

REDUCING FALSE POSITIVE RESPONSES IN LUNG NODULE DETECTOR SYSTEM BY ASYMMETRIC ADABOOST

Martin Dolejš^{*}, Jan Kybic[†]

