# Lung Tissue Classification in HRCT data Integrating the Clinical Context

Adrien Depeursinge, Jimison Iavindrasana, Gilles Cohen
University & Hospitals of Geneva
Service of Medical Informatics
24, rue Micheli–du–Crest
CH–1211 Geneva 14, Switzerland
adrien.depeursinge@sim.hcuge.ch

Alexandra Platon, Pierre–Alexandre Poletti
University & Hospitals of Geneva
Service of Emergency Radiology
24, rue Micheli–du–Crest
CH–1211 Geneva 14, Switzerland

Henning Müller
University of Applied Sciences
Sierre, Switzerland
University & Hospitals of Geneva
24, rue Micheli–du–Crest
CH–1211 Geneva 14, Switzerland

## Abstract

*In this paper, we investigate the influence of the clinical context of high–resolution computed tomography (HRCT) images of the chest on tissue classification. Evaluation of the classification performance is based on a high–quality multimedia data extracted from clinical routine. The clinical attributes with highest information gain ratio show to be relevant and consistent for the classification of lung tissue patterns. A combination of visual and clinical attributes allowed a mean of 93% of correct predictions of testing instances among the five classes of lung tissue with optimized support vector machines (SVM), which represents a significant benefit of 8% compared to a pure visually–based classification.*

## 1 Introduction

The interpretation of high–resolution computed tomography (HRCT) images of the chest from patients affected with interstitial lung diseases (ILDs) is challenging and time–consuming even for experienced radiologists. The term interstitial lung disease accounts for around 150 illnesses of which many forms are rare. Images play an important role for building the diagnosis and patients may not require surgical lung biopsy when the clinical and radiographic (HRCT) impression is consistent with a safe diagnosis [10]. The most common imaging procedure used is the chest x–ray because of its low cost and weak radiation exposure. However, chest x–rays are negative in a large portion of diseases and often unspecific where HRCT of the chest contains essential visual data for the characterization of lung tissue patterns associated with ILDs [23]. The three–dimensional form of HRCT data requires significant reading time, effort, and experience for a correct interpretation. Owing to this intrinsic complexity of the interpretation of HRCTs, an image–based computerized diagnostic aid tool can bring quick and precious information to less experienced radiologists and non–chest experts [2, 8, 22]. Moreover, radiologist's ability to interpret HRCT images is likely to change based on the domain–specific experience, human factors and time of the day where computerized image analysis is 100% reproducible.

### 1.1 The clinical context of HRCT images

When analyzing an image, one interprets its content according to a given context. For example, an image showing a heavenly beach is not interpreted in the same manner if seen in a holiday leaflet or in a report on tsunamis. This is particularly true when analyzing medical images. Radiologists do never interpret HRCT images without taking into account the clinical context. For example, discovering some fibrotic findings in a lung belonging to an 80–years–old patient is not as surprising as finding some in a lung of a 25–years–old young man. Several clinical parameters – in particular the age of the patient (see Fig. 1) – have major influence on the visual aspect (density) of HRCT images of the chest [17, 21]. In Figure 1, one can see that

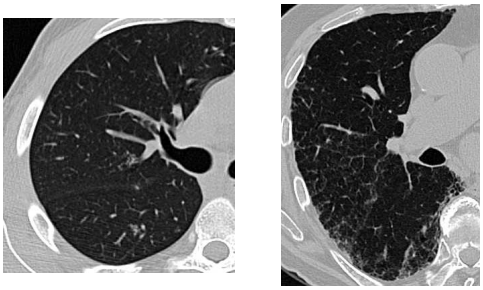

**Figure 1. Healthy tissues from a 25–years–old man on the left, and from an 88–years–old man on the right. Both images have identical window level settings.**

healthy tissue from the 88–years–old man has lower mean density with more pre–fibrotic lesions compared to the homogeneous healthy tissue of the 25–years–old man. To accurately analyze HRCT images an image–based computerized diagnostic aid system for ILDs must integrate the clinical context of the images.

