Automatic abnormal region detection in lung CT images for visual retrieval

Henning Müller (1), Samuel Marquis (1), Gilles Cohen (1), Pierre-Alexandre Poletti (2), Christian Lovis (1), Antoine Geissbuhler (1)

(1) Service of medical informatics, University and hospitals of Geneva
(2) Emergency radiology, University hospitals of Geneva

Abstract
Image management, analysis, and retrieval are currently very active research fields mainly because of the large amount of visual data being produced in modern hospitals, and the lack of applications dealing with these data. Most often, the goal is to aid the diagnostic process. Unfortunately, only very few medical image retrieval systems are currently used in clinical routine. One application domain with a high potential for automatic image retrieval is the analysis and retrieval of lung CTs. A first user study in the United States (Purdue University) shows that these systems allow improving the diagnostic quality significantly.
This article describes an approach to an aid for lung CT diagnostics. The analysis incorporates several steps and the goal is to automate the process as much as possible for easy integration into clinical processes. Thus, several automatic steps are proposed from a selection of the most characteristic slices, to an automatic segmentation of the lung tissue and a classification on the segmented area into visual observation classes. Feedback to the MD is given in the form of marked regions in the images that appear to be different from the norm of healthy tissue. We currently work on a small set of training images with marked and annotated regions but a larger set of images for the evaluation of our algorithm is in work. The article currently only contains a short quantitative evaluation.
For most tasks we use existing open source software such as Weka, GIFT, and itk. This allows an easy reproduction of the search results and limits the need for costly redevelopments.

Keywords
Content-based image retrieval, high-resolution lung CT, diagnostic aid, classification

1. Introduction
Content-based image retrieval (CBIR) has been an extremely active domain in the fields of computer vision and images processing for more than 20 years [1]. In the medical field, this domain is also starting to become active, as an increasing amount of visual data is being produced in hospitals and made available in digital form [2,3]. General medical image retrieval in PACS-like databases is in this context very different from specialized retrieval in a very focused domain. In the medical field, the main goal is the use as a diagnostic aid and for access to medical teaching files. Current medical use is on the retrieval of tumors by shapes [4] as well as on histological images [5], and in other specific fields (dermatology, pathology). A domain where textures play a very important role in the diagnostic process is the analysis of high-resolution lung CTs [6]. In [7], a user test among radiologists shows that an image retrieval system can improve the diagnostic quality significantly, especially for less experienced radiologists. Still, most of these systems either rely on much complicated interaction with the user, which makes them hard to introduce into a clinical context, or they are too broad to be used as a diagnostic aid in a specialized domain.
This article details a solution for helping with the interpretation of high-resolution lung CTs, which is a domain where diagnostics are fairly hard especially for non-chest specialists. A large number of unrelated diseases exist with often unspecific symptoms as well as many classes of visual observations and other data to be integrated (lab results, age, environmental exposure, …). The diagnostic result strongly depends on the overall texture of the lung tissue, so automatic analysis seems possible. Our project limits the direct interaction with the user and performs as many tasks as possible in an automatic fashion, so a minimum of time is needed to operate the system and get responses for feedback. Several steps of the process have already been integrated, whereas a few are still in their first tests.

2. Diagnostic aid on lung CT interpretation

This section describes the steps that are necessary for a complete diagnostic aid system for lung diagnostics and their degree of automation.

2.1 Generation of a test database and acquisition of representative samples

The first and most important part is the creation of a database of thin-section lung CTs. This database needs to include healthy cases as well as pathologic cases. Characteristic regions need to be marked by a radiologist to allow learning the characteristics of a disease (or visual observation) with respect to healthy tissue. Currently, we only have a fairly small database of 12 CT series containing a total of 326 images. 112 regions of varying size are marked in 69 of the images by a radiologist to represent the following classes of visual observations for the classification step:
- Healthy tissue (52 sample regions);
- Emphysema (21);
- Micro nodules (19);
- Macro nodules (3);
- Interstitial syndrome (5);
- Ground glass attenuation (1);
- Fibrosis (6).

It is important to note that prototypically healthy regions have to be annotated by the radiologist as well so that a classifier can get a good idea of healthy tissue. Other systems in the literature often use only pathologic classes for the classification but the first step in the diagnostic process is to find out whether the tissue is abnormal or not. We are currently creating a larger database projected to contain at least 100-200 series of 50-60 images that will allow us a better representation of the classes. In Figure 1 a screenshot of our tool for image annotation can be seen. It generates a simple XML file containing regions of interests as a set of points (outline) and a label for each outline. The files are then fed into the system along with the images at the training step.

Figure 1: A screenshot of our utility for the annotation of image regions.

2.2 Analysis of blocks of lung tissue

In Figure 3 the partitioning of the lungs into smaller blocks (size 16x16 pixels in the image) for further detailed texture analysis can be seen. A block is taken into account if three edges are inside of the area marked by the expert or automatically segmented by the system. These lung blocks are stored as references together with the original image. This avoids artifacts of the filters that can occur due to missing border pixels as we can take into account the entire block environment.

