Image Dissimilarity-Based Quantification of Lung Disease from CT

Lauge Sørensen[1], Marco Loog[1,2], Pechin Lo[1], Haseem Ashraf[3], Asger Dirksen[3], Robert P.W. Duin[2], and Marleen de Bruijne[1,4]

[1] The Image Group, Department of Computer Science,
University of Copenhagen, Denmark
`lauges@diku.dk`
[2] Pattern Recognition Laboratory, Delft University of Technology, The Netherlands
[3] Department of Respiratory Medicine, Gentofte University Hospital, Denmark
[4] Biomedical Imaging Group Rotterdam, Departments of Radiology & Medical
Informatics, Erasmus MC, The Netherlands

Abstract. In this paper, we propose to classify medical images using dissimilarities computed between collections of regions of interest. The images are mapped into a dissimilarity space using an image dissimilarity measure, and a standard vector space-based classifier is applied in this space. The classification output of this approach can be used in computer aided-diagnosis problems where the goal is to detect the presence of abnormal regions or to quantify the extent or severity of abnormalities in these regions. The proposed approach is applied to quantify chronic obstructive pulmonary disease in computed tomography (CT) images, achieving an area under the receiver operating characteristic curve of 0.817. This is significantly better compared to combining individual region classifications into an overall image classification, and compared to common computerized quantitative measures in pulmonary CT.

1 Introduction

Quantification of abnormality in medical images often involves classification of regions of interest (ROIs), and combination of individual ROI classification outputs into one global measure of disease for the entire image [1,2,3,4,5,6,7]. These measures may, e.g., express a probability of the presence of certain abnormalities or reflect the extent or severity of disease.

A global image measure based on the fusion of several independent ROI classifications disregards the fact that the ROIs belong to a certain image in the classification step. Moreover, in some cases only global image labels are available, while the images are still represented by ROIs in order to capture localized abnormalities. In some studies, this is handled by propagating the image label to the ROIs within that image, which again allows fusion of individual ROI classifications, to obtain a global image measure [4,5,6]. However, an image showing abnormality will generally comprise both healthy and abnormal regions, and the above approach, incorrectly, labels ROIs without abnormality in such an image as abnormal.

In this paper, we propose to classify medical images using dissimilarities computed directly between the images, where the images are represented by a collection of regions. In this approach, all ROIs in an image contribute when that image is compared to other images, thereby taking into account that the ROIs collectively constitute that image. Further, problems where only a global image label is available are handled automatically since the classification is done at the image level. The images are mapped into a dissimilarity space [8] in which a standard vector space-based classifier can be directly applied, and the soft output of this classifier is used as quantitative measure of disease. The measure used to compute the dissimilarity between two images is the crucial component in this approach, and we evaluate four different image dissimilarity measures in the experiments.

The proposed approach is applied to quantify chronic obstructive pulmonary disease (COPD) in volumetric pulmonary computed tomography (CT) images using texture. Several general purpose classifiers built in the obtained image dissimilarity spaces are evaluated and compared to image classification by fusion of individual ROI classifications as was used in [6].

2   Image Dissimilarity Space

We propose to represent a set of images $\{I_1, \ldots, I_n\}$ by their pair-wise dissimilarities $d(I_i, I_j)$ and build classifiers on the obtained dissimilarity representation [8]. From the matrix of pair-wise image dissimilarities $D = [d(I_i, I_j)]_{n \times n}$ computed from the set of images, there exist different ways of arriving at a feature vector space where traditional vector space methods can be applied. In this work, we consider the dissimilarity space approach [8]. An image dissimilarity space is constructed of dimension equal to the size of the training set $|T| = |\{J_1, \ldots, J_m\}| = m$, where each dimension corresponds to the dissimilarity to a certain training set image J. All images I are represented as single points in this space, and are positioned according to their dissimilarities to the training set images $D(I, T) = [d(I, J_1), \ldots, d(I, J_m)]$. The image dissimilarity measure is a function from two images, represented as sets of ROIs, to a non-negative scalar $d(\cdot, \cdot) : \mathcal{P}(S) \times \mathcal{P}(S) \rightarrow \mathbb{R}_+$ where S is the set of all possible ROIs and $\mathcal{P}(S)$ is the power set of S. It is in this part of the proposed approach that the ROIs are taken collectively into account.