### 1.2 Contextual image analysis

Although fundamental, the context is rarely used in computer vision applications. On the one hand, collecting contextual information beside images is usually time–consuming. One the other hand, a high–level of knowledge of the application domain is required to find relevant contextual parameters. The selection of parameters for contextual medical image analysis has to be carried out based on domain–specific literature along with knowledge bases of computer–based diagnostic decision support systems [11].

An early review of the integration of context for pattern recognition can be found in [24] where Bayesian models are used to implement compound decision theory.

Context has been used in content–based image retrieval (CBIR) where information from textual annotations of images is fused with image features [28] (i.e. grey–level histograms and texture features). In [3], a CBIR system combines visual statistics with textual statistics directly in the feature vector space representation. Inter–media medical image retrieval is carried out in [16] using textual features semantically parsed and described with the Unified Medical Language System (UMLS) along with color and texture features. The visual and textual information is combined in the similarity measure. Investigation of the effectiveness of combining text and image for retrieval including medical image retrieval is one of the main goals of the CBIR benchmarking campaign ImageCLEF[1] [12, 18].

A combination of radiologic findings on chest radiographs and clinical parameters to provide probability output of 11 possible ILDs using an artificial neural network is carried out in [1]. By using these probabilities, radiologists were able to significantly improve their diagnosing accuracies. However automatic detection of relevant patterns in the chest radiographs was not investigated.

Utilization of knowledge of disease location to improve detection of *fibrosis* patterns in HRCT data is carried out in [30]. The location of the patterns show to significantly improve detection performance but requires an accurate segmentation of the anatomy of the lung.

Many image–based diagnostic aid systems for ILDs achieved high classification accuracy of lung tissue patterns in HRCT data [22, 25, 26] and showed to be effective in clinical routine [2]. Yet, most of these systems are based on the visual data only (HRCT images). To our knowledge no system attempted to integrate clinical parameters for automatic detection of lung tissue patterns associated with ILDs in HRCT data.

Texture analysis of lung images using Wavelet frames was investigated in [7] and support vector machines (SVM) showed to be optimal for the categorization of lung tissue using quincunx Wavelet frames in [6]. In this paper, we study the influence of the integration of the clinical context of HRCT images on classification performance of 5 lung tissue patterns associated with ILDs.

## 2 Methods

The dataset used to investigate the influence of clinical parameters on classification accuracy of lung tissue is part of an internal multimedia database of ILD cases containing HRCT images created in the Talisman project. 99 relevant clinical parameters were chosen according to the 15 most frequent ILDs [13] based on the literature [14, 27], along with knowledge bases of computer–based diagnostic decision support systems [11]. Discussions and remarks from lung specialists, radiologists and the medical informatics research group (SIM[2]) at the University Hospitals of Geneva (HUG) allowed an iterative review of the selected parameters as well as standardized units and data format to be used. The parameters that were not available from the electronic health record (EHR) were removed. An HTML form and PHP scripts were used to collect the clinical parameters and to store them into a MySQL database. When multiples instances of clinical parameters (e.g. laboratory data) were available in the EHR, the instance as near as possible to HRCT examinations was retained. 96 patients with confirmed diagnosis have been retrospectively collected at the HUG since the year 2003. For each patient, clinical param-

[1] http://www.imageclef.org

[2] http://www.sim.hcuge.ch/

**Table 1. Distribution of the ROIs & patients per class of lung tissue pattern**

| | healthy | emphysema | ground glass | fibrosis | micronodules |
|---|---|---|---|---|---|
| # of ROIs | 63 | 58 | 148 | 312 | 155 |
| # of patients | 5 | 4 | 14 | 28 | 5 |

eters were filled as much as possible and a total of 1104 regions of interest (ROIs) of lung tissue patterns were drawn in full–resolution DICOM images by two experienced radiologists. The slice thickness of the images is limited to $1mm$. A graphical user interface implemented in Java was developed in order to meet the needs of the radiologists for the various annotation tasks.

736 ROIs from healthy and four pathologic lung tissue patterns belonging to 56 patients with filled clinical parameters are selected for this study (see Table 1). Patterns that are represented by less than 4 patients are left aside. The selected patterns are *healthy*, *emphysema*, *ground glass*, *fibrosis* and *micronodules*. Distributions of the classes are highly imbalanced as the largest class *fibrosis* contains 312 ROIs and the smallest class only 58 ROIs. There is a mean of 147.2 ROIs per class.

