KU LEUVEN

Shape-based feature analysis for nodule detection in lung images

Mission definition

Kim Nuyts (s0200745)

Sven Van Hove (s0190440)

March 22, 2014

Task specification

Goal

The goal of this project is to develop an automatic method for detection and segmentation of lung nodules, based on assumptions about the shape and appearance of their features. An algorithm based on machine learning techniques will be used to tackle this goal. One of the challenges is the selection of the features. The fact that some nodules are embedded in surrounding tissue (e.g. lung wall, blood vessel, etc.) presents another important challenge.

Given the limited amount of time we have for this project, we decided to focus our work on CT scans. Future research may also explore the opportunities for MRI scans or other imaging modalities. The CT images are acquired from public databases like the LIDC-IDRI database or the datasets made available by ANODE09 (van Ginneken *et al.*, 2010). ANODE09 is an initiative to compare lung nodule detection systems with a single evaluation protocol, based on a single common dataset of 5 annotated training CT scans and 50 test scans (Consortium for Open Medical Image Computing, 2014). This allows researchers to make an objective comparison of the performances of their algorithm with that of other teams, which was an issue before due to the use of differences in training and test datasets, in performance evaluation methods and in targeted nodule groups. The detection algorithm developed in this project will be trained based on the LIDC-IDRI database images. Later on, it will also be trained and tested on the images of ANODE09. The hit criterion of ANODE09 requires a set of findings for a processed scan, a set of 3D positions and a degree of suspicion for all the detected nodules (van Ginneken *et al.*, 2010). This criterion will also be used for assessing the performances of our algorithm.

The literature review revealed that there are a few main – although not mandatory – steps in a lung nodule detection algorithm: pre-processing, lung (nodule) segmentation, nodule detection and false positive reduction (Lee *et al.*, 2012). During the pre-processing step noise and artefacts are reduced to improve the quality of the images. During this step the scans can also be subsampled to isotropic resolution etc. Then the scans are segmented. The lung lobe regions are identified and the rest of the image can be removed. This increases the accuracy of the nodule detection and reduces the computational cost at the same time. The lung nodule detection step can be a separate phase in the process or can coincide with the segmentation step. However, it is suggested that including both a segmentation and a detection step in the algorithm might increase the accuracy of the system (Lee *et al.*, 2010). There are three major detection techniques: a classification based approach, a template matching method and a method that relies on clustering. A combination of these three techniques is also a possibility. The last step is the false positive reduction. Reducing the amount of false positives at the end by using specific features decreases the computational cost of the overall process compared to when this step would be integrated in the nodule detection. This is a very important step as it increases the sensitivity of the algorithm a lot. Our algorithm will implement all steps mentioned in order to achieve a performance as high as possible. If the algorithm turns out to perform well compared to the other methods that attended the ANODE09 challenge, we might consider participating.

Task Decomposition

Figure 1: Task Decompositionshows the high-level decomposition of our task in multiple subcomponents.

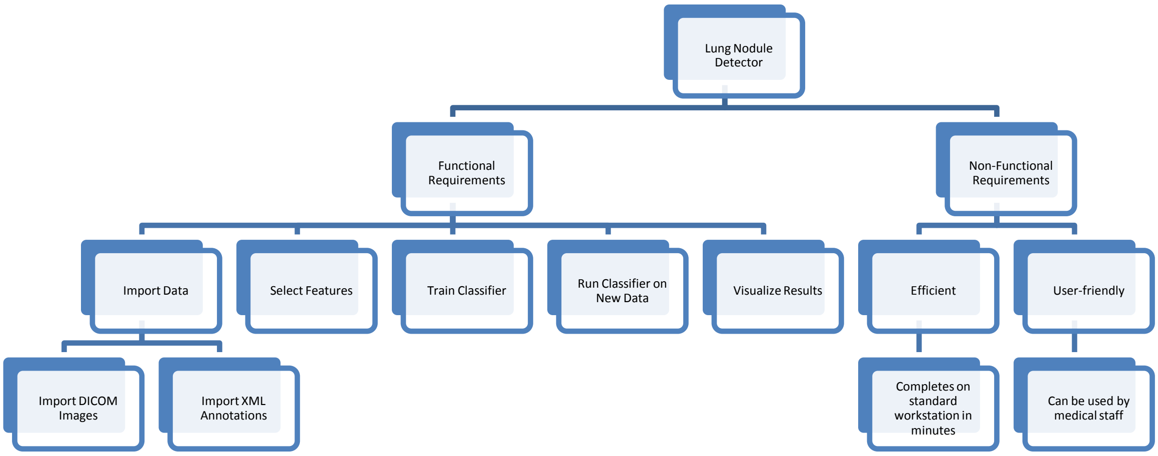

Figure 1: Task Decomposition

First of all, we make a distinction between functional and non-functional requirements. Functional components actually perform some specific task, while non-functional components simply state additional requirements such as efficiency or user-friendliness. Considering that this project aims to be a proof of concept rather than a commercially ready product, the functional components will require most of our attention.

