Lung tissue classification in severe advanced cystic fibrosis from CT scans

Francesco Ciompi[2,3], Alexios Palaioroutas[1], Martine Loeve[4,5], Oriol Pujol[2,3], Petia Radeva[2,3], Harm Tiddens[4,5] and Marleen de Bruijne[1,5,6], on behalf of the CT CFS study group

[1]Biomedical Imaging Group Rotterdam, Dept. of Medical Informatics, Erasmus MC, Rotterdam, The Netherlands
[2]Dept. of Applied Mathematics and Analysis, University of Barcelona, Spain
[3]Computer Vision Center, Campus UAB, Barcelona, Spain
[4]Dept. of Pediatric Pulmonology and Allergology, Erasmus MC - Sophia Children's Hospital, Rotterdam, The Netherlands
[5]Dept. of Radiology, Erasmus MC, Rotterdam, The Netherlands
[6]Dept. of Computer Science, University of Copenhagen, Copenhagen, Denmark
{fciompi; oriol; petia}@maia.ub.es
alexpaleoroutas@yahoo.com
{m.loeve; h.tiddens; marleen.debruijne}@erasmusmc.nl

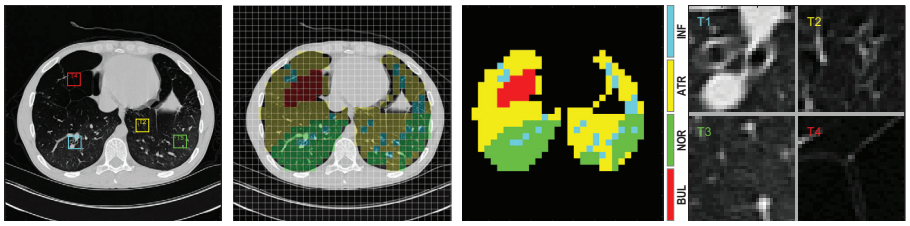
Abstract. A framework for lung tissue classification in Computed Tomography (CT) scans is presented. The method combines supervised and unsupervised learning techniques, with the aim of classifying four tissue types in lung: (*i*) inflammation, (*ii*) air-trapped / hypoperfused, (*iii*) normal / hyperperfused and (*iv*) bulla / cyst. The framework has been tested on a large heterogeneous dataset, collected over the last 20 years from 17 sites worldwide. The overall accuracy of the proposed methodology is 72.6% and the mean tissue sensitivity and specificity are 70.1% and 89.6% respectively: these results suggest that the framework can be used for the assessment of Severe Advanced Lung Disease (SALD) in patients affected by Cystic Fibrosis.

1 Introduction

Computed Tomography (CT) is a non-invasive imaging technique commonly used for the analysis of lung parenchyma. Voxels in CT scans are displayed according to the mean attenuation of the tissue, measured in Hounsfield Units (HU).

Cystic Fibrosis (CF) is a genetic disease that affects several parts of the human body, often leading to death. The leading cause of mortality in CF patients is *Severe Advanced Lung Disease* (SALD) [1]. The abnormalities that are observed by CT in the lung of patients with CF can be split into the following categories: (1) *bronchiectasis*, consisting in the destruction or widening of the airways; (2) *bronchial wall thickening*, caused by the excess of fluid or mucus in the small airway passages of the lung; (3) *mucus*; (4) *pulmonary consolidations*,

due to the presence of liquid (instead of gas) in lung alveolar space; (5) *air trapping*, which is the abnormal retention of gas in the lung after full expiration; (6) *bullae and cysts*, consisting in large pockets of either air or liquid/semi-solid inside the lung, respectively.



(a) (b) (c) (d)

Fig. 1. Example of lung CT slice labeling. Given the CT slice (a), a grid is superimposed (b) in order to guide the expert in the labeling process. Four labels, corresponding to four different colors, are assigned (c). In (d) the tissue samples (2×2 cells per sample) highlighted in (a) are depicted.

The quality of life for patients with SALD can improve thanks to a lung transplant (LTx), providing survival benefit as well. As a standard clinical procedure, patients are listed when their clinical condition is found severe enough to require LTx. For this reason, accurate and objective measures to evaluate the severity of abnormalities are required. *Loeve et al.* recently proposed a *SALD scoring* procedure [2], based on the assessment of the pulmonary tissue composition and amount, estimating relative volume of the different abnormalities. Furthermore, the assessment of disease severity is useful for monitoring purposes and for personalized treatment as well.