Implementation of the SVMs' $C$–support vector classification is taken from the open source Java library *Weka*[3] using a wrapper for *LIBSVM*[4]. The image feature extraction and the optimization of SVMs is implemented in Java. Quincunx Wavelet frames are implemented in Java [9].

# 3 Results

## 3.1 The clinical feature space

The clinical parameters entered in the MySQL database are not directly usable for data–mining. Pre–processing steps are required to build a workable feature space. Nominal variables are divided into binary features. Textual variables and binary variables that contained one single modality are left aside. After having gathered binary and continuous variables, the created clinical feature space contains (anew) 99 attributes (78 binary and 21 continuous). The filling rate of the retained parameters is 64.8%, which signifies that for each patient, approximately 35 attributes of 99 have missing values. Since leaving aside cases with missing values is not conceivable, mean values were substituted. For example, over the 56 selected patients, the parameter *host_HIV* has 4 *yes*, 51 *no* and 1 *unknown* value. *yes* values are coded with 1, *no* with 0 and thus the missing value is substituted by the mean: 0.073.

[3] http://www.cs.waikato.ac.nz/ml/weka/

[4] http://www.csie.ntu.edu.tw/ cjlin/libsvm/

## 3.2 Combining features

Visual features consist of texture features with 22 bins of grey–level histograms of Hounsfield Units (H.U.) within the ROIs, along with mean $\mu_i$ and variance $\sigma_i$ of quincunx Wavelet frame (QWF) coefficients extracted at 8 scales $i$. An additional feature *airpix* measuring the number of pixels of the ROI with value inferior to $-1000$ H.U. (which corresponds to the density of air) is used. A complete description and evaluation of the visual feature space can be found in [6, 7].

In order to create a multimodal feature space, clinical attributes described in Section 3.1 and visual features are normalized and concatenated into one single feature vector $\mathbf{v}$. The number of clinical attributes selected is studied. Indeed, due to missing values, binarization or irrelevance according to the studied diseases, some features might introduce noise by scattering homogeneous clusters of instances in the feature space. A feature selection is thus required to build an effective set of attributes. The information gain ratio $I_{G_{ratio}}$ is used to rank the clinical attributes [20]. Gain ratio is derived from the information gain measure $I_G$ originally used by Quinlan in decision trees in [19]. Compared to $I_G$ the gain ratio will not give advantage to attributes with a high range of possible values. As the clinical feature space is populated with binary as well as continuous attributes, it is highly preferable to use the $I_{G_{ratio}}$ for ranking. Table 2 lists the first 20 clinical attributes with highest $I_{G_{ratio}}$. When using $I_G$, the 10 attributes with highest $I_G$ are continuous, which confirms that $I_G$ is biased by the nature of the variable. The correlation matrix of the feature space containing the visual features along with the first 20 clinical attributes is shown in Figure 2.

## 3.3 Influence of clinical parameters on the classification accuracy

In order to study the effect of the integration of the clinical context of HRCT images on the classification accuracy of the lung tissue patterns, optimized SVMs with a Gaussian kernel are used to categorize ROIs from the multimodal feature space described in Section 3.2. SVMs with a Gaussian kernel have shown to be effective to categorize lung tissue patterns from visual features in [6] and are adapted to mine clinical parameters as shown in [5] where SVMs are used to detect nosocomial infection from a clinical feature space.

The methodology used for the evaluation of the influence of clinical parameters on classification accuracy is organized as follows: 50 patients (90%) are randomly drawn from the full dataset and used to train and optimize the

