Textured Graph–based Model of the Lungs: Application on Tuberculosis Type Classification and Multi–Drug Resistance Detection

Yashin Dicente Cid[1,2], Kayhan Batmanghelich[3], and Henning Müller[1,2]

[1] University of Applied Sciences Western Switzerland (HES–SO), Sierre, Switzerland;
[2] University of Geneva, Switzerland;
[3] University of Pittsburgh, USA
yashin.dicente@hevs.ch

Abstract. Tuberculosis (TB) remains a leading cause of death worldwide. Two main challenges when assessing computed tomography scans of TB patients are detecting multi–drug resistance and differentiating TB types. In this article we model the lungs as a graph entity where nodes represent anatomical lung regions and edges encode interactions between them. This graph is able to characterize the texture distribution along the lungs, making it suitable for describing patients with different TB types. In 2017, the ImageCLEF benchmark proposed a task based on computed tomography volumes of patients with TB. This task was divided into two subtasks: multi–drug resistance prediction, and TB type classification. The participation in this task showed the strength of our model, leading to best results in the competition for multi–drug resistance detection (AUC = 0.5825) and good results in the TB type classification (Cohen's Kappa coefficient = 0.1623).

Keywords: lung graph model, 3D texture analysis, tuberculosis

1 Introduction

Tuberculosis (TB) is an infectious disease that remains a persistent threat and a leading cause of death worldwide. An important task is to detect when the TB organisms become resistant to standard drugs. The multi–drug resistant (MDR) form of the disease is a difficult and expensive form to recover from. The gold–standard methods for MDR detection are expensive and may take up to several months [1]. Therefore, there is a need for quick and cheap methods of MDR detection. The identification of TB types (TBT) is another important task, as different types of TB may require different treatments. Several visual patterns can be seen in a Computed Tomography (CT) volume of a patient with TB in the lungs, some of them characteristic of a specific TB type. However, the final classification of the disease requires additional analyses, besides the CT images [2]. An automatic image analysis system that can identify holistic patterns of lungs with TB, not evident through simple visual assessment of CT images, can be very useful for radiologists.

Graph modeling is a complete framework that was previously proposed for brain connectivity analysis but has rarely been applied to other organs [3]. Graph methods divide the brain into fixed anatomical regions and compare neural activations between regions [4]. In [5], we presented a basic graph model of the lungs capable of differentiating between pulmonary hypertension and pulmonary embolism. Both diseases present similar visual defects in lung CT scans. However, they differ in their distribution throughout the lung. The graph was based on dividing the lung into several regions and using these regions as nodes of a graph. The regions were described using Hounsfield Unit (HU) distributions, extracted from Dual Energy CT (DECT) scans. Preliminary results showed that a single CT did not contain enough information about the HU distribution to differentiate between the diseases and a more advanced description of the regions was needed.

In this article, we present a more complex version of the graph model to characterize the lungs. The new model describes each region of the lung using state–of–the–art 3D texture descriptors. Our hypothesis is that a holistic analysis of the relations between regional texture features is able to encode subtle differences between patients with separate TB types and to assist in an early detection of drug resistance patients. We tested our texture–based graph model of the lungs in the ImageCLEF 2017 TB challenge, where it was compared against 8 other methods, obtaining the best results in the MDR detection task. The following section contains a brief overview of the subtasks and dataset of the ImageCLEF 2017 TB task. More detailed information on the task can be found in the overview article [6]. Section 3 explains the process of building the texture graph model of the lungs and all the variations tested for this task in detail. The results obtained by this approach in the task are shown in Section 4. Finally, Section 5 concludes with lessons learned working on the data with our approach.

2 ImageCLEF Challenge

The ImageCLEF (Image retrieval and image analysis evaluation campaign of the Cross-Language Evaluation Forum, CLEF) has organized challenges on image classification and retrieval since 2003 [7]. Since 2004, medical image retrieval and analysis tasks have been organized [8, 9]. The ImageCLEF 2017 [10] challenge included a task based on CT volumes of patients with TB, the ImageCLEF 2017 TB task [6]. In this task, a dataset of lung CT scans was provided and two subtasks were proposed. For both subtasks volumetric chest CT images with different voxel sizes and automatic segmentations of the lungs were provided.

MDR detection task: This subtask was a 2–class problem that consisted on detecting MDR based on a series of CT images. The dataset was composed of 444 CT volumes, divided into training and test sets as shown in Table 1. By visual inspection, the CT volumes of this task did not present any relevant visual difference that distinguished MDR from drug–sensitive (DS) patients.

Table 1. Number of CT images for the MDR detection and TBT classification tasks.

