For Classification	PCA	Variance retained/%	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy	Area Under ROC
SVM-balanced	8	All texture with clinical	Yes	95	13	31	12	1	0.923	0.721	0.772	0.873
SVM-balanced	8	All texture with clinical	No	n/a	13	33	10	1	0.929	0.767	0.807	0.867
SVM-balanced	16	All texture with clinical	Yes	99	11	36	7	3	0.786	0.837	0.825	0.804
SVM-balanced	8	GLCM Features	Yes	99	12	33	10	2	0.857	0.767	0.790	0.823
SVM-balanced	64	Gabor Filter Features	Yes	99	12	28	15	2	0.857	0.651	0.702	0.804
SVM	8	All texture features	Yes	99	12	35	8	2	0.857	0.814	0.825	0.794
SVM	16	All texture features	No	n/a	11	31	12	3	0.786	0.721	0.737	0.779
SVM	16	All texture features	Yes	99	11	36	7	3	0.786	0.837	0.825	0.770
SVM	64	All texture features	Yes	99	11	36	7	3	0.786	0.837	0.825	0.736
SVM	8	All texture features	No	n/a	11	33	10	3	0.786	0.767	0.772	0.732

Table 2 displays the average performance of the SVMs trained with balanced data sets. The values have been averaged over the number of grey levels used and the different values of PCA used.

Table 2 shows that there is a increase in performance when the texture features are combined with the clinical features compared to using either set independently. A t-test was used to determine whether the increase in performance, when using both feature sets together, was statistically significant. These results can be seen in Table 3.

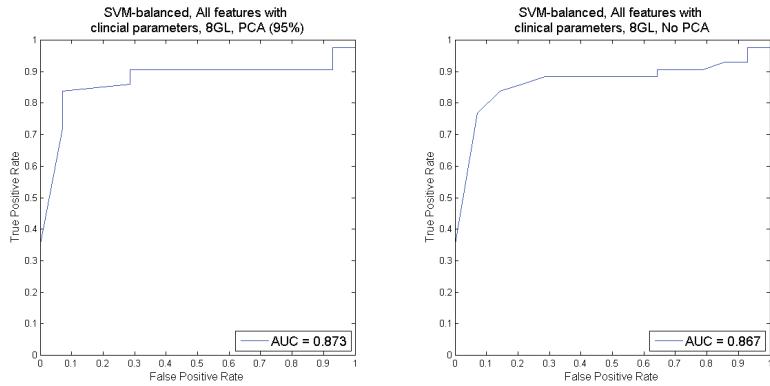


Figure 2. Receiver Operator Curves for the two best classifiers from Table 1.

Table 2. Average performance of the SVMs trained with a balanced training set. The table displays an improvement in average performance when the texture features are combined with the clinical features.

	Specificity	Sensitivity	Accuracy	Area Under ROC
All texture features with clinical	0.6899	0.7321	0.7003	0.6990
All texture features	0.6705	0.6964	0.6769	0.6521
Clinical features	0.6570	0.5655	0.6345	0.5815

4 Discussion

The results presented in Section 3 are very promising and demonstrate that an SVM trained with the features described has the potential to predict the risk of a patient developing radiation induced pneumonitis to a reasonable degree. This suggests that there is a significant difference in the lung tissues between patients who go on to develop the condition versus those who do not and that this underlying difference can be characterised by the use of texture calculated from radiotherapy planning CT images.

There is more work to be done in this area to build on the preliminary results presented in this paper. An obvious next step would be to verify the work with a much larger cohort of patients. More data would also remove the need to perform leave one out classification, which was used to simultaneously increase the amount of training and test data available. However, leave one out classification would not be suited to a clinical implementation. Other extensions of this work would be to incorporate an NTCP method, to develop a grading system that predicts the severity of radiation pneumonitis or to add additional features from the CT scans to improve classification performance.

It would also be useful to consider the severity of the induced lung injury. There are varying degrees of radiation induced pneumonitis and it would be an improvement to the method proposed in this paper to incorporate this in the form of a multi-class classifier, rather than a binary classifier. At the current

Table 3. p-values from a t-test comparing the performance of SVMs that were trained with either the texture features combined with the clinical features or trained with one of the individual feature sets. The results presented are for the SVMs trained with a balanced training set. The p-values demonstrate a significant increase in performance.

"Texture with Clinical" vs "Clinical"		Specificity	Sensitivity	Accuracy	Area Under ROC
Significant Increase	Yes	Yes	No	Yes	
p-value	0.2067	0.0101	0.0084	0.0136	
"Texture with Clinical" vs "Texture"		Specificity	Sensitivity	Accuracy	Area Under ROC
Significant Increase	Yes	Yes	Yes	Yes	
p-value	0.2839	0.2304	0.2127	0.1144	

time, the patient data does not include this grading information, though we are hopeful that this will be available in the near future.

The clinical application of this work would be to assist clinicians assess which patients are at a higher risk of developing clinically significant radiation induced pneumonitis and to adjust therapeutic management accordingly. The work presented in this paper displays a promising avenue for further research into the prediction of the risk of developing radiation induced pneumonitis with the overall aim of improving the clinical outcome for patients with lung cancer.