**Table 2. List of the first 20 clinical parameters with highest $I_{G_{ratio}}$. Abbreviations:** crp: $C$ reactive protein, wbc: white blood cell, vs: sedimentation speed**.**

| rank | $I_{G_{ratio}}$ | clinical parameter | type |
|---|---|---|---|
| 1 | 0.5486 | findings_physical_generals_lymph | binary |
| 2 | 0.5194 | host_hemopathy | binary |
| 3 | 0.5194 | past_medical_lymphom | binary |
| 4 | 0.5062 | host_chemotherapy | binary |
| 5 | 0.4772 | past_medical_neoplasm | binary |
| 6 | 0.4685 | findings_physical_generals_fever | binary |
| 7 | 0.4641 | findings_physical_visual | binary |
| 8 | 0.4419 | laboratory_angiotensin | continuous |
| 9 | 0.4377 | laboratory_hemoglobin | continuous |
| 10 | 0.4212 | age | continuous |
| 11 | 0.4184 | findings_physical_generals_arthralgy | binary |
| 12 | 0.406 | medication_gold | binary |
| 13 | 0.4025 | laboratory_hematocrit | continuous |
| 14 | 0.4006 | laboratory_crp | continuous |
| 15 | 0.3948 | medication_cyclophosphamides | binary |
| 16 | 0.3853 | laboratory_wbc | continuous |
| 17 | 0.3814 | medication_cyclines | binary |
| 18 | 0.3721 | medication_corticosteroid_duration | continuous |
| 19 | 0.3716 | laboratory_wbc_neutrophilic | continuous |
| 20 | 0.3644 | laboratory_vs | continuous |

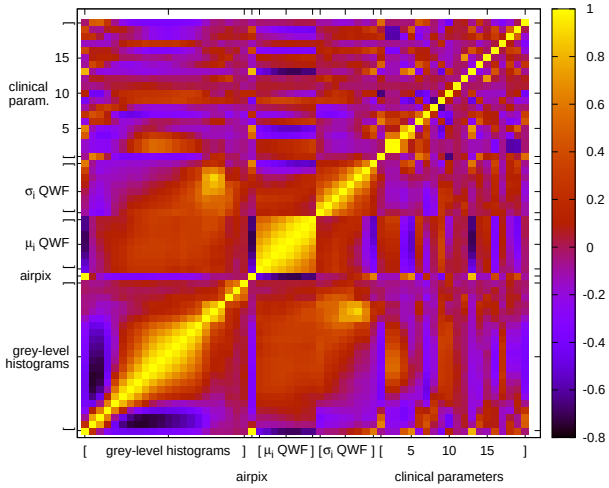


**Figure 2. Correlation matrix of the combined feature space. Indexes of the clinical parameters corresponds to their rank described in Table 2.**

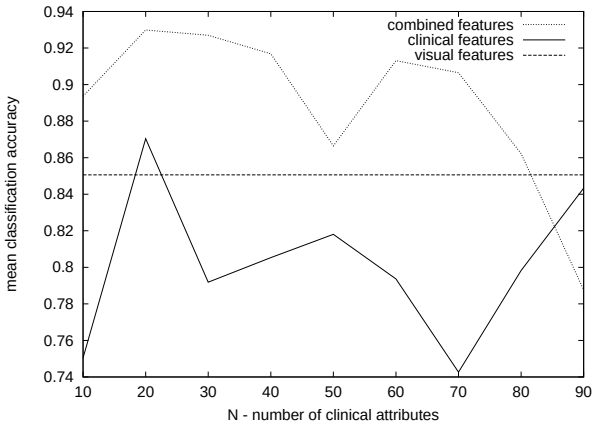


**Figure 3. Mean accuracies based on the number $N$ of clinical parameters ordered by $I_{G_{ratio}}$. Combined features with optimal $N = 20$ clinical parameters allowed a mean of 93% correct predictions of testing instances among the five lung tissue classes.**

SVM. The remaining patients are used for testing. As classes are highly imbalanced, the training set is built from stratification, which consists at selecting at least one patient per class for training. This avoids testing with unseen classes. The training set is used both for grid search for optimal parameters and adjustment of the maximum–margin hyperplane of the SVMs. Optimized parameters of the SVM are the cost of the errors $C$ and the width $\sigma$ of the Gaussian kernel. A grid search is carried out within the intervals $C \in [1; 100]$ and $\sigma \in [10^{-2}; 10^{2}]$. For every coordinate of the grid, a 10–fold cross–validation (CV) is carried out on the training set. Optimal parameters $(C_{opt}, \sigma_{opt})$ that allowed best mean CV accuracy $A^{cv}$ are used to train the final model on the entire training set. A preliminary coarse grid search was performed to locate regions of the space with high $A^{cv}$ values.