There are several of these functional components, with feature selection, training and running of the classifier being the three most important ones. Importing the data and visualising the results are also necessary but not so interesting from an academic point of view.

The part on efficiency will be elaborated on below, but in general terms we aim to make this program finish on a modern workstation in just a couple of minutes. Because the analysis is fully automated, processing can start right after data acquisition. It needs to be ready by the time the radiologist wants to start examining the scans.

Furthermore, in a real-world scenario the software must be user-friendly enough for medical staff to operate it. However, for our proof of concept it is much less important.

Creative alternatives

A non-exhaustive literature review made clear that a lot of algorithms have already been invented to detect lung nodules, some with excellent performances (Table 1). Although most show impressive results at first sight, they should be taken with a grain of salt. Because every study uses another dataset, the results are very difficult to compare.

Table 1: Comparison of the worst and best performing studies according to the nodule detection method that was implemented.

Method	Performance	Dataset	Authors
Segmentation detection	Sensitivity: 90-98% Specificity: 85-87%	10 scans 300 nodules (isolated, juxta-vascular, peripheral)	Fetita *et al.* (2003)
	Sensitivity: 89.3% 0.21 FPs/slice	29 scans 393 nodules (juxta-pleural, pleural tail, juxta-vascular)	Lin *et al.* (2002)
Classification detection	Sensitivity: 100% 0.88 FPs/scan	8 scans 98 nodules	Chang *et al.* (2004)
	Sensitivity: 94% 5 FPs/slice	32 scans 62 nodules	Suzuki *et al.* (2006)
	Sensitivity: 77.71% Specificity: 87.18%	154 nodules	Pereira *et al.* (2006)
Segmentation - Template matching detection	Sensitivity: 100% 13.375 FPs/scan	16 scans 16 nodules	Ozekes *et al.* (2008)
	Sensitivity: 67% 10 FPs/slice	11 scans 67 nodules	Lee *et al.* (1997)
Segmentation-classification detection	Sensitivity: 100% 0.27 FPs/slice	47 scans 85 nodules (solid)	Dehmeshki *et al.* (2004)
	Sensitivity: 59% 19.2 FPs/scan	33 scans 57 nodules	Oda *et al.* (2002)
Classification aided by clustering	Sensitivity : 98.33% Specificity: 97.11%	32 scans 1203 nodules	Lee *et al.* (2012)
	Sensitivity 84.0% 1.74 FPs/slice	34 scans 63 nodules	Gurcan *et al.* (2002)

Intelligent choice

As was mentioned before, it is believed that combining several detection methods yields better results (Lee *et al.*, 2010). A few methods showed good results overall and were considered for this project. The list below shows advantages and disadvantages for several methods.

<u>Support vector machines</u>

- Better than Random Forest according to Ashfaq *et al.* (2013)
- Complexity is $O(n^2)$ at best in number of samples, based on Chang *et al.* (2011)
- Problems with noisy data (which will be the case for this project)
- Features need to be determined in advance (there is no standard feature set)
- Problems if features do not have the same dimensions/magnitudes

<u>Calibrated boosted trees</u>

- Slightly better than Random Forest according to Caruana *et al.* (2006)

- Problems with noise: choice and tuning of parameters of boosted trees method should be done carefully (which may take some time)
- Training of the algorithm is inherently sequential (which makes it slower)

Methode E, ISI-CAD

- Best method according to ANODE09 comparison by van Ginneken *et al.* (2010)
- False positive detection based on k-nearest neighbour classifier (new approach)
- Problems if features do not have the same magnitudes (e.g. diameter of sphere or distance of nodule to certain point)

Random forests

- Complexity is O(n log n), which makes is suitable for larger datasets
- No problem if features are not known in advance (RF can determine the main features for us)
- Not extremely sensitive to noise

By process of elimination, it was decided to use the Random Forest approach. An extra advantage of Random Forests is that this method has not been exhaustively explored yet in this field of research which makes it interesting to use. To increase the sensitivity and keep the computational cost as low as possible, this approach can be combined with a false positive reduction technique.