MDR dataset

Patient set	Train	Test
DS	134	101
MDR	96	113
Total patients	230	214

TBT dataset

Patient set / TBT	Train	Test
Infiltrative	140	80
Focal	120	70
Tuberculoma	100	60
Miliary	80	50
Fibro–cavernous	60	40
Total patients	500	300

TBT classification task: The TBT subtask was a classification problem with five classes, corresponding to five TB types: Infiltrative, focal, tuberculoma, miliary, and fibro–cavernous. The patterns present in the several TBT patients were already quite discriminative, *e.g.* the patients with fibro–cavernous TB presented distinctive caverns in their CT image (see Figure 1). The dataset for this subtask consisted of 800 CT volumes. The detailed number of patients for each class is shown in Table 1.

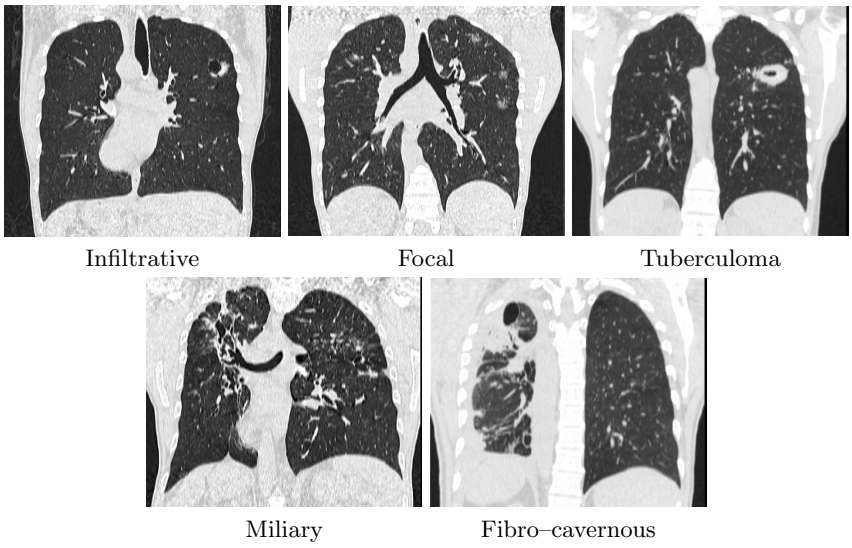


Fig. 1. Examples of the five tuberculosis types in the TBT subtask. The CT slices are shown using a HU window with center at -500 HU and width of 1400 HU

3 Texture–based Graph Model of the Lungs

We propose a general pipeline (see Figure 2) to automatically obtain a texture–based graph model of the lungs that is composed of four steps: 1) automatic

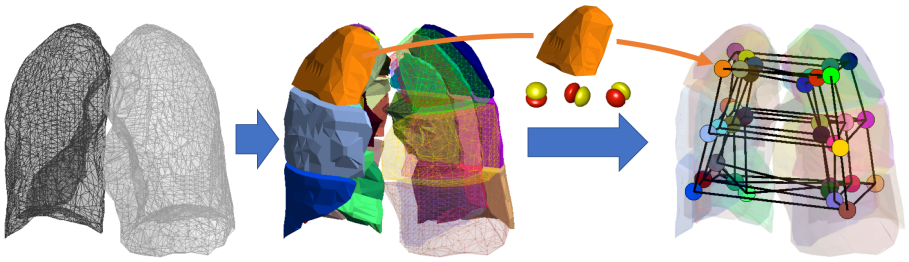


Fig. 2. Construction pipeline of a texture–based graph model of the lungs: First, the lungs are automatically segmented. Then, they are divided using a geometric atlas with 36 regions. From each region, texture features are extracted. Finally, the graph is built using the regions in the atlas as nodes. The edges contain the similarities between the 3D texture descriptors.

segmentation of the lung fields; 2) division of the lung mask into regions; 3) extraction of local biomedical texture features in each region; and 4) construction of the lung graph encoding the comparison between the regional features. Following this general pipeline, several graph models were investigated in this work and tested on the ImageCLEF 2017 TB task. The models obtained were produced varying the texture descriptors used in each lung region, the number of connections used to build the graph (edges), and the type of comparison between the regional features (weights).

Preprocessing pipeline: The graphs built were based on 3D texture features that require having isometric voxels. Therefore, we first resampled the CT volumes and the masks. After analyzing the multiple resolutions and the inter–slice distances found in the dataset, a voxel size of 1 mm was used to capture a maximum of information. We used the lung masks provided by the organizers, that were obtained with the method described in [11].