In order to determine the optimal number $N$ of clinical attributes to be added to the feature vector $\mathbf{v}$, the global experimentation described above is repeated 30 times for each $N \in [10; 90]$. Means of the classification accuracies obtained with the test set using visual features, clinical features and combined features according to $N$ are shown in Figure 3.

## 4 Interpretation

### 4.1 Relevance of clinical attributes

The relevance of the clinical attributes for classifying lung tissue patterns in HRCT data is subject to many ex-

ternal factors such as the availability of the parameters in the EHR, its binarization required to be added to $\mathbf{v}$ and relevance according to the studied diseases. Indeed a parameter such as the result of a lung biopsy is obviously highly informative for characterizing the lung tissue but is rarely carried out and available in the EHR (the parameter *biopsy_lung_interstitial_fibrosis* has 87% missing values and is nevertheless ranked to $26^{th}$ position according to $I_{G_{ratio}}$ value). The categorization and binarization has also major influence on the quality of clinical data. At last, the relevance of the parameter according to the studied diseases is of course primordial.

The presence of the parameter *findings_physical_general_lymph* (which stands for the enlargement of lymph node(s)) at the top of the list is not surprising as it usually highlights the presence of a host illness. Idem for the parameter *findings_physical_general_fever* located at the $6^{th}$ rank. As observed in Section 1.1, the age has an important influence on the visual aspect of lung tissue (see Figure 1) and it is not astonishing to find it at the $10^{th}$ rank. However, the presence of the parameter *findings_physical_visual* at the $7^{th}$ rank which characterizes visual anomalies (i.e. exophtalmy, cataract, ...) is most certainly the result of a coincidence.

The study of the correlation of the multimodal feature space is carried out in Figure 2. Correlations among the visual features are detailed in [6]. A first look at the correlation matrix shows that clinical features are very little correlated with the visual features. On the second look one can observe that clinical feature number 13, *laboratory_hematocrit*, is strongly anticorrelated with the means $\mu_i$ of the QWF and highly correlated to *airpix* the number of pixels of air within the ROIs. An explication for this is that large homogeneous regions of air, characterizing *emphysema* patterns, will cause hypoxia and may elicit an increased production of red blood cells by the kidney, and thus increase the level of hematocrit. This phenomenon is indeed commonly observed in cases affected with chronic obstructive pulmonary disease (COPD) [4], characterized by HRCT images showing *emphysema* patterns.

### 4.2 Influence of the clinical context on lung tissue classification

Influence of the clinical context of HRCT images on lung tissue classification accuracy is studied in Figure 3. As baseline performance, let us consider the mean accuracy of 85% achieved by using visual features only. It seems clear that the optimal number $N$ of clinical attributes both for classifications based on clinical features only and combined features is 20. It allowed a mean of 93% correct predictions of testing instances among the five classes of lung tissue, which represents a significant benefit of 8% compared to a pure visually–based classification. Fluctuations of the performances according to $N$ are the result of interactions among the various groups of features. Variations of the performance are consistent until $N = 30$. For $N = 50$, the clinically–based classifications obtain a local maximum whereas classifications using combined featured have a local minimum. This highlights the fact that the clinical parameters scatter homogeneous clusters of instances in the feature space. By adding 10 more parameters ($N = 60$), clusters are re–organized and the mean accuracy with combined features rises again. The observed fluctuations suggest that features issued from distinct sources must be mined separately, and the fusion of the results has to be carried out by combining the output of hierarchical classifiers as in [15], or by using a different kernel for each data source [29].

## 5 Conclusions

In this paper, the influence of the clinical context on lung tissue classification from HRCT data was investigated. A combination of visual and clinical attributes allowed a mean of 93% of correct predictions of testing instances among the five classes of lung tissue with optimized SVMs. This represents a significant benefit of 8% compared to a pure visually–based classification. Accuracy values are trustworthy for further usage in clinical routine as we never train and test with ROIs that belongs to the same patient. The fluctuations of performances according to the number of clinical attributes used suggests that features issued from distinct sources must be mined separately in order to preserve homogeneous clusters of instances in the feature space.