In this paper we present a methodology for automatic tissue classification in lung CT scans. The presented technique is oriented to the definition of a fully automatic and objective model for the quantitative description of the abnormalities found in CF and the computation of SALD scoring. Several techniques for automatic lung tissue classification in CT images have been presented so far. In [3] a set of circular ROIs from CT images are used to train a Bayesian classifier in a supervised manner, to classify four tissues by texture analysis and using statistical moments as features. A similar approach is presented in [4], where the supervised learning is devolved upon a Knn classifier to classify six tissues by texture analysis: ROIs segmented by the unsupervised method proposed in the paper are classified; in [5] textons are used as features to train a Support Vector Machine to discriminate three tissues: the classification is performed at voxel level; in [6] Local Binary Patterns are used as texture descriptors, and ROIs are classified by Knn. In [7] the dissimilarity between ROIs in lung CT images is used as a measure to quantify chronic obstructive pulmonary disease. In [8] the fusion of features based on both texture analysis and clinical parameters leads to a better accuracy classifying five tissues by SVM. Finally, in [9] the superi-

ority, in terms of accuracy, of 3D features with respect to 2D features has been demonstrated.

The technique proposed in this paper is close to the ones in [3–6], since it relies on a pattern recognition approach where tissue is classified by texture analysis. Differently from previous approaches, in order to cope with differences in inspiration level, inter-patients and inter-scanners variability, we formulate the classification process into three steps: (1) first, a multi-class classifier is trained in a *supervised* manner, by using the manual labeling provided by experts as ground truth; (2) additionally, an image-based *unsupervised* clustering is applied, in order to take into account per-image differences in tissues appearance; finally (3), a labels post-processing based on a-priori knowledge on morphological tissues properties is applied. To the best of our knowledge, the combination of supervised and unsupervised learning has never been used in lung tissue classification in CT scans.

2 Method

The automatic tissue classification is formulated as a pattern recognition problem, consisting in the extraction of features as descriptors of the lung tissue, followed by classification.

Based on a recently presented clinical study on SALD [10], four classes are considered, consisting in: (i) inflammation / infection (INF); (ii) air-trapped / hypoperfused tissue (ATR); (iii) normal / hyperperfused tissue (NOR); (iv) bulla / cyst (BUL). Labels are assigned to tissues after partitioning the CT slice into square regions (cells) of 1 $cm \times 1$ cm (see Fig. 1(b)).

Due to differences in inter-patients inspiration levels and in inter-scanners imaging parameters, the discrimination between ATR and NOR class is often challenging. The two tissues can often appear with similar texture and they may also share values in the HU range. Furthermore, due to the dark appearance and texture homogeneity of ATR cells, they can often be confused with BUL. Finally, in presence of extremely enlarged airways, the inner region of INF itself can be confused as BUL as well.

In order to cope with these objective problems, we formulate the tissue classification into three steps. The first one consists in a supervised learning process, where general tissue properties acquired by different scanners and from different patients are learned, and a first tissue labeling is obtained. Labels provided by a clinical study [2] are used to both train the classifier in a supervised manner and to validate the entire framework. As a second step, a *per-slice* unsupervised classification solely based on the tissue HU value is applied. Finally, regions previously classified as BUL are post-processed by applying a-priori knowledge on bullae morphology. The entire classification framework is detailed in next sessions.

2.1 Features extraction

Let us define as G_i a region in the CT slice corresponding to the i^{th} cell, $i = 1, \ldots, N_G$, where N_G is the total cells number per slice. For each cell, let us define as $(\mathbf{x}_i, y_i)$ the pair of feature vector $\mathbf{x}_i$ and its corresponding label y_i, where $y_i = 1, \ldots, 4$. In the entire CT slice, the data and label fields can be expressed as $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_{N_G}\}$ and $Y = \{y_1, \ldots, y_{N_G}\}$, respectively.

In order to characterize the properties of the tissue contained in each cell, as in [4] three gaussian kernels have been defined, with standard deviations $(\sigma_1, \sigma_2, \sigma_3)$. The gradient of the kernels has been then used to compute the Gradient Magnitude, the Laplacian and the Hessian matrix for every voxel in the slice. The eigenvalues of the Hessian matrix have been considered as tissue descriptors, along with the values of gradient magnitude and the response of the laplacian operator. Additionally, the value of the *Hounsfield Unit* (HU) and its median value in the region G_i have been also considered as tissue descriptors. The total number of descriptors for each voxel in the slice is $N_F = 17$. A feature vector $\mathbf{x}_i \in \Re^{257}$ is finally assigned to each cell G_i by computing the 16 bins histogram of all voxels in the cell for each descriptor, except for the median value; the histogram equalization procedure described in [6] have been applied.