Geometric atlas of the lungs: In this article we use as a lung division the atlas introduced by Depeursinge *et al.* [12]. This atlas contains 36 geometric regions produced by intersecting four axis segmentations: coronal (right/left), sagital (anterior/posterior), vertical (apical/central/basal), and axial (peripheral/middle/central). These regions are based on the 3D model of the lung presented by Zrimec *et al.* [13]. Each region of the atlas is referred to as r. Figures 2, 3, and 4 contain a 3D visualization of this atlas.

3.1 3D Texture Features

Two state–of–the–art 3D texture feature types were selected to describe the texture in each atlas region r. The first method is a histogram of gradients based on the Fourier transform HOG (FHOG) introduced in [14]. We used 28 3D directions for the histogram obtaining a 28–dimensional feature vector per image

voxel v ($\mathbf{f}_H(v) \in \mathbb{R}^{28}$). The second descriptor is the locally–oriented 3D Riesz–wavelet transform introduced by Dicente *et al.* in [15]. The parameters used in this work correspond to the ones obtaining the best classification results of synthetic 3D textures in the above mentioned article. These are: 3rd–order Riesz transform, 4 scales and 1st–order alignment. This configuration provides 40–dimensional feature vectors for each image voxel. The feature vector for a single voxel was then defined as the weighted sum of the absolute Riesz response along the 4 scales, obtaining a 10–dimensional feature vector ($\mathbf{f}_{\mathcal{R}}(v) \in \mathbb{R}^{10}$). Finally, the average (μ) and standard deviation (σ) of these descriptors were extracted from each region r, hence obtaining four region descriptors (see Equation 1).

$$\begin{aligned} \boldsymbol{\mu}_H(r) &= \mu_{v\in r}(\mathbf{f}_H(v)) \qquad & \boldsymbol{\mu}_{\mathcal{R}}(r) &= \mu_{v\in r}(\mathbf{f}_{\mathcal{R}}(v)) \\ \boldsymbol{\sigma}_H(r) &= \sigma_{v\in r}(\mathbf{f}_H(v)) \qquad & \boldsymbol{\sigma}_{\mathcal{R}}(r) &= \sigma_{v\in r}(\mathbf{f}_{\mathcal{R}}(v)) \end{aligned} \tag{1}$$

3.2 Graph Model of the Lungs

A graph is a structure that contains a set of nodes $\mathcal{N}$ and a set of relations between the nodes, called set of edges $\mathcal{E}$. In particular, edge–weighted graphs are graphs in which a value is assigned to each edge, *i.e.*, there is a function $w : \mathcal{E} \to \mathbb{R}$. From now on, the graphs in this work are considered to be edge–weighted graphs with no self–loops.

Given a division of the lungs with n regions $\{r_1, \ldots, r_n\}$, we define a *graph model of the lungs* $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ as a set of n nodes $\mathcal{N} = \{N_1, \ldots, N_n\}$ connected by a set of m edges $\mathcal{E}$. $E_{i,j}$ is defined as the edge connecting nodes N_i and N_j with associated weight $w_{i,j}$. The weights are functions of the texture features extracted in the lung regions. Using the 36–region atlas as a lung division, the graphs were finite with 36 nodes, *i.e.*, $\mathcal{N} = \{N_1, \ldots, N_{36}\}$. Figure 3 contains a 3D visualization of the graph elements using this atlas.

Using a fixed number of nodes and the same connections for all patients allow us to compare the patient graphs by comparing their adjacency matrices. Given a graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ with 36 nodes, its adjacency matrix $\mathcal{A}$ is defined as the 36×36 square matrix with elements $a_{i,j} = w_{i,j}$ if $E_{i,j}$ exists, and $a_{i,j} = 0$ otherwise. Since no self–loops are allowed, $a_{i,i} = 0\ \forall i \in \{1, \ldots, 36\}$. These matrices can be characterized by the ordered list of their elements in a vector form. Then, the comparison of graphs of different patients can be reduced to a vector comparison, where standard machine learning techniques can be directly applied. For a patient p with graph $\mathcal{G}_p$ and adjacency matrix $\mathcal{A}_p$, we define the *patient descriptor* $\boldsymbol{w}_p$ as the ordered list of weights $w_{i,j}$, using the order induced by the vectorization of the matrix $\mathcal{A}_p$.