## 6 Acknowledgments

This work was supported by the Swiss National Science Foundation (FNS) with grant 200020–118638/1, the equalization fund of University and Hospitals of Geneva (grant 05–9–II) and the EU $6^{th}$ Framework Program in the context of the KnowARC project (IST 032691).

## References

[1] H. Abe, K. Ashizawa, F. Li, N. Matsuyama, A. Fukushima, J. Shiraishi, H. MacMahon, and K. Doi. Artificial neural networks (ANNs) for differential diagnosis of interstitial lung disease : results of a simulation test with actual clinical cases. *Academic Radiology*, 11:29–37, 2004.

[2] A. M. Aisen, L. S. Broderick, H. Winer-Muram, C. E. Brodley, A. C. Kak, C. Pavlopoulou, J. Dy, C.-R. Shyu, and A. Marchiori. Automated storage and retrieval of thin–section CT images to assist diagnosis: System description and preliminary assessment. *Radiology*, 228:265–270, 2003.

[3] M. L. Cascia, S. Sethi, and S. Sclaroff. Combining textual and visual cues for content-based image retrieval on the world wide web. In *CBAIVL '98: Proceedings of the IEEE Workshop on Content - Based Access of Image and Video Libraries*, page 24, Washington, DC, USA, 1998. IEEE Computer Society.

[4] A. Chambellan, E. Chailleux, and T. Similowski. Prognostic value of the hematocrit in patients with severe COPD receiving long–term oxygen therapy. *Chest*, 128:1201–1208, 2005.

[5] G. Cohen, M. Hilario, H. Sax, S. Hugonnet, and A. Geissbuhler. Learning from imbalanced data in surveillance of nosocomial infection. *Artificial Intelligence in Medicine*, 37(1):7–18, May 2006.

[6] A. Depeursinge, J. Iavindrasana, A. Hidki, G. Cohen, A. Geissbuhler, A. Platon, P.-A. Poletti, and H. Müller. A classification framework for lung tissue categorization. In *SPIE Medical Imaging*, San Diego, CA, USA, February 2008–to appear.

[7] A. Depeursinge, D. Sage, A. Hidki, A. Platon, P.-A. Poletti, M. Unser, and H. Müller. Lung tissue classification using Wavelet frames. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, Lyon, France, August 2007.

[8] K. Doi. Current status and future potential of computer-aided diagnosis in medical imaging. *British Journal of Radiology*, 78:3–19, 2005.

[9] M. Feilner, D. Van De Ville, and M. Unser. An orthogonal family of quincunx wavelets with continuously adjustable order. *IEEE Transactions on Image Processing*, 14(4):499–510, April 2005.

[10] K. R. Flaherty, T. E. King, J. Ganesh Raghu, J. P. Lynch III, T. V. Colby, W. D. Travis, B. H. Gross, E. A. Kazerooni, G. B. Toews, Q. Long, S. Murray, V. N. Lama, S. E. Gay, and F. J. Martinez. Idiopathic interstitial pneumonia: What is the effect of a multidisciplinary approach to diagnosis? *American Journal of Respiratory and Critical Care Medicine*, 170:904–910, July 2004.

[11] C. P. Friedman, A. S. Elstein, F. M. Wolf, G. C. Murphy, T. M. Franz, P. S. Heckerling, P. L. Fine, T. M. Miller, and V. Abraham. Enhancement of clinician's diagnostic reasoning by computer-based consultation. *Journal of the American Medical Association*, 282(19):1851–1856, November 1999.

[12] W. Hersh, J. Kalpathy-Cramer, and J. Jensen. Medical image retrieval and automated annotation: OHSU at ImageCLEF 2006. In C. Peters, P. Clough, F. C. Gey, J. Karlgren, B. Magnini, D. W. Oard, M. de Rijke, and M. Stempfhuber, editors, *CLEF*, volume 4730 of *Lecture Notes in Computer Science*, pages 660–669. Springer, 2006.