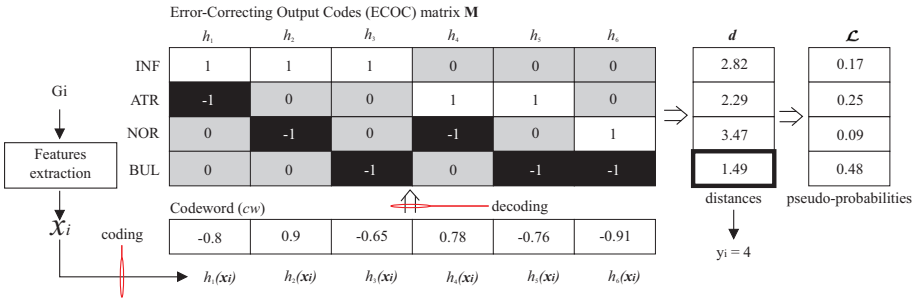


Fig. 2. Example of supervised classification of a cell G_i by the ECOC framework using OneVsOne coding and Euclidean decoding. Each dichotomy is indicated as h_j, while the decoding distances and the pseudo-probabilities are indicated as vectors $\mathbf{d}$ and $\boldsymbol{\mathcal{L}}$, respectively.

2.2 Supervised classification

The automatic classification of the four defined tissues represents a multi-class problem in pattern recognition. The approach adopted to cope with the multi-class problem is the *Error-Correcting Output Codes* (ECOC) [11]. ECOC is a technique that combines $\mathcal{N}$ binary classifiers (dichotomies) to solve a $\mathcal{K}$-classes classification problem. For each class a code $c_k = \{1, 0, -1\}^{1 \times \mathcal{N}}$ is designed $(k = 1, \ldots, \mathcal{K})$ according to a coding technique. The $\mathcal{K}$ codes are then combined as rows of a matrix $\mathbf{M}$ where each column represents a dichotomy $h_j(\mathbf{x})$ and each

row represents a class (see Fig. 2). A value 1 in position $\mathbf{M}(k, j)$ means that the j^{th} dichotomy classifies an unknown example $\mathbf{x}$ as belonging to the class k, a value -1 means that it belongs to the class $q \neq k$ and a 0 value means that we do not care about the classification result, regarding class k. Therefore, in order to classify $\mathbf{x}$ a codeword cw is computed by testing the feature vector with all the dichotomies $\boldsymbol{h}$. The vector $\mathbf{d} = [d_1, \ldots, d_k]$ of the distances between cw and each row m_k of the matrix $\mathbf{M}$ is computed: the inferred class is the value k reporting the minimum distance.

It has been recently demonstrated [12] that the distance in the ECOC space can be used as an information related to the classification reliability, when a basic *margin classifier* is used as dichotomy; in this work both Adaptive Boosting (AdaBoost) [13] and Support Vector Machine (SVM) [14] binary margin classifiers are used.

During the decoding process, for each cell G_i with feature vector $\mathbf{x}_i$ we obtain a vector $\boldsymbol{\mathcal{L}}(\mathbf{x}_i) = [\mathcal{L}_1, \ldots, \mathcal{L}_K]$, where each value $\mathcal{L}_k$ can be considered as related to the probability, or pseudo-probability of the cell to belong to a class: $\mathcal{L}_k \approx P(y_i = k|\mathbf{x}_i)$ (see Fig. 3(b,c)). Given the vector $\boldsymbol{\mathcal{L}}$, the label y_i is assigned to the cell through the Maximum A-Posteriori Probability (MAP) over the $\mathcal{L}$ values; it is worth to note that this procedure is equivalent to consider the index of the row with the minimum distance with respect to cw in the ECOC space.