3.3 Graph Architectures

Using the geometric atlas with 36 regions as a base, several undirected weighted graphs were defined varying the number of edges and their weights. The number of edges varied according to the connections considered between the nodes. The different configurations of node connections correspond to different pruning levels

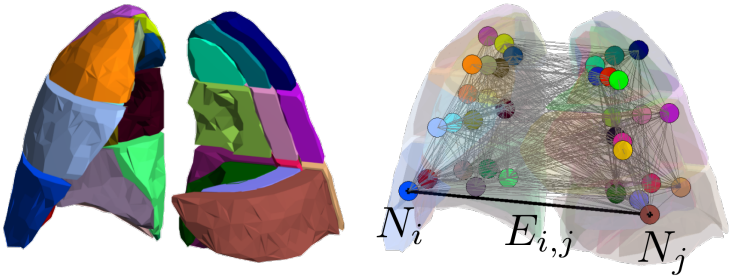


Fig. 3. Prototype visualization of the graph elements defined in Section 3.2. Left: 3D visualization of the geometric atlas used where six regions are not visible to show the atlas interior divisions. Right: Complete graph built from the geometric atlas. N_i and N_j are the nodes corresponding to regions r_i and r_j, respectively. $E_{i,j}$ is the edge connecting the nodes N_i and N_j. All the other edges are shown in light gray.

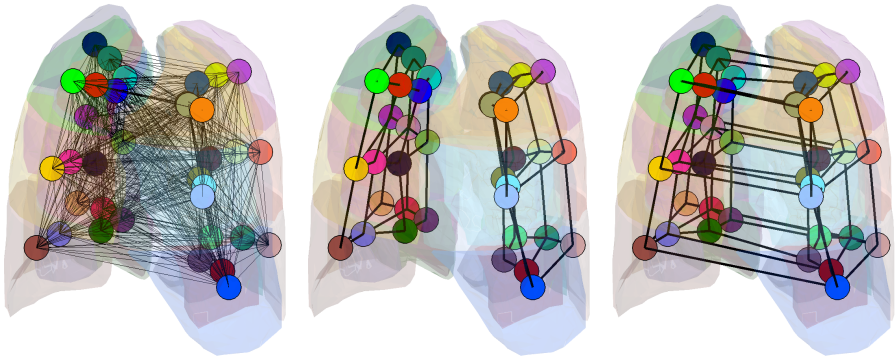
of the complete graph with 36 nodes and we refer to them as *graph architectures*. Three graph architectures were designed (shown in Figure 4), each one with a different adjacency matrix $\mathcal{A}$. Figure 5 shows the matrix $\mathcal{A}$ for each graph architecture using the same local features. The difference between them is the number of elements $a_{i,j} \in \mathcal{A}$ informed.

- *Graph_Complete*: This is the complete 36–node graph. For every pair of nodes N_i and N_j with $i \neq j$ an undirected edge $E_{i,j}$ exists. The total number of edges in this case is 630 ($\frac{36 \cdot 35}{2}$).
- *Graph_66*: Based on the region adjacency defined by the geometric atlas, there is an edge $E_{i,j}$ between nodes N_i and N_j if regions r_i and r_j are neighbors in the atlas, *i.e.*, if the regions are 3D adjacent. This graph contains in total 66 edges.
- *Graph_84*: This graph architecture has the same 66 edges as Graph_66. In addition, it includes 18 edges connecting each pair of nodes representing symmetric regions in the atlas with respect the left–right division of the lungs.

3.4 Graph–based Patient Descriptors

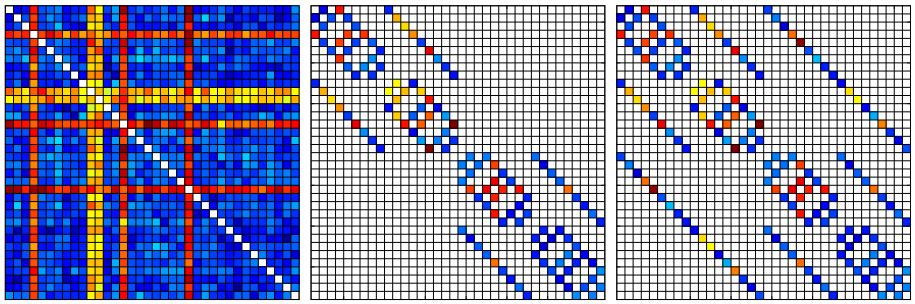
The weight $w_{i,j}$ of an edge $E_{i,j}$ was defined using four different measures between the features of the corresponding nodes N_i and N_j. Considering $\mathbf{f}_i$ and $\mathbf{f}_j$ the feature vectors of regions r_i and r_j respectively, the measures used are:

- *Correlation distance (corr)*: $w_{i,j} = 1 - corr(\mathbf{f}_i, \mathbf{f}_j)$.
- *Cosine similarity (cos)*: $w_{i,j} = 1 - cos(\mathbf{f}_i, \mathbf{f}_j)$.
- *Euclidean distance (euc)*: $w_{i,j} = \|\mathbf{f}_i - \mathbf{f}_j\|_2$.
- *Norm of the sum (sumNorm)*: $w_{i,j} = \|\mathbf{f}_i + \mathbf{f}_j\|_2$.