Planning of the project

The timing of the project was summarised in a Gantt chart (see attachment). This was roughly based on the time schedule that was provided at the beginning of the project:

- Analysis of the existing methods (15%)
- Design of the method and its components (15%)
- Implementation of the method (50%)
- Validation on open data and reporting (20%)

Technical specifications

Goals: quantification

In order to compare various studies researches de facto agreed on some standard statistical measures, the most important of which are sensitivity (or “true positive rate”), specificity (or “true negative rate”) and Receiver Operating Characteristic (ROC). Other measures such as accuracy and precision are also occasionally mentioned, but are less relevant.

$$sensitivity = \frac{\#\, true\ positives}{\#\, true\ positives + \#\, false\ negatives}$$

$$specificity = \frac{\#\, true\ negatives}{\#\, true\ negatives + \#\, false\ positives}$$

The ROC is a curve which plots the sensitivity versus the false positive rate ($1 - \text{specificity}$). When a ROC value is mentioned, it refers to the area under this curve.

A study performed by Lee *et al.* (2012) used the random forest approach and reached a sensitivity and specificity of respectively 98.33% and 97.11% (ROC value of 0.9786). However, the comparative study of van Ginneken *et al.* (2010) indicated ROC values like these are not to be expected in this project. The sensitivities at seven levels of false positive detection were calculated and these sensitivities were then averaged. The best performing method in this study yielded an average sensitivity of 0.632 for the detection of all kinds of nodules. The sensitivity per nodule type was also provided: small nodules (0.634), large nodules (0.628), isolated nodules (0.609), vascular nodules (0.693), pleural nodules (0.435) and peri-fissural nodules (0.766).

These results also show that the ease of nodule detection also depends on the type of nodule. As this information is not available in the annotations of the scans and as we will not cooperate with a radiologist, it is not possible to differentiate between the nodules in this project. Therefore, our aim will be to detect all types of nodules, regardless of their size or anatomical location. However, the annotations assign a probability of malignancy for each nodule. Separating the detected nodules into a malignancy or benignancy class is not the main aim of this project, but this might be implemented as an extra feature.

In the ideal case, the algorithm would be able to do the processing in a couple of minutes. This would be very interesting for a commercial software product. However, considering the available computational power (a laptop) and the scripting language that is used, this would not be feasible. Python is an interpreted language which makes it inherently slower than compiled languages such as C++. Nevertheless, Python was chosen for its rapid prototyping abilities. Future work may implement our algorithm in C++ or another compiled language to speed up the computational process.

Selection of the features

An non-exhaustive literature review revealed commonly used features used in automatic nodule detection. Tartar *et al.* (2013) primarily used morphological features: area, perimeter, diameter, solidity, eccentricity, aspect ratio, compactness, roundness, circularity and ellipticity. To select these features the minimum Redundancy Maximum Relevance (mRMR) method was applied. Selecting relevant features is import to improve the accuracy of the algorithm and to reduce the processing time. For this project we might consider using a similar feature selection method or we can select the features based on an analysis of the most import features used by the Random Forest algorithm. Regardless of the method to select the features, this selection will always be executed in a cascaded manner.

Murphy *et al.* (2009) used 3D local image features which were calculated per voxel: shape index and curvedness. Chen *et al.* (2012) calculated the size, margins, contours and internal characteristics of the candidate nodules. Keshani *et al.* (2013) used 2D stochastic features – grey level values and intensity values- as well as 3D anatomical features to remove the bronchioles from the list of candidate nodules. The features selected by Teramoto and Fujita (2013) were area, surface area, volume, CT value, convergence, diameter and overlapping area. The algorithm of Ozekes and Osman (2010) implemented 3D features such as straightness, thickness, vertical and horizontal widths, regularity and vertical and horizontal black pixel ratios.

The process

We decided to divide the implementation process over three phases.

Phase I [27/02/2014 – 21/03/2014]: Explorative Research

- Develop XML parser to read annotations
- Find python package to parse DICOM files
- Get hands-on experience with medical images using MeVisLab
- Experiment with classifiers in python on well-known datasets

Phase II [24/03/2014 – 20/04/2014]: Basic Classifier Implementation

- Make first features selection
- Train and run classifier on imaging data
- Meanwhile continuously validate the results by comparing standard measures
- First focus on "easy-to-spot" nodules, later consider all nodules

Phase III [21/04/2014 – 08/05/2014]: Improve algorithm efficiency

- Further extend the algorithm
- Improve the feature selection
- Enhance the accuracy
- (lower the processing time)
- …

References

Ashfaq, A.K., Aljahdali, S., and Hussain, S.N. (2013). Comparative prediction performance with support vector machine and random forest classification techniques. International Journal of Computer Applications 69(11): 12-16.

Caruana, R. and Niculescu-Mizil, A. (2006). An empirical comparison of supervised learning algorithms. In: Proceedings of the 23rd International Conference on Machine Learning, Pittsburgh, PA.

Chang, Chih-Chung, and Chih-Jen Lin. (2011). LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology 2(3): 27.