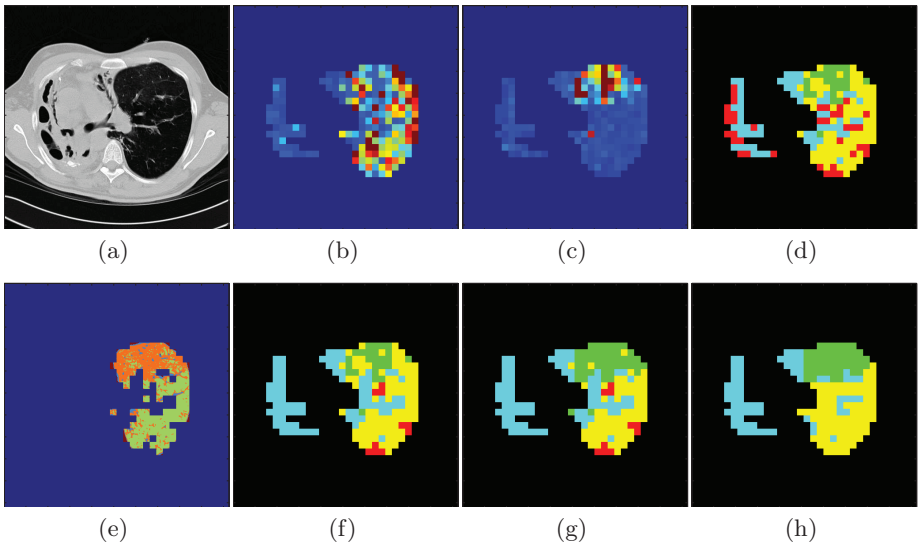


Fig. 3. Example of automatic lung CT image labeling. Original CT image (a); Pseudo probabilities $\mathcal{L}_{AT}$ (b) and $\mathcal{L}_{NOR}$ (c) as output of the ECOC-based framework. MAP labeling by ECOC (d). Output of the unsupervised clustering, with $C = 4$ (e). MAP after post-processing (f). Final automatic labeling (g), compared with the manual labeling (h).

2.3 Unsupervised labeling

Owing to differences in disease severity, inspiration level and scanner imaging parameters, ATR tissue in one scan may look very similar to NOR tissue in another scan. This phenomenon causes the supervised classification to confuse the two classes. However, within a scan, regions of ATR and NOR tissues do appear different. This motivated us to apply a further labeling step based on unsupervised clustering.

Given the output labels map $Y = \{Y_{INF}, Y_{ATR}, Y_{NOR}, Y_{BUL}\}$ for a CT slice (see Fig. 3(d)), we consider as input for the unsupervised labeling solely the subset of cells in $\widetilde{Y} = \{Y_{ATR}, Y_{NOR}\}$. The assumption is that the classification output based on supervised learning for the class INF and BUL is sufficiently reliable. In the cells contained in $\widetilde{Y}$, features related to the tissue HU value are considered, i.e., the gaussian smoothing at the three different scales along with the original image value and the median value. Differently from what has been done in the supervised learning, the features values are now considered *per voxel*, assigning a feature vector $\mathbf{x}_{HU} \in \Re^5$ to each voxel of the slice. In this way, it is easier to cluster even small sets of points that are likely to belong to ATR and NOR tissue per slice. The set of all features vector $\mathbf{X}_{HU}$ in $\widetilde{Y}$ are then used as input of a *Fuzzy C-Means* unsupervised clustering [15]. Since the lung region in $\widetilde{Y}$ is affected by area quantization, voxels belonging to either the almost homogeneous body area or to small bright vessels are also included in $\widetilde{Y}$. For this reason two backup classes are foreseen: the number of chosen clusters is $C = 4$.

Once the voxels are clustered we need to assign the correct label for ATR and NOR tissues. For this purpose, the pseudo-probability $\boldsymbol{\mathcal{L}}$ is used (Fig. 3(b,c)). The hypothesis is that, despite the fact that the two tissues can be misclassified in some cells, the information provided by the pseudo-probability is correct for most of the cells. Given the number of clusters C, let us define as $\mathcal{Q}$ the set of all the combinations of pairs of C elements. Let us also define as V_A and V_B the set of voxels belonging to the first and second group (A and B) of each pair in $\mathcal{Q}$, respectively. The correct unsupervised labeling for $\widetilde{Y}$ is the one that maximizes the product of the mean pseudo-probability for V_A and V_B:

$$(\widetilde{Y}_{ATR}, \widetilde{Y}_{NOR}) = argmax \frac{1}{|V_A||V_B|} \left(\sum_{\boldsymbol{p} \in V_A} \mathcal{L}_{ATR}(\boldsymbol{p}) \sum_{\boldsymbol{q} \in V_B} \mathcal{L}_{NOR}(\boldsymbol{q}) \right),$$

where $\boldsymbol{p}$ and $\boldsymbol{q}$ represent the voxel location in the bi-dimensional CT image domain, and $|\cdot|$ indicates the cardinality of a set. The same criterion used in [2] for labeling the cells is then applied: each cell takes the label represented by more than 50% of the enclosed voxels.