Graph_Complete Graph_66 Graph_84

Fig. 4. 3D visualization of the three graph architectures (or pruning levels) designed.



$\mathcal{A}$ for the Graph_Complete $\mathcal{A}$ for the Graph_66 $\mathcal{A}$ for the Graph_84

Fig. 5. Example of adjacency matrices for the three graph architectures shown in Figure 4. The matrices were created using the same features in each atlas region. Hence, the elements present in the three matrices contain the same values. The adjacency matrices differ in the non informed elements (blanks).

The feature vector $\boldsymbol{w}_p$ of a patient p is defined as the ordered list of weights $w_{i,j}$ (see Section 3.2). The adjacency matrices $\mathcal{A}_p$ are symmetric, and only half of the elements are needed to characterize them. Depending on the graph used, this feature vector is 630–, 66–, or 84–dimensional.

3.5 Graph–based Patient Descriptor Fusion and Classification

Given a graph model of the lungs, a patient descriptor vector $\boldsymbol{w}_p$ was obtained. In particular, for each of the regional texture descriptors extracted ($\boldsymbol{\mu}_H$, $\boldsymbol{\mu}_{\mathcal{R}}$, $\boldsymbol{\sigma}_H$, and $\boldsymbol{\sigma}_{\mathcal{R}}$), a different graph model was obtained, thus generating a different patient descriptor $\boldsymbol{w}_p$. For the classification experiments, we tested several combinations of these patient descriptors. Therefore, we defined the *derived patient descriptor vector* $\hat{\boldsymbol{w}}_p$, containing a combination of these patient descriptors $\boldsymbol{w}_p$.

In this section, the different steps designed to obtain the derived patient descriptor vectors $\hat{\boldsymbol{w}}_p$ are explained. Moreover, the classification setup is detailed.

Patient descriptor normalization: The patient descriptors $\boldsymbol{w}_p$ were normalized with respect to the set of training patients P_{trn}. Two normalizations were tested: Z–score and box normalization between 0 and 1, referred to as *Gauss-Norm* and *[0,1]* respectively. Since each component of a vector $\boldsymbol{w}_p$ corresponds to the weight of a different edge in the graph, the normalizations were performed over all the vector components together to preserve the relations induced by the graph structure. $\bar{\boldsymbol{w}}_p$ denotes the normalized patient descriptor of a patient p.

Patient descriptor concatenation: Fixing a graph structure (Graph_Complete, Graph_66, or Graph_84) and a measure between the regional features (*corr*, *cos*, *euc*, or *sumNorm*) (see Sections 3.3 and 3.4), four normalized patient descriptor vectors $\bar{\boldsymbol{w}}_p$ were obtained. These are: $\bar{\boldsymbol{w}}_{\boldsymbol{\mu}_H}$, $\bar{\boldsymbol{w}}_{\boldsymbol{\sigma}_H}$, $\bar{\boldsymbol{w}}_{\boldsymbol{\mu}_\mathcal{R}}$, and $\bar{\boldsymbol{w}}_{\boldsymbol{\sigma}_\mathcal{R}}$. Five concatenations of these descriptors were tested in our experiments in order to better describe each patient:

- Mean and std of FHOG: $\hat{\boldsymbol{w}} = (\bar{\boldsymbol{w}}_{\boldsymbol{\mu}_H} || \bar{\boldsymbol{w}}_{\boldsymbol{\sigma}_H})$.
- Mean and std of 3DARiesz: $\hat{\boldsymbol{w}} = (\bar{\boldsymbol{w}}_{\boldsymbol{\mu}_\mathcal{R}} || \bar{\boldsymbol{w}}_{\boldsymbol{\sigma}_\mathcal{R}})$.
- Mean of FHOG and 3DARiesz: $\hat{\boldsymbol{w}} = (\bar{\boldsymbol{w}}_{\boldsymbol{\mu}_H} || \bar{\boldsymbol{w}}_{\boldsymbol{\mu}_\mathcal{R}})$.
- Std of FHOG and 3DARiesz: $\hat{\boldsymbol{w}} = (\bar{\boldsymbol{w}}_{\boldsymbol{\sigma}_H} || \bar{\boldsymbol{w}}_{\boldsymbol{\sigma}_\mathcal{R}})$.
- Mean and std of FHOG and 3DARiesz: $\hat{\boldsymbol{w}} = (\bar{\boldsymbol{w}}_{\boldsymbol{\mu}_H} || \bar{\boldsymbol{w}}_{\boldsymbol{\sigma}_H} || \bar{\boldsymbol{w}}_{\boldsymbol{\mu}_\mathcal{R}} || \bar{\boldsymbol{w}}_{\boldsymbol{\sigma}_\mathcal{R}})$.