3   Image Dissimilarity Measures

The main issue in obtaining the image dissimilarity space, is the definition of $d(\cdot, \cdot)$. Since the application in this paper is quantification of COPD in pulmonary CT images based on textural appearance in the ROIs, we will focus on image dissimilarity measures suitable for this purpose. In texture-based classification of lung tissue, the texture is sometimes assumed stationary [3,4,6,7]. We will make the same assumption and, therefore, disregard the spatial location of the

ROIs within the lungs. The following are then desirable properties of an image dissimilarity measure for quantification of COPD:

1. Spatial location within the image does not matter. ROIs should be compared solely based on the textural appearance within those regions.
2. The amount of diseased tissue does matter. An image with many abnormal regions is more diseased than an image with few abnormal regions.
3. The appearance of abnormal tissue does matter. Two images with abnormal regions of the same size but with different types of abnormalities should be considered different.

A simple and straightforward image dissimilarity measure between two images, I_1 and I_2, having the above properties is the sum of all pair-wise ROI dissimilarities:

$$d_{sum}(I_1, I_2) = \sum_{i,j} \Delta(\mathbf{x}_{1i}, \mathbf{x}_{2j}) \tag{1}$$

where $\mathbf{x}_{1i}$ is the i'th ROI in I_1 and $\Delta(\cdot, \cdot)$ is a texture appearance dissimilarity measure between two ROIs. However, when all ROIs in one image are compared to all ROIs in the other image, the discriminative information of abnormality present in only a few ROIs may be lost. One way to avoid this is to match every ROI in one image with the most similar ROI in the other image. This is the minimum sum distance [9]:

$$d_{ms}(I_1, I_2) = \sum_i \min_j \Delta(\mathbf{x}_{1i}, \mathbf{x}_{2j}) + \sum_j \min_i \Delta(\mathbf{x}_{2j}, \mathbf{x}_{1i}). \tag{2}$$

However, this image dissimilarity measure allows several ROIs in one image to be matched with the same ROI in the other image. This may not be desirable for quantifying COPD since an image with a small abnormal area is considered similar to an image with a large abnormal area. The image dissimilarity measure proposed in the following is a trade-off between d_{sum} and d_{ms}; it is the sum of several pair-wise ROI dissimilarities, where only one-to-one matchings are allowed, thereby considering images with a small abnormal area as dissimilar to images with a large abnormal area.

3.1 Bipartite Graph Matching-Based Image Dissimilarity Measure

The dissimilarity between two images, or sets of ROIs, $I_1 = \{\mathbf{x}_{1i}\}_n$ and $I_2 = \{\mathbf{x}_{2i}\}_n$, can be expressed as the minimum linear sum assignment between the two sets according to $\Delta(\cdot, \cdot)$. This can be seen as assigning the ROIs in one set to the ROIs in the other set in a way such that the two sets are as similar as possible while only allowing one-to-one matchings. Let $G = (I_1 \cup I_2, E)$ be a weighted undirected bipartite graph with node sets I_1 and I_2 where $|I_1| = |I_2| = n$, edge set $E = \{\{\mathbf{x}_{1i}, \mathbf{x}_{2j}\} : i, j = 1, \ldots, n\}$, and with weight $\Delta(\mathbf{x}_{1i}, \mathbf{x}_{2j})$ associated with each edge $\{\mathbf{x}_{1i}, \mathbf{x}_{2j}\} \in E$. The resulting graph is illustrated in Figure 1. A subset M of E is called a perfect matching, or assignment, if every node of
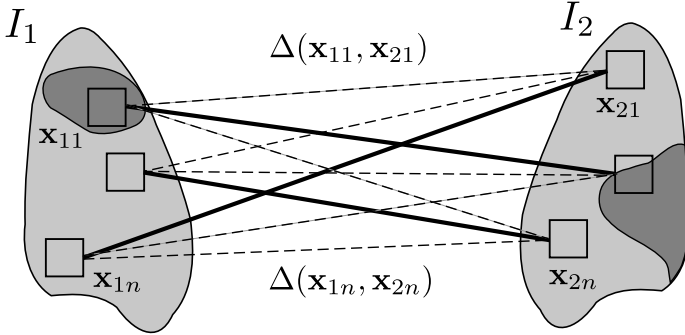
Fig. 1. Illustration of the graph considered when computing the dissimilarity between two images, I_1 and I_2, in (4). All edges have an associated weight $\Delta(\mathbf{x}_{1i}, \mathbf{x}_{2j})$ that expresses the textural dissimilarity between the two corresponding ROIs $\mathbf{x}_{1i}$ and $\mathbf{x}_{2j}$. The edges in the perfect matching with minimum weight M^* are shown as solid lines, and the remaining edges, not in M^*, are shown as dashed lines.