2.4 Post processing

In some cases the ATR tissue, due to its local dark appearance, is confused with BUL tissue. For this reason, in order to filter out misclassified BUL cells, we

consider the *Minimal Bulla Area* (MBA) parameter (in mm^2) in the training set as an a-priori knowledge on the BUL tissue morphology. The area A_{BUL} of each region (set of cells) labeled as BUL is computed: if the condition $A_{BUL} \leq \frac{MBA}{2}$ holds, the label of those cells is changed. The new label corresponds to the class next to BUL in the *ranked* pseudo-probabilities $\boldsymbol{\mathcal{L}}_{rank}$ for those cells.

Finally, it can be observed that the inner part of extremely enlarged airways, scored as inflammation by the observer, often appears as a dark and homogeneous region in the CT image: this may cause misclassification between INF and BUL tissue. In order to avoid this problem, we consider the tissue surrounding each region labeled as BUL: if the amount of surrounding tissue has been classified as INF in at least a percentage $\mathcal{T}$, learned by training examples, the labels of cells labeled as BUL are switched to INF.

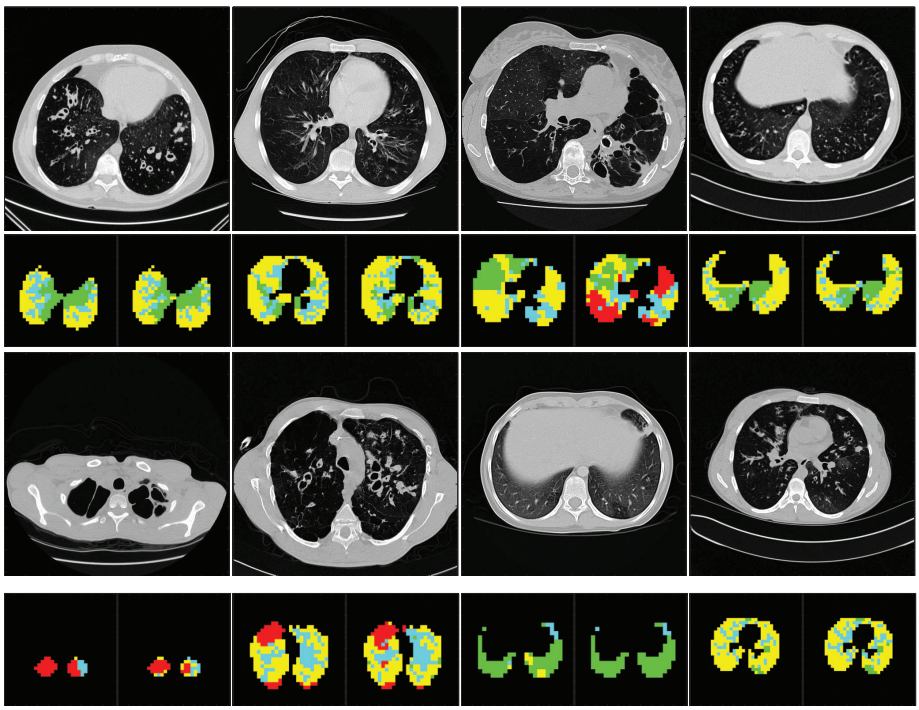


Fig. 4. Examples of automatic tissue classification on eight different scans. First / third row: CT images; second / forth row: manual labeling (a,c,e,g) and automatic labeling (b,d,f,h).

3 Experimental results

Materials Clinical data and chest CT scans of CF patients screened for LTx between 1990 and 2005 were collected from 17 centers worldwide. Scans have