Feature space reduction: The dimension of the feature space was much larger than the number of patients in some of the experiments, *e.g.* when using the Graph_Complete architecture or the feature concatenations. To avoid the known problems of overfitting, two feature space reduction techniques were tested, both applied in the training phase. The first one selected the dimensions that best correlated with the training labels. The second one only kept those dimensions with a standard deviation higher than the mean standard deviation of all dimensions. Both techniques reduced the size of the feature space by two approximately and are referred as to *mostCorr* and *mostStd*, respectively.

Classification: Multi–class SVM classifiers with RBF kernel were used in both subtasks, particularly, 2–class SVMs for the MDR task and 5–class SVMs for the TBT task. Grid search over the RBF parameters (cost C and gamma γ) was applied. Since the data were normalized, both C and γ varied in $\{2^{-10}, 2^{-9}, \ldots, 2^{10}\}$. The best C and γ combination for a run was set as the one with highest cross–validation (CV) accuracy (10–fold) in the training set of each subtask.

4 Experiments

The ImageCLEF 2017 TB task was divided into two phases. In the first phase, the organizers released for each subtask a set of patient CT volumes as training

set with their lung masks and ground truth labels. In the second phase, the test set with the corresponding lung segmentations were provided. However, the test labels were never released. The evaluation of the methods was performed by the organizers of the task based on the predicted labels submitted by the participants. In this section we detail the tested and submitted runs. Moreover the results of other participants provided by the organizers of the task are detailed.

4.1 Tested Runs

Considering all the different configurations explained in Section 3, 648 runs were obtained per subtask. Table 2 summarizes all possible options for each configuration step using the same codename as in the result tables.

Table 2. Possible configurations for each step. With these variations there were 648 combinations: 3 graph architectures × 4 edge weights × 3 texture features × 3 feature measures × 2 feature normalizations × 3 feature reductions.

Graph model property	**Options**
Graph architecture	Graph_Complete, Graph_66, Graph_84
Edge weight	corr, cos, euc, sumNorm
Texture feature	FHOG, 3DARiesz, FHOG and 3DARiesz
Feature measure	mean, std, mean and std
Feature normalization	[0,1], GaussNorm
Feature reduction	none, mostCorr, mostStd

4.2 Submitted Runs

A total of ten runs could be submitted in the ImageCLEF 2017 TB task, considering the submitted runs of both subtasks. Therefore, five runs were submitted for each subtask. For both subtasks, we first selected the five runs with best scores considering only the CV accuracy on the training set (Acc_{trn}). Tables 3 and 4 show the identifier and run setup of the five selected runs with top Acc_{trn} for each subtask, respectively. Then, subgroups of these five runs were combined using late fusion to obtain new run files. Four new run files were obtained per subtask, identified by the suffixes *TopBest2*, *TopBest3*, *TopBest4*, and *TopBest5*. The late fusion was computed using the probabilities that the SVM classifier returned and the mean probability of belonging to each class. Finally, we submitted three original runs with the best scores and two fused runs per subtask.

The following tables show the results obtained set by the submitted runs on the training (Acc_{trn}) and the final performance in the competition (Acc_{tst}). The final ranking was based on the AUC for the MDR subtask and on the unweighted Cohen's Kappa coefficient (Kappa) for the TBT task. Table 5 shows the results for the MDR subtask provided by the task organizers. The run identifiers *MDR_TopBest3* and *MDR_TopBest5* were obtained by late fusion of the 3

Table 3. Runs for the MDR subtask with the best scores based on the CV accuracy in the training set.

Run Id.	Graph	Texture features	F. measure	E. weight	F. norm.	F. reduct.	Acc_{trn}
MDR_Top1	Graph_84	FHOG and 3DARiesz	mean and std	corr	GaussNorm	mostCorr	0.6900
MDR_Top2	Graph_66	FHOG and 3DARiesz	std	cos	[0,1]	mostCorr	0.6856
MDR_Top3	Graph_84	FHOG	mean	corr	[0,1]	none	0.6812
MDR_Top4	Graph_66	FHOG and 3DARiesz	mean and std	corr	[0,1]	mostCorr	0.6725
MDR_Top5	Graph_66	FHOG	mean	corr	GaussNorm	mostCorr	0.6725

Table 4. Runs for the TBT subtask with the best scores based on the CV accuracy in the training set.