[13] A. Hidki, H. Müller, A. Depeursinge, P.-A. Poletti, and A. Geissbuhler. Putting the image into perspective: The need for domain knowledge when performing image–based diagnostic aid. In *Swiss conference on medical informatics (SSIM 2006)*, Basel, Switzerland, April 2006.

[14] T. E. King. Approach to the adult with interstitial lung disease. *UpToDate*, August, 2004.

[15] J. Kludas, E. Bruno, and S. Marchand-Maillet. Information fusion in multimedia information retrieval. In *Proceedings of 5th international Workshop on Adaptive Multimedia Retrieval (AMR)*, Paris, France, July 5-6 2007.

[16] C. Lacoste, J.-P. Chevallet, J.-H. Lim, D. T. H. Le, W. Xiong, D. Racoceanu, R. Teodorescu, and N. Vuillenemot. Inter-media concept-based medical image indexing and retrieval with UMLS at IPAL. In C. Peters, P. Clough, F. C. Gey, J. Karlgren, B. Magnini, D. W. Oard, M. de Rijke, and M. Stempfhuber, editors, *CLEF*, volume 4730 of *Lecture Notes in Computer Science*, pages 694–701. Springer, 2006.

[17] F. Mitsunobu, T. Mifune, K. Ashida, Y. Hosaki, H. Tsugeno, M. Okamoto, S. Harada, S. Takata, and Y. Tanizaki. Influence of age and disease severity on high resolution CT lung densitometry in asthma. *Thorax*, 56:851–856, 2001.

[18] H. Müller, N. Michoux, D. Bandon, and A. Geissbuhler. A review of content–based image retrieval systems in medicine – clinical benefits and future directions. *International Journal of Medical Informatics*, 73:1–23, 2004.

[19] R. J. Quinlan. Induction of decision trees. *Machine Learning*, 1(1):81–106, March 1986.

[20] R. J. Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers Inc., 1993.

[21] J. H. Ryu, E. J. Olson, D. E. Midthun, and S. J. Swensen. Diagnostic approach to the patient with diffuse lung disease. In *Mayo Clin Proc.*, volume 77, pages 1221–1227, 2002.

[22] C.-R. Shyu, C. E. Brodley, A. C. Kak, A. Kosaka, A. M. Aisen, and L. S. Broderick. ASSERT: A physician–in–the–loop content–based retrieval system for HRCT image databases. *Computer Vision and Image Understanding (special issue on content–based access for image and video libraries)*, 75(1/2):111–132, July/August 1999.

[23] P. Stark. High resolution computed tomography of the lungs. *UpToDate*, September, 2007.

[24] G. T. Toussaint. The use of context in pattern recognition. *Pattern Recognition*, 10:189–204, 1978.

[25] Y. Uchiyama, S. Katsuragawa, H. Abe, J. Shiraishi, F. Li, Q. Li, C.-T. Zhang, K. Suzuki, and K. Doi. Quantitative computerized analysis of diffuse lung disease in high–resolution computed tomography. *Medical Physics*, 30(9):2440–2454, September 2003.

[26] R. Uppaluri, E. A. Hoffman, M. Sonka, P. G. Hartley, G. W. Hunninghake, and G. McLennan. Computer recognition of regional lung disease patterns. *American Journal of Respiratory and Critical Care Medicine*, 160(2):648–654, August 1999.

[27] W. R. Webb, N. L. Müller, and D. P. Naidich, editors. *High-Resolution CT of the Lung*. Lippincott Williams & Wilkins, Philadelphia, PA, USA, 2001.

[28] T. Westerveld. Image retrieval: Content versus context. In *Recherche d'Informations Assistée par Ordinateur (RIAO'2000) Computer–Assisted Information Retrieval*, volume 1, pages 276–284, Paris, France, April 2000.

[29] A. Zien and C. S. Ong. Multiclass multiple kernel learning. In *ICML '07: Proceedings of the 24th international conference on Machine learning*, pages 1191–1198, New York, NY, USA, 2007. ACM.

[30] T. Zrimec and J. Wong. Improving computer aided disease detection using knowledge of disease appearance. *Studies in Health Technology and Informatics*, 129:1324,1328, 2007.