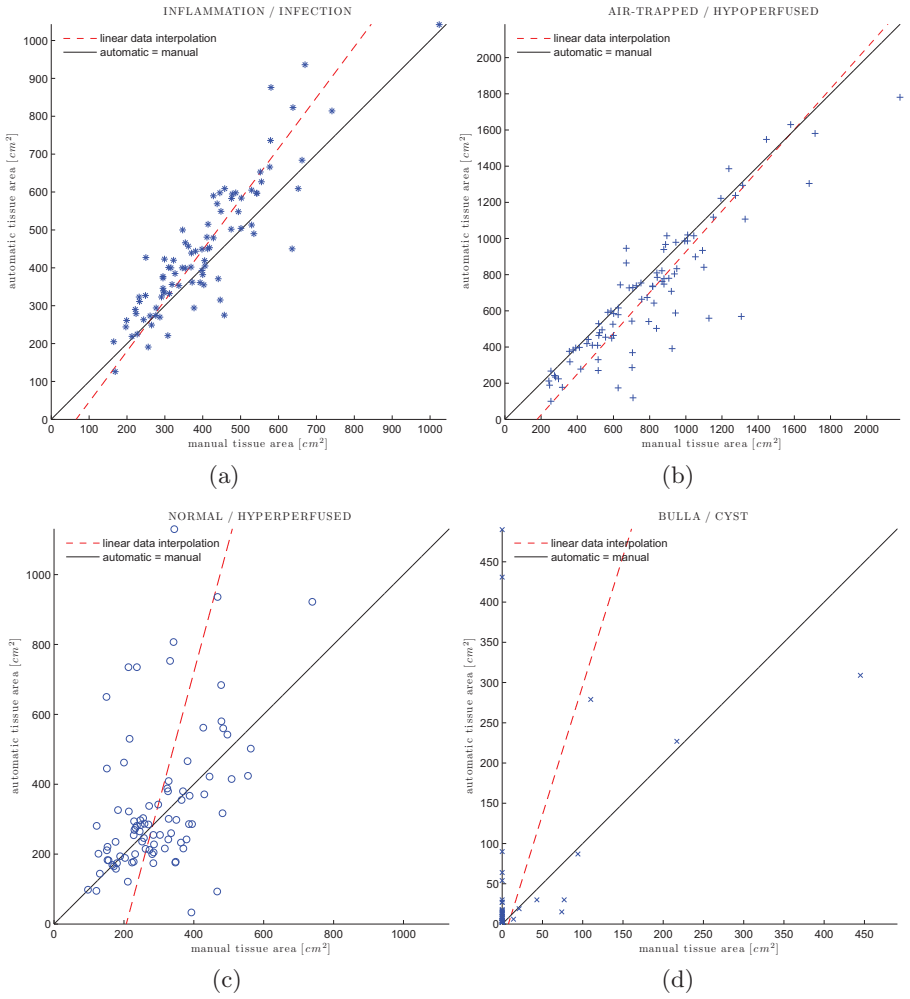


Fig. 5. Scatter plots for the amount of classified tissues (in cm^2) per scan, compared with manual annotations. Linear interpolation of data is also provided.

	INF_{auto}	ATR_{auto}	NOR_{auto}	BUL_{auto}
INF_{gt}	27423	3602	3198	881
ATR_{gt}	8186	45522	13064	2231
NOR_{gt}	3009	4326	18747	27
BUL_{gt}	83	129	7	876

(a)

	INF_{auto}	ATR_{auto}	NOR_{auto}	BUL_{auto}
INF_{gt}	27883	4149	2818	254
ATR_{gt}	8194	49708	9628	1473
NOR_{gt}	2937	6126	17042	4
BUL_{gt}	233	141	23	698

(b)

Table 1. Confusion matrix for tissue labeling by SVM (a) and $\widehat{SVM}$ (b). The amount of manual (gt) and automatic ($auto$) labeled cells are indicated.

been acquired with different models of scanners from four vendors; $N_P = 89$ volumetric scans were randomly selected. For each scan, slices at a distance of 3 *cm* were selected and labeled, according to the procedure described in [10]: the total number of available CT slices is $N_S = 654$.

Validation For each region, the set of features described in section 2.1 was extracted by using the kernel parameters $(\sigma_1; \sigma_2; \sigma_3) = (0.43; 0.86; 1.3)$ mm. The margin classifiers used in the ECOC framework are Adaptive Boosting (AdaBoost) [13] with Decision Stumps as weak learner and Support Vector Machine (SVM) [14] with Radial Basis kernel. AdaBoost was trained with $T = 300$ iterations, while the kernel parameters for SVM were estimated from the training set as described in [16]. In both cases the ECOC matrix was designed according to the *One-vs-One* coding and the Attenuated Euclidean Distance decoding, resulting in a 4×6 ECOC matrix (see Fig. 2). The number of samples per class used in the training set was 1500, randomly selected from the training set; this value was chosen in order to keep balanced samples from different classes. The threshold for the bulla filtering $\mathcal{T}$ was set to 40% from training set, while the minimal bulla area was computed as $MBA = 203.12$ mm^2.

In order to completely separate the test samples from the training set, the *leave-one-scan-out* cross-validation technique was used. A number of N_P folds were considered for training: at each fold, all the samples belonging to one scan were discarded, and the rest used for training. The values of the 4×4 confusion matrix computed at each fold were cumulated into a global confusion matrix, used for the computation of the classification performance (see Table 1).