G is incident with exactly one edge in M. The perfect matching with minimum weight M^* is given by

$$M^* = \underset{M}{\operatorname{argmin}} \sum_{\{\mathbf{x}_{1i}, \mathbf{x}_{2j}\} \in M} \Delta(\mathbf{x}_{1i}, \mathbf{x}_{2j}) : M \text{ is a perfect matching.} \qquad (3)$$

This problem can be solved efficiently using the Hungarian algorithm [10]. The resulting image dissimilarity measure is thus

$$d_{la}(I_1, I_2) = \sum_{\{\mathbf{x}_{1i}, \mathbf{x}_{2j}\} \in M^*} \Delta(\mathbf{x}_{1i}, \mathbf{x}_{2j}) \qquad (4)$$

where M^* is obtained via (3). No normalization is needed since the images contain an equal amount of ROIs, i.e., n ROIs. Although not used in this work, the formulation can also be relaxed to handle images containing a varying number of ROIs. This will result in an image dissimilarity measure that does not obey the triangle inequality due to partial matches of images. However, this is no problem in the dissimilarity space appraoch.

4   Experiments

4.1   Data

The data consists of 296 low-dose volumetric CT images from the Danish Lung Cancer Screening Trial with the following scan parameters: tube voltage 120 kV, exposure 40 mAs, slice thickness 1 mm, and in-plane resolution ranging from 0.72 to 0.78 mm. 144 images are from subjects diagnosed as healthy and 152 images are from subjects diagnosed with moderate to very severe COPD. Both groups are diagnosed according to spirometry [11].

4.2   Evaluation

The image dissimilarity-based approach is applied by building classifiers in the CT image dissimilarity spaces obtained using $d(\cdot, \cdot)$. This is compared to using $d(\cdot, \cdot)$ directly as distance in a k nearest neighbor classifier (kNN), which for $k = 1$ corresponds to template matching, and to fusing individual ROI classifications, classified using kNN, for image classification [6]. A posterior probability of each image being positive is obtained using leave-one-out estimation, and receiver operating characteristic (ROC) analysis is used to evaluate the different methods by means of the area under the ROC curve (AUC). The CT image dissimilarity spaces considered in each leave-out trial are of dimension equal to the size of the training set, i.e., 295-dimensional.

Apart from the three image dissimilarity measures described in Section 3, (1), (2), and (4), we also experiment with the Hausdorff distance [9], d_h. This is a classical point set distance measure that do not obey the second property described in Section 3, since it ultimately rely on the dissimilarity between two single ROIs, or points, one from each image. Thus, a total of four different CT image dissimilarity representations are considered in the experiments, one based on each of the four image dissimilarity measures d_{sum}, d_{ms}, d_{la}, and d_h.

4.3   Classifiers

All CT images are represented by a set of 50 ROIs of size $41 \times 41 \times 41$ voxels that each are described by three filter response histograms capturing the local image texture. The filters are: Laplacian of Gaussian (LG) at scale 0.6 mm, gradient magnitude (GM) at scale 4.8 mm, and Gaussian curvature (GC) at scale 4.8 mm. The ROI size as well as the filters are selected based on the results in [6]. The ROI dissimilarity measure used in all experiments is based on the L1-norm between the filter response histograms: $\Delta(\mathbf{x}_1, \mathbf{x}_2) = L_1(h_{LG}(\mathbf{x}_1), h_{LG}(\mathbf{x}_2)) + L_1(h_{GM}(\mathbf{x}_1), h_{GM}(\mathbf{x}_2)) + L_1(h_{GC}(\mathbf{x}_1), h_{GC}(\mathbf{x}_2)))$ where $h_i(\mathbf{x})$ is the response histogram of filter i computed in ROI $\mathbf{x}$.

A SVM with a linear kernel and trade-off parameter $C = 1$ is applied in the obtained CT image dissimilarity spaces. kNN is applied in the following three ways: in the image dissimilarity spaces, using the image dissimilarities directly as distance, and using ROI dissimilarity directly for ROI classification followed by fusion. $k = 1$ is used as well as $k = \sqrt{n}$ where n is the number of proto-types [12]. When classifying CT images, this is $k = \lfloor\sqrt{295}\rfloor = 17$, and when classifying ROIs, this is $k = \lfloor\sqrt{(295 \times 50)}\rfloor = 121$. The following combination rules are considered for fusing individual ROI classifications into image classifications: quantile-based fusion schemes with quantiles ranging from 0.01, i.e., the minimum rule, to 1.00, i.e., the maximum rule, and the mean rule [13]. We also compare to two common CT-based quantitative measures, namely, relative area of emphysema (RA) and percentile density (PD) using the common thresholds of -950 Hounsfield units (HU) and 15% respectively [14]. These measures are computed from the entire lung fields and are denote RA_{950} and PD_{15}, respectively.