References

1. Bağcı, U., Bray, M., Caban, J., Yao, J., Mollura, D.J.: Computer-assisted detection of infectious lung diseases: A review. *Computerized Medical Imaging and Graphics* 36(1), 72–84 (2012)
2. Brunelli, A., Charloux, A., Bolliger, C., Rocco, G., Sculier, J.P., Varela, G., Licker, M., Ferguson, M., Faivre-Finn, C., Huber, R.M., et al.: Ers/ests clinical guidelines on fitness for radical therapy in lung cancer patients (surgery and chemo-radiotherapy). *European Respiratory Journal* 34(1), 17–41 (2009)
3. Chang, C.C., Lin, C.J.: Libsvm: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* 2, 27:1–27:27 (2011), software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
4. Claude, L., Pérol, D., Ginechet, C., Falchero, L., Arpin, D., Vincent, M., Martel, I., Hominal, S., Cordier, J.F., Carrie, C.: A prospective study on radiation pneumonitis following conformal radiation therapy in non-small-cell lung cancer: clinical and dosimetric factors analysis. *Radiotherapy and oncology* 71(2), 175–181 (2004)
5. Cortes, C., Vapnik, V.: Support-vector networks. *Machine learning* 20(3), 273–297 (1995)
6. Daugman, J.G., et al.: Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Optical Society of America, Journal, A: Optics and Image Science* 2(7), 1160–1169 (1985)
7. Ferlay, J., Shin, H.R., Bray, F., Forman, D., Mathers, C., Parkin, D.M.: Estimates of worldwide burden of cancer in 2008: Globocan 2008. *International journal of cancer* 127(12), 2893–2917 (2010)

8. Gabor, D.: Theory of communication. part 1: The analysis of information. Electrical Engineers-Part III: Radio and Communication Engineering, Journal of the Institution of 93(26), 429–441 (1946)
9. Galloway, M.M.: Texture analysis using gray level run lengths. Computer graphics and image processing 4(2), 172–179 (1975)
10. Gangeh, M.J., Sørensen, L., Shaker, S.B., Kamel, M.S., De Bruijne, M., Loog, M.: A textron-based approach for the classification of lung parenchyma in ct images. In: Medical Image Computing and Computer-Assisted Intervention—MICCAI 2010, pp. 595–602. Springer (2010)
11. Haralick, R.M., Shanmugam, K., Dinstein, I.H.: Textural features for image classification. Systems, Man and Cybernetics, IEEE Transactions on (6), 610–621 (1973)
12. Higginson, J., Muir, C.S., Munoz, N.: Human cancer: epidemiology and environmental causes. Cambridge University Press (1992)
13. Hope, A.J., Lindsay, P.E., El Naqa, I., Alaly, J.R., Vicic, M., Bradley, J.D., Deasy, J.O.: Modeling radiation pneumonitis risk with clinical, dosimetric, and spatial parameters. International Journal of Radiation Oncology* Biology* Physics 65(1), 112–124 (2006)
14. Lakshmi, D., Santhosham, R., Ranganathan, H.: Comparison of texture analysis in the differentiation of carcinoma from other lung abnormalities using low-dose ct images. In: Point-of-Care Healthcare Technologies (PHT), 2013 IEEE. pp. 271–274. IEEE (2013)
15. Mehta, V.: Radiation pneumonitis and pulmonary fibrosis in non-small-cell lung cancer: Pulmonary function, prediction, and prevention. International Journal of Radiation Oncology* Biology* Physics 63(1), 5–24 (2005)
16. Rancati, T., Ceresoli, G.L., Gagliardi, G., Schipani, S., Cattaneo, G.M.: Factors predicting radiation pneumonitis in lung cancer patients: a retrospective study. Radiotherapy and oncology 67(3), 275–283 (2003)
17. Rodrigues, G., Lock, M., D'Souza, D., Yu, E., Van Dyk, J.: Prediction of radiation pneumonitis by dose–volume histogram parameters in lung cancer—A systematic review. Radiotherapy and oncology 71(2), 127–138 (2004)
18. Thibault, G., Fertil, B., Navarro, C., Pereira, S., Cau, P., Levy, N., Sequeira, J., Mari, J.L.: Texture indexes and gray level size zone matrix: application to cell nuclei classification. Pattern Recognition and Information Processing (PRIP), Minsk, Belarus pp. 140–145 (2009)
19. Tsoutsou, P.G., Koukourakis, M.I.: Radiation pneumonitis and fibrosis: mechanisms underlying its pathogenesis and implications for future research. International Journal of Radiation Oncology* Biology* Physics 66(5), 1281–1293 (2006)
20. Uppaluri, R., Hoffman, E.A., Sonka, M., Hartley, P.G., Hunninghake, G.W., McLENNAN, G.: Computer recognition of regional lung disease patterns. American Journal of Respiratory and Critical Care Medicine 160(2), 648–654 (1999)
21. Xu, Y., Sonka, M., McLennan, G., Guo, J., Hoffman, E.A.: Mdct-based 3-d texture classification of emphysema and early smoking related lung pathologies. Medical Imaging, IEEE Transactions on 25(4), 464–475 (2006)
22. Zavaletta, V.A., Bartholmai, B.J., Robb, R.A.: High resolution multidetector ct-aided tissue analysis and quantification of lung fibrosis. Academic radiology 14(7), 772–787 (2007)