Figure 2: Partitioning of lung tissue into small blocks for feature extraction and classification.

The framework is designed to facilitate finding the optimal block size for analysis and classification. From each block we extract and store the following visual features:
·   average grey level, standard deviation, skewness, kurtosis and min-max of the grey levels;
·   grey level histogram using 32 grey levels;
·   features derived from co-occurrence matrices (four directions, two distances);
·   responses of Gabor filters in four directions and at three scales;
·   Run length after thresholding and other features based on mathematical morphology.

A small number of grey levels in the histogram is sufficient for this kind of classification as has been shown in image retrieval applications.

2.3 Training

The features together with the region label become a sample in a classification problem [8]. Based on the acquired training data, the weights of the features for classification are calculated. To develop an optimal classification strategy, several classifiers are tested and their performance is evaluated on the currently available data. An open source utility allowing us to compare several classifiers is Weka , which has also the advantage of being able to connect directly to the feature database (mySQL). We perform cross-validation using various classifiers to get an idea of how discriminative our features are.

While Weka is an external tool, we have also included libsvm [9] into, an easy-to-use Support Vector Machine (SVM) classifier. The SVMs had finally the best overall results, whereas the tested classifiers k-nearest neighbor, naïve Bayes and C4.5 perform slightly worse.

2.4 Lung segmentation as data preparation for classification

When submitting a new image for analysis and as diagnostic aid, we concentrate on the part of the image that we are interested in, the lung tissue. While manual region selection of the image is still possible (using the tool described in 2.1 – only without a label), automatic segmentation is desired to minimize user interaction in the final diagnostic aid step. To this aim we use an algorithm described in [10] to find an optimal threshold for lung tissue segmentation, which works on DICOM images having a full 12-bit resolution as well as on the jpeg images from our radiology teaching file. As basis for the segmentation we use itk. In Figure 2, a lung CT, its segmented version and a view of the outline discovered by the software can be seen.

Figure 3: A lung CT and the tissue of the two lung halves segmented.

For the final texture classification we do not plan to take into account the entire lung tissue but rather the diagnostically interesting part, which is the outside part of the lung with less vessels that can change the texture strongly and introduce noise for the classification. The inside part with the vessels is automatically removed from further analysis.

2.5 Classification of lung blocks

For the classification step of a new lung image, a partitioning of the image into blocks is performed. Then, the features of each block in the (manually or automatically) marked regions corresponding to lung tissue are extracted. The samples are created by the block's features and have no label attached to them, yet. The integrated classifier performs the classification and attaches a label to each block based on previous learning data. Best results are obtained when using a block size in between 16x32 pixels with 32 (resulting in a total of 833 blocks) being slightly better. A larger size of the blocks left two few regions. We also performed tests with several overlaps but with such a small number of regions this could lead to having the training data overlapping too much with the test data and resulting in incorrectly good results.

2.6 Slice selection from a CT volume

This step is currently not implemented and will likely be the last part to be done, as its development is not crucial. Goal is basically to perform the task of the medical doctor to find the slice(s) that best characterize a disease. Once a large database of labeled tissue samples is ready, it will be fairly easy to process the entire volume slice by slice and select those slices with the largest part of the tissue being marked as non-healthy for further inspection. To select several slices, we can give the system a combination of maximum number of slices and threshold of unhealthy tissue. Selected slices can be marked in the volume data directly by highlighting the part containing most pathologic blocks.

2.7 Result presentation

The goal of the results presentation is not to make the decision for the radiologist but rather to highlight parts of the lung tissue that are classified by the system as pathologic. Highlighting of the background in colored shades is planned instead of grey scales in parts of the images being classified as pathologic. Each color presents one of the classes detected in the classification step. Currently, the results are only presented in a 2D view and one image at a time, meaning that the slices with the largest pathogenic parts are taken and displayed. It can also be imagined to present the results in 3D, where the entire pathogenic area over several slices can be highlighted within the volume. Retrieval of similar cases from the reference database will enable the MD to verify a diagnosis. This is easily possible through an image retrieval application using visual features from the currently active case and comparing them those stored and labeled for past cases in the database.

3. Results

We currently have a framework in place allowing to acquire knowledge from the radiologists in the form of marked regions and annotations within images. The database is still small but a larger number of cases is planned. The acquired data with labels is used to train the classifiers. This means that with a growing number of judged cases from the radiologist, the system is expected to perform better. Lung segmentation works reliably and stable as well as the partitioning of the lung tissue into small blocks and feature extraction. All these steps work in a completely automated fashion. The MD can simply feed a volume of lung CT images into the system. The images are segmented and portioned into smaller blocks automatically. These blocks are classified and unhealthy tissue is marked in a different color in the images so the MD has a feedback for regions to inspect further.