Run Id.	Graph	Texture features	F. measure	E. weight	F. norm.	F. reduct.	Acc_{trn}
TBT_Top1	Graph_66	FHOG and 3DARiesz	mean and std	sumNorm	GaussNorm	none	0.5276
TBT_Top2	Graph_84	FHOG and 3DARiesz	mean and std	sumNorm	GaussNorm	none	0.5174
TBT_Top3	Graph_66	FHOG and 3DARiesz	mean and std	sumNorm	[0,1]	none	0.5112
TBT_Top4	Graph_66	FHOG and 3DARiesz	mean and std	sumNorm	GaussNorm	mostCorr	0.5112
TBT_Top5	Graph_84	FHOG and 3DARiesz	mean and std	sumNorm	[0,1]	none	0.5092

and 5 best runs respectively. The results for the TBT task are shown in Table 6. Again, the run identifiers *TBT_TopBest3* and *TBT_TopBest5* correspond to the late fusion of the 3 and 5 best runs respectively.

Table 5. Results of the MDR detection task. We participated as the MedGIFT group.

Group Name	Run Id	AUC	Acc_{tst}	Acc_{trn}	#Rank
MedGIFT	MDR_Top1	0.5825	0.5164	0.6900	1
MedGIFT	MDR_TopBest3	0.5727	0.4648	–	2
MedGIFT	MDR_TopBest5	0.5624	0.4836	–	3
SGEast	MDR_LSTM_6_probs	0.5620	0.5493	–	4
SGEast	MDR_resnet_full	0.5591	0.5493	–	5
MedGIFT	MDR_Top2	0.5337	0.4883	0.6856	10
MedGIFT	MDR_Top3	0.5112	0.4413	0.6725	17

5 Discussion and Conclusions

This article presents a novel graph–based framework to model the lung fields based on regional 3D texture features. The parts of this framework can be adapted to describe multiple diseases affecting the lung parenchyma. In particular, more than 600 configurations were tested to describe patients with TB. The participation in the ImageCLEF 2017 TB task provides an objective comparison between methods, since the ground truth for the test set was never released. The global description of the lungs provided by the graph model allowed the detection of MDR patients better than any other approach submitted in this challenge. Moreover, it also showed to be useful in the distinction of the different TB types. According to the results in the ImageCLEF 2017 TB task,

Table 6. Results of the TBT classification task. We participated as the MedGIFT group.

Group Name	Run Id	Kappa	Acc_{tst}	Acc_{trn}	#Rank
SGEast	TBT_resnet_full	0.2438	0.4033	–	1
SGEast	TBT_LSTM_17_wcrop	0.2374	0.3900	–	2
MEDGIFT UPB	TBT_T_GNet	0.2329	0.3867	–	3
SGEast	TBT_LSTM_13_wcrop	0.2291	0.3833	–	4
Image Processing	TBT-testSet-label-Apr26-XGao-1	0.2187	0.4067	–	5
MedGIFT	TBT_Top1	0.1623	0.3600	0.5276	10
MedGIFT	TBT_TopBest3	0.1548	0.3500	–	12
MedGIFT	TBT_TopBest5	0.1410	0.3367	–	15
MedGIFT	TBT_Top4	0.1352	0.3300	0.5112	16
MedGIFT	TBT_Top2	0.1235	0.3200	0.5174	17

the new representation of the lungs as a graph entity showed to be promising, reaching better results than for example deep learning approaches. Our method was robust enough to provide a better characterization of the several classes in both subtasks only with the available number of patients in the task. If added to the clinical workflow, physicians can benefit of a new way of visualizing and interpreting the lung parenchyma, in a systematic and schematic fashion.

For the MDR subtask, the graph model participated with five runs and obtained the 1st, 2nd and 3rd place in the challenge. The results obtained by the participants confirmed the difficulty of this subtask. Independently of the technique applied, all runs remained close to the performance of a random classifier, meaning that there is likely a high potential for improvements.

On the other hand, the results support the suitability of the imaging techniques for the TBT task. Five runs were also submitted to the TBT subtask but the best rank obtained by the texture–based graph model was 10. For this particular task, deep learning methods worked better than other approaches, obtaining the 6 best results. The results underline the difficulty of both tasks and the suitability of the graph model for describing TB patients. However, the strong differences in the accuracies obtained for the training and test sets (see Tables 5 and 6) suggest some overfitting in the training phase. The graph model describes each patient with a single vector in a relativity large feature space. Therefore, more training data may be needed to build a stable model of each class.

Acknowledgements

This work was partly supported by the Swiss National Science Foundation in the project PH4D (320030–146804).