Chang, S., Emoto, H., Metaxas, D.N., *et al.* (2004). Pulmonary micronodule detection from 3D chest CT. Medical image computing and computer-assisted intervention. Vol. 3217, Springer, Berlin. 821-828 p.

Chen, H., Zhang, J., Xu, Y., Chen, B., and Zhang, K. (2012). Performance comparison of artificial neural network and logistic regression model for differentiating lung nodules on CT scans. Expert Systems with Applications 39: 11503-11509.

Consortium for Open Medical Image Computing (2014). Automatic nodule detection 2009. Available at http://anode09.grand-challenge.org/ [date of search: 18/03/2014]

Dehmeshki, J., Chen, J., Casique, M.V., *et al.* (2004). Classification of lung data by sampling and support vector machine. In: 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol 2: 3194-3197.

Fetita, C.I., Prêteux, F.J., Beigelman-Aubry, C., *et al.* (2003). 3D automated lung nodule segmentation in HRCT. Lecture notes in computer science. Vol 2878/2003, Springer, Berlin. 626-634 p.

Gurcan, M.N., Sahiner, B., Petrick, N., *et al*. (2002). Lung nodule detection on thoracic computed tomography images: preliminary evaluation of a computer-aided diagnosis system. Med. Phys. 29: 2552-2558.

Keshani, M., Azimifar, Z., Tajeripour, F., and Boostani, R. (2013). Lung nodule segmentation and recognition using SVM classifier and active contour modelling: a complete intelligent system. Computers in Biology and Medicine 43: 287-300.

Lee, S.L.A., Kouzani, A.Z., Hu, E.J. (2012). Automated detection of lung nodules in computed tomography: a review. Machine Vision and Applications 23:151-163.

Lee, Y., Hara, T., Fujita, H. et al. (1997). Nodule detection on chest helical CT scans by using a genetic algorithm. In: Proceedings of Intelligent Information Systems, IIS ’97. Grand Bahama Island, Bahamas. 67-70 p.

Lin, D.T. and Yan, C.R. (2002). Lung nodules identification rules extraction with neural fuzzy network. In: Wang, L., Rajapakse, J.C., Fukushima, K., Lee, S.Y., Yao, X. (eds.): 9th International Conference of Information Processing (ICONIP), IEEE, Singapore.

Murphy, K., van Ginneken, B., Schilham, A.M.R., de Hoop, B.J., and Prokop, M. (2009). A large-scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbour classification. Medical Image Analysis 13: 757-770.

Oda, T., Kubo, M., Kawata, Y., *et al.* (2002). A detection algorithm of lung cancer candidate nodules on multi-slice CT images. In: Proceedings of SPIE, vol. 4684. 1354-1361 p.

Ozekes, S. and Osman, O. (2010). Computerised lung nodule detection using 3D feature extraction and learning based algorithms. J. Med. Syst. 34: 185-194.

Ozekes, S., Osman, O., and Ucan, O.N. (2008). Nodule detection in a lung region that is segmented with using genetic cellular neural networks and 3D template matching with fuzzy rule based thresholding. Korean J. Radiol. 9:1-9.

Pereira, C.S., Alexandre, L.A., Mendonça, A.M., *et al.* (2006). A multiclassifier approach for lung nodule classification. In: Campilho, A., Kamel, M. (eds.): Image Analysis and Recognition, vol. 4142/2006, Springer Berlin/Heidelberg, Berlin. 612-623 p.

Suzuki, K., Li, F., Li, Q., *et al.* (2006). Comparison between 2D and 3D massive-training ANNs (MTANNs) in CAD for lung nodule detection on MDCT. Int. J. Comput. Assist. Radiol. Surg. 1: 354-355.

Tartar, A, Killic, N., and Akan, A. (2013). Classification of pulmonary nodules by using hybrid features. Hindawi Publishing Corporation, Computational and Mathematical Methods in Medicine. Available at http://dx.doi.org/10.1155/2013/148363

Teramoto, A. and Fujita, H. (2013). Fast lung nodule detection in chest CT images using cylindrical nodule-enhancement filter. Int. J. CARS 8: 193-205.

van Ginneken, B., Armato III, S.G., de Hoop, B., van Amelsvoort-van de Vorst, S., Duindam, T., Niemeijer, M., Murphy, K., Schilham, A., Retico, A., Fantacci, M.E., Camarlinghi, N., Bagagli, F., Gori, I., Hara, T., Fujita, H., Gargano, G., Bellotti, R., Tangaro, S., Bolanos, L., De Carlo, F., Cerello, P., Cheran, S.C., Lopez Torres, E., and Prokop, M. (2010). Comparing and combining algorithms for computer-aided detection of pulmonary nodules in computed tomography scans: the ANODE09 study. Medical Image Analysis 14: 707-722.