4.4   Results

Table 1 shows the estimated AUCs for all the classifiers. The best CT image-dissimilarity based classifier, SVM built in CT image dissimilarity space using d_{la}, achieves an AUC of 0.817. This is better than the best performing mean rule ROI fusion-based classifier, 121NN, which achieves an AUC of 0.751. The common CT-based measures, RA_{950} and PD_{15}, perform worse than all the texture-based measures. The quantile-rule only performed better than the mean rule in the ROI classification fusion in one case, 121NN using maximum rule achieved an AUC of 0.757, and they are therefore not reported in Table 1. SVM in image dissimilarity space using d_{la} or d_{sum} is significantly better, with $p = 0.0028$ and $p = 0.0270$, respectively, than 121NN using the mean rule, while SVM using d_{ms} is not, with $p = 0.085$, according to DeLong, DeLong, and Clarke-Pearson's test [15].

Table 1. AUCs for COPD diagnosis. **Left:** The results of classification in image dissimilarity space, as well as using the image dissimilarities directly in kNN. **Right:** The results of fusion of individual ROI classification outputs for image classification using the mean rule. The best performing classifier in each approach is marked in bold-face.

	in image dissimilarity space			using $d(\cdot,\cdot)$ directly		fusion of ROI classifications	
	SVM	1NN	17NN	1NN	17NN	1NN	0.721
d_h	0.609	0.522	0.624	0.566	0.668	121NN	**0.751**
$d_{sum}(1)$	0.793	0.619	0.643	0.504	0.663	RA_{950}	0.585
d_{ms} (2)	0.795	0.632	0.725	0.600	0.768	PD_{15}	0.589
d_{la} (4)	**0.817**	0.612	0.671	0.593	0.741		

5   Discussion

Image dissimilarity measures that match each ROI of one image to an ROI of the other image, under some restrictions, are expected to work well for quantification of COPD within the proposed framework, mainly because more information is taken into account, but also due to increased robustness to noisy ROIs. This is in contrast to measures relying on the match between two ROIs only, such as the Hausdorff distance that is included in the experiments for the sake of completeness. Further, the main argument for building a more global decision rule, such as SVM in a dissimilarity space, instead of applying kNN using the dissimilarity directly as distance is better utilization of the training data and therefore reduced sensitivity to noisy prototypes [8]. This may explain why SVM with a linear kernel built in the dissimilarity space obtained using d_{la} is the best performing of the CT image dissimilarity-based approaches. However, validation on an unseen data set would be needed to draw a final conclusion on this. The experiments showed that SVM with a linear kernel built in the CT image dissimilarity space obtained using d_{la} performed significantly better than using kNN for ROI classification together with the mean rule for CT image classification

($p < 0.05$). This implies that performing the classification at image level, taking into account that an image is in fact a collection of ROIs that collectively constitute that image, is beneficial compared to classifying ROIs individually, while disregarding the fact that they do belong to a certain image.

The computational complexity of the proposed approach using either of the image dissimilarities (1), (2), or (4), in terms of the number of times $\Delta(\cdot, \cdot)$ is evaluated in order to classify a CT image, is the same compared to using the image dissimilarities directly as distance in kNN and to fusion of ROI classifications that are classified using kNN. All approaches require a total of $50 \times 50 \times 295$ evaluations of $\Delta(\cdot, \cdot)$ for classification of a CT image.

When an image is represented by a collection of ROIs and only a label for the entire image is available, the problem of classifying the image can be formulated as a multiple instance learning (MIL) problem [16]. Fusion of independent ROI classifications in order to arrive at an overall image classification can be seen as a "simple" algorithm for solving such a problem. In this paper, we propose to use the dissimilarity-based approach of Pekalska *et al.* [8] on image dissimilarities for solving MIL problems in medical imaging. The approach is similar in spirit to various kernel-based MIL algorithms, such as [17]. The dissimilarity-based approach, however, puts less restrictions on the proximity measure used for comparing objects. Kernel-based approaches require the kernel to be positive definite, which excludes well-known proximity measures such as the Hausdorff distance [9] as well as the bipartite graph matching image dissimilarity measure proposed in this work. Within our framework such measures can be used without any problem.