References

1. Bento, J., Silva, A.S., Rodrigues, F., Duarte, R.: Diagnostic tools in tuberculosis. Acta medica portuguesa **24**(1) (2011) 145–154

2. Jeong, Y.J., Lee, K.S.: Pulmonary tuberculosis: up-to-date imaging and management. American Journal of Roentgenology **191**(3) (2008) 834–844
3. Richiardi, J., Achard, S., Bunke, H., Van De Ville, D.: Machine learning with brain graphs: Predictive modeling approaches for functional imaging in systems neuroscience. IEEE Signal Processing Magazine **30**(3) (2013) 58–70
4. Richiardi, J., Eryilmaz, H., Schwartz, S., Vuilleumier, P., Van De Ville, D.: Decoding brain states from fMRI connectivity graphs. NeuroImage **56**(2) (2011) 616–626
5. Dicente Cid, Y., Müller, H., Platon, A., Janssens, J.P., Lador, F., Poletti, P.A., Depeursinge, A.: A lung graph–model for pulmonary hypertension and pulmonary embolism detection on DECT images. In: MICCAI Workshop on Medical Computer Vision: Algorithms for Big Data, MCV 2016. (2016)
6. Dicente Cid, Y., Kalinovsky, A., Liauchuk, V., Kovalev, V., , Müller, H.: Overview of ImageCLEFtuberculosis 2017 - predicting tuberculosis type and drug resistances. In: CLEF 2017 Labs Working Notes. CEUR Workshop Proceedings, Dublin, Ireland, CEUR-WS.org <http://ceur-ws.org> (September 11-14 2017)
7. Müller, H., Clough, P., Deselaers, T., Caputo, B., eds.: ImageCLEF – Experimental Evaluation in Visual Information Retrieval. Volume 32 of The Springer International Series On Information Retrieval. Springer, Berlin Heidelberg (2010)
8. Kalpathy-Cramer, J., García Seco de Herrera, A., Demner-Fushman, D., Antani, S., Bedrick, S., Müller, H.: Evaluating performance of biomedical image retrieval systems: Overview of the medical image retrieval task at ImageCLEF 2004–2014. Computerized Medical Imaging and Graphics **39**(0) (2015) 55 – 61
9. Villegas, M., Müller, H., Gilbert, A., Piras, L., Wang, J., Mikolajczyk, K., García Seco de Herrera, A., Bromuri, S., Amin, M.A., Kazi Mohammed, M., Acar, B., Uskudarli, S., Marvasti, N.B., Aldana, J.F., Roldán García, M.d.M.: General overview of ImageCLEF at the CLEF 2015 labs. In: Working Notes of CLEF 2015. Lecture Notes in Computer Science. Springer International Publishing (2015)
10. Ionescu, B., Müller, H., Villegas, M., Arenas, H., Boato, G., Dang-Nguyen, D.T., Dicente Cid, Y., Eickhoff, C., Garcia Seco de Herrera, A., Gurrin, C., Islam, B., Kovalev, V., Liauchuk, V., Mothe, J., Piras, L., Riegler, M., Schwall, I.: Overview of ImageCLEF 2017: Information extraction from images. In: Experimental IR Meets Multilinguality, Multimodality, and Interaction 8th International Conference of the CLEF Association, CLEF 2017. Volume 10456 of Lecture Notes in Computer Science., Dublin, Ireland, Springer (September 11-14 2017)
11. Dicente Cid, Y., Jimenez-del-Toro, O., Depeursinge, A., Müller, H.: Efficient and fully automatic segmentation of the lungs in CT volumes. In Orcun Goksel, Jimenez-del-Toro, O., Foncubierta-Rodriguez, A., Müller, H., eds.: Proceedings of the VISCERAL Challenge at ISBI. Number 1390 in CEUR Workshop Proceedings (Apr 2015) 31–35
12. Depeursinge, A., Zrimec, T., Busayarat, S., Müller, H.: 3D lung image retrieval using localized features. In: Medical Imaging 2011: Computer–Aided Diagnosis. Volume 7963., SPIE (2011) 79632E
13. Zrimec, T., Busayarat, S., Wilson, P.: A 3D model of the human lung. In LNCS, S., ed.: Proceedings of MICCAI 2004. Volume 3217. (October 2004) 1074–1075
14. Liu, K., Skibbe, H., Schmidt, T., Blein, T., Palme, K., Brox, T., Ronneberger, O.: Rotation-invariant hog descriptors using fourier analysis in polar and spherical coordinates. International Journal of Computer Vision **106**(3) (2014) 342–364
15. Dicente Cid, Y., Müller, H., Platon, A., Poletti, P.A., Depeursinge, A.: 3–D solid texture classification using locally–oriented wavelet transforms. IEEE Transactions on Image Processing **26**(4) (April 2017) 1899–1910