The method has been validated through the following experiments. First, the performance of the ECOC framework was compared with the one of a Knn classifier where either the 17 tissue descriptor or the HU-related features are used; in this case AdaBoost was used as binary classifier in ECOC. The aim of this experiment was two-fold: first, it established whether the texture information, additionally to the HU-based descriptors, is relevant in tissue classification. Furthermore, it directly compares two different approaches for solving multi-class problems; in all the experiments the value $K = 100$ was chosen for Knn. The ECOC framework using SVM was then compared with previous approaches. In order to assess the validity of both unsupervised labeling and post-processing, the two steps were applied, independently (SVM' and SVM") and combined ($\widehat{SVM}$), to the map labels obtained by SVM. The results of the experiments are shown in Table 2, where the following classification performance are considered: overall accuracy A, sensitivity S and specificity K per tissue.

4 Discussion

The experimental results demonstrated the benefit of using texture features in the supervised classification (see Table 2): the use of 17 tissue descriptors improved classification performance for both Knn and AdaBoost, when compared with the ones obtained by solely using HU-based features (Knn_{HU},$Boost_{HU}$).

	A	S_{INF}	S_{AT}	S_{NOR}	S_{BUL}	K_{INF}	K_{AT}	K_{NOR}	K_{BUL}
Knn_{HU}	65.1%	78.7%	56.2%	70.1%	76.1%	84.6%	89.8%	80.5%	96.8%
Knn	67.1%	82.5%	60.0%	64.9%	72.8%	82.0%	88.1%	85.8%	97.2%
$Boost_{HU}$	63.8%	74.9%	57.1%	65.7%	78.6%	83.1%	86.0%	81.9%	97.1%
$Boost$	66.6%	76.2%	62.0%	65.5%	79.4%	84.2%	85.5%	84.5%	97.5%
SVM	70.5%	78.1%	65.9%	71.8%	80.0%	88.3%	87.0%	84.5%	97.6%
SVM'	71.5%	79.6%	66.6%	73.2%	63.1%	87.9%	87.3%	84.5%	98.8%
SVM"	71.7%	77.9%	70.3%	66.4%	78.3%	88.3%	84.0%	87.7%	97.7%
$\widehat{SVM}$	72.6%	79.4%	72.0%	65.3%	63.7%	88.2%	83.3%	88.1%	98.7%

Table 2. Classification performance parameters for the methodologies considered in the comparison. The overall accuracy A as well as the sensitivity S and the specificity K per tissue are indicated. SVM' indicates the labeling obtained by using solely SVM and the post processing; SVM" indicates the labeling obtained by using solely SVM and the unsupervised labeling; $\widehat{SVM}$ indicates the output of the whole framework.

No remarkable difference has been noticed in the comparison between Knn and the ECOC framework when AdaBoost is used as margin classifier. An evident improvement is observed when SVM is used in the ECOC framework instead. This fact corroborates the assumption of heterogeneity in the dataset. Since AdaBoost with Decision Stumps performs features selection, it tends to discard features that mislead the classification of the training set, while they can be significant in test samples. On the other hand, SVM does consider all the features: this fact, combined with the fine kernel tuning and the error-correcting capability of the ECOC framework, leads to better classification results.

The effectiveness of both unsupervised labeling and bulla refinement has been demonstrated, since classification performance improves when using them both separately and in conjunction. Owing to the lack of a lung segmentation, voxel-based classification in the unsupervised labeling has been necessary, in order to avoid the influence of regions belonging to the body in the characterization of cells placed at the lung border; this problem can be avoided by using an accurate lung segmentation.

The sensitivity and specificity for the INF tissue are 79.6% and a specificity of 87.9%, respectively. Since it has been recently demonstrated that the INF score has a significant predictive value for mortality [2], the proposed methodology can be used as en effective system for automatic SALD scoring for risk assessment.

The SALD scoring takes into account the amount of tissue per volume. For this reason, we computed the amount of tissue labeled by the automatic framework and we compared it with the manual annotation (see Fig. 5). It can be noticed that INF and ATR tissues highly correlate with manual scoring, thus corroborating our hypothesis. An excessive reduction of True Positives for BUL regions is experienced in both SVM' and $\widehat{SVM}$, while the number of False Positives remains high. In particular, though improved with respect to SVM

(see Table 1), the confusion with ATR is still remarkable. As a straightforward improvement, along with the assumption on minimal bulla area, the filtering process could be guided by the context in the CT slice. In this way, not only the local properties of cells is considered, but a more global relationship with surrounding tissues. Such approach could also include the knowledge acquired by unsupervised labeling, to further reduce the confusion between ATR and NOR tissues (see Table 1).