In conclusion, dissimilarities computed directly between medial images, where the images are represented by a collection of ROIs, was proposed for image classification. This is an alternative to fusion of individual ROI classifications within the images. A SVM built in a dissimilarity space using an image dissimilarity measure based on a minimum sum perfect matching in a weighted bipartite graph, with ROIs as nodes and the textural dissimilarity between two ROIs as edge weight, achieved an AUC of 0.817 on a COPD quantification problem in volumetric pulmonary CT.

References

1. Müller, N.L., Staples, C.A., Miller, R.R., Abboud, R.T.: "Density mask". An objective method to quantitate emphysema using computed tomography. Chest 94(4), 782–787 (1988)
2. van Ginneken, B., Katsuragawa, S., ter Haar Romeny, B., Doi, K., Viergever, M.: Automatic detection of abnormalities in chest radiographs using local texture analysis. IEEE Trans. Med. Imag. 21(2), 139–149 (2002)

3. Park, Y.S., Seo, J.B., Kim, N., Chae, E.J., Oh, Y.M., Lee, S.D., Lee, Y., Kang, S.H.: Texture-based quantification of pulmonary emphysema on high-resolution computed tomography: comparison with density-based quantification and correlation with pulmonary function test. Invest Radiol. 43(6), 395–402 (2008)

4. Raundahl, J., Loog, M., Pettersen, P., Tanko, L.B., Nielsen, M.: Automated effect-specific mammographic pattern measures. IEEE Trans. Med. Imag. 27(8), 1054–1060 (2008)

5. Arzhaeva, Y., Hogeweg, L., de Jong, P.A., Viergever, M.A., van Ginneken, B.: Global and local multi-valued dissimilarity-based classification: Application to computer-aided detection of tuberculosis. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009, Part II. LNCS, vol. 5762, pp. 724–731. Springer, Heidelberg (2009)

6. Sørensen, L., Lo, P., Ashraf, H., Sporring, J., Nielsen, M., de Bruijne, M.: Learning COPD sensitive filters in pulmonary CT. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009, Part I. LNCS, vol. 5761, pp. 699–706. Springer, Heidelberg (2009)

7. Sørensen, L., Shaker, S.B., de Bruijne, M.: Quantitative analysis of pulmonary emphysema using local binary patterns. IEEE Trans. Med. Imag. 29(2), 559–569 (2010)

8. Pekalska, E., Duin, R.P.W.: Dissimilarity representations allow for building good classifiers. Pattern Recog. Lett. 23(8), 943–956 (2002)

9. Eiter, T., Mannila, H.: Distance measures for point sets and their computation. Acta Inf. 34(2), 109–133 (1997)

10. Kuhn, H.W.: The hungarian method for the assignment problem. Naval Research Logistic Quarterly 2, 83–97 (1955)

11. Rabe, K.F., Hurd, S., Anzueto, A., Barnes, P.J., Buist, S.A., Calverley, P., Fukuchi, Y., Jenkins, C., Rodriguez-Roisin, R., van Weel, C., Zielinski, J.: Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. Am. J. Respir. Crit. Care Med. 176(6), 532–555 (2007)

12. Kittler, J., Alkoot, F.M.: Moderating k-NN classifiers. Pattern Anal. Appl. 5(3), 326–332 (2002)

13. Loog, M., Van Ginneken, B.: Static posterior probability fusion for signal detection: applications in the detection of interstitial diseases in chest radiographs. In: ICPR (1), pp. 644–647. IEEE Computer Society, Los Alamitos (2004)

14. Webb, W.R., Müller, N., Naidich, D.: High-Resolution CT of the Lung, 3rd edn. Lippincott Williams & Wilkins (2001)

15. DeLong, E.R., DeLong, D.M., Clarke-Pearson, D.L.: Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44(3), 837–845 (1988)

16. Dietterich, T.G., Lathrop, R.H., Lozano-Pérez, T.: Solving the multiple instance problem with axis-parallel rectangles. Artif. Intell. 89(1-2), 31–71 (1997)

17. Gärtner, T., Flach, P.A., Kowalczyk, A., Smola, A.J.: Multi-instance kernels. In: ICML, pp. 179–186. Morgan Kaufmann, San Francisco (2002)