Current results using the 112 regions and block sizes of 32x32 pixels (833 blocks) are 83%, when using cross validation between healthy and non-healthy tissue with SVM classifiers. Nearest neighbors lead to 75.4% correctly classified, Bayes to 73% and C4.5 to xxx. When classifying into the 8 available classes, the nearest neighbor classifier reaches 85.5% correctly classified. Problems occur especially for classes with a very small number of representative blocks. Another problem is that the radiologist marked regions with a margin, leading to several blocks that are actually healthy tissue but close to a pathologic region and labeled as pathologic. This explains part of the errors.

We currently run the framework on a simply desktop computer with a Pentium IV processor with 2,8 GHz and 1 MB of RAM. On this computer, the segmentation takes around 5 seconds per slice and the subsequent cutting into blocks, feature extraction and classification another 2 seconds. Thus the analysis of a single slice is almost interactive whereas an entire volume takes a few minutes before results can be displayed. We still need to experiment with the classification part and also with the features that we extract from the images to obtain an optimal feature set for classification. The current framework is by now a research tool, designed to ease experimentation of features, classifiers, parameters, etc. The final system will probably discard a lot of these options, be much simpler and focus more on the user interface and the results display to the user.

4 Conclusions

This article presents a framework to aid the diagnostic process for lung diseases using lung CTs. The domain has shown its potential in studies and our current cross validations leads to good results. The steps of the diagnostic process are performed in an automatic way. Abnormalities are highlighted in the original images by a change of color. Once we have a larger database accessible, more quantitative evaluation is needed to evaluate the algorithm quality and show the usefulness of the application in a clinical environment. Many parameters need to be optimized, from the feature extraction phase to the training step and the classifiers employed. We also need to think about optimal block size of lung tissue and whether we should rather take overlapping blocks to avoid misclassifying small parts of the texture and reduce false positives.

Lung CT analysis has shown its usefulness in practice by improving the diagnostic quality especially of non-experts. Now it is important to create large reference databases and evaluate the many visual descriptors and techniques available to create a robust framework for routine use that needs to have as many steps of the process in an automatic way as possible. Several questions still need to be solved before routine use, for example the handling of other available data on the patients. The age can play an important role for the texture of the lung tissue. For the classification we need to integrate all these data into the framework. The possibility to compare the images with annotated cases from the reference databases is expected to further increase acceptance of the technology because the system does not make a decision by itself but rather points out interesting areas of the lung tissue and gives evidence on these areas by supplying similar past cases.

References
[1] AWM Smeulders, M Worring, S Santini, A Gupta and R Jain, Content-Based Image Retrieval at the End of the Early Years, IEEE Transactions on Pattern Analysis and Machine Intelligence 22(12) pp 1349-1380, 2000.
[2]·H Müller, N Michoux, D Bandon, A Geissbuhler, A r eview of content-based image retrieval systems in medicine – clinical benefits and future directions, International Journal of Medical Informatics, 73, pp 1-23, 2004.

[3] TM Lehmann, MO Güld, C Thies, B Fischer, K Spitzer, D Keysers, H Ney, M Kohnen, H Schubert, BB Wein, Content-based image retrieval in medical applications, Methods of Information in Medicine,43, pp 354-361, 2004.

[4] P Korn, N Sidiropoulos, C Faloutsos, E Siegel, Z Protopapas, Fast and effective retrieval of medical tumor shapes, IEEE Transactions on Knowledge and Data Engineering, 10(6) 889—904, 1998.

[5] LHY Tang, R Hanka, HHS Ip, A review of intelligent content-based indexing and browsing of medical images, Health Informatics Journal 5, 40—49, 1998.

[6] C.-R. Shyu, CE Brodley, AC Kak, A Kosaka, AM Aisen, LS Broderick, ASSERT: A physician-in-the-loop content-based retrieval system for HRCT image databases, Computer Vision and Image Understanding 75 (1—2), pp. 111—132, 1999.

[7] AM Aisen, LS Broderick, H Winer-Muram, CE Brodley, AC Kak, C Pavlopoulou, J Dy, CR Shyu, A Marchiori, Automated storage and retrieval of thin-section CT images to assist diagnosis: System description and preliminary assessment, Radiology, 228, pp. 265-270, 2003.

[8] AK Jain, RPW Dvi, J Mao, Statistical Pattern Recognition: A Review, IEEE Transactions of Pattern Analysis and Machine Intelligence, 22 (1) pp. 4-37, 2000.

[9] CC Chang, CJ Lin, libsvm, a library for support vector machines, Technical report, available with software at http://www.csie.ntu.edu.tw/~cjlin/libsvm/

[10] S Hu, EA Hoffman, JM Reinhardt, Automatic lung segmentation for accurate quantification of volumetric X-ray CT images. IEEE Transactions on Medical Imaging, 20(6), pp. 490-498, 2001.

7. Address for correspondence

Henning Müller
University and Hospitals of Geneva, Service of Medical Informatics
24, rue Micheli-du-Crest, CH-1211 Geneva 14, Switzerland
henning.mueller@sim.hcuge.ch, http://www.sim.hcuge.ch/medgift/
tel ++ 41 22 372 6175, fax ++41 22 372 8680