Finally, it is worth to note that only one expert has been involved in the manual labeling: both training and validation process may be then affected by the lack of agreement / disagreement on tissue definition. An exhaustive comparison with inter- and intra-observer variability using labels from multiple observers represents a meaningful analysis in our research.

5 Conclusions

A framework for lung tissue classification in Computed Tomography (CT) scans has been presented. The method demonstrated to perform accurately on a large heterogeneous database, collected by multiple hospitals and multiple scanners. Furthermore, the presented technique can be considered as a valid tool for the automatic assessment of SALD scoring.

6 Acknowledgement

CT CFS study Group: Aitken ML, Assael BM, Aurora P, Brink M, Dodd JD, Dupont LJ, Faro A, Ferkol T, Fraioli F, Gallagher CC, van de Graaf EA, Grasemann H, Hadjiliadis D, van Hal P, Kanaan R, Kraan J, Lala B, Lindblad A, Mallory Jr. GB, Nossent GD, Ocampo M, Quattrucci S, Robinson P, Sermet I, Solomon M, Spencer H, Staab D, Stephen MJ, Verschuuren EA, Williams TJ.

References

1. Levy, H., Kalish, L., Huntington, I., Weller, N., Gerard, C., Silverman, E., Celedon, J., Pier, G., Weiss, S.: Inflammatory markers of lung disease in adult patients with cystic fibrosis. Pediatric Pulmonology **42** (2007) 256 – 262
2. Loeve, M., Hal, P.V., Robinson, P., Aitken, M., Dodd, J., de Bruijne, M., Hop, W., Tiddens, H.: Ct scores correlate with survival in cf patients awaiting lung transplantation. European Cystic Fibrosis Conference (2011)
3. Chabat, F., Yang, G.Z., Hansell, D.M.: Obstructive lung diseases: Texture classification for differentiation at ct. Radiology **228** (2003) 871 – 877
4. Sluimer, I.C., Prokop, M., van Ginneken, I.H.B.: Automated classification of hyperlucency, fibrosis, ground glass, solid, and focal lesions in high-resolution ct of the lung. Medical Physics **33 (7)** (2006) 2610 – 2620
5. Gangeh, M.J., Sørensen, L., Shaker, S.B., Kamel, M.S., de Bruijne, M., Loog, M.: A texton-based approach for the classification of lung parenchyma in ct images. MICCAI 2010, Part III, LNCS **6363** (2010) 596 – 603

6. Sørensen, L., Shaker, S.B., de Bruijne, M.: Quantitative analysis of pulmonary emphysema using local binary patterns. IEEE Transactions on Medical Imaging **29** (2010) 559 – 569
7. Sørensen, L., Loog, M., Lo, P., Ashraf, H., Dirksen, A., Duin, R.P.W., de Bruijne, M.: Image dissimilarity-based quantification of lung disease from ct. MICCAI2010, LNCS (2010)
8. Depeursinge, A., Racoceanu, D., Iavindrasana, J., Cohen, G., Platon, A., Poletti, P.A., Muller, H.: Fusing visual and clinical information for lung tissue classification in high-resolution computed tomography. Artificial Intelligence in Medicine **50** (2010) 13 – 21
9. 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. IEEE Transactions on Medical Imaging **25 (4)** (2006) 464 – 475
10. Loeve, M., van Hal, P.T.W., Robinson, P., de Jong, P.A., Lequin, M.H., Hop, W.C., Williams, T.J., Nossent, G.D., Tiddens, H.A.: The spectrum of structural abnormalities on ct scans from patients with cf with severe advanced lung disease. Thorax **64** (2009) 876 – 882
11. Dietterich, T.G., Bakiri, G.: Solving multiclass learning problems via error-correcting output codes. JAIR **2** (1995) 263–286
12. Ciompi, F., Pujol, O., Radeva, P.: A meta-learning approach to conditional random fields using error-correcting output codes. IEEE International Conference on Pattern Recognition (ICPR) (2010) 710 – 713
13. R.Schapire: The boosting approach to machine learning: An overview. MSRI (2001)
14. C.Cortes, V.Vapnik: Support-vector networks. Machine Learning **20(3)** (1995) 273–297
15. Bezdek, J.: Pattern Recognition with Fuzzy Objective Function Algorithms. New York (1981)
16. R.Rifkin, A.Klautau: In defense of one-vs-all classification. JMLR **5** (2004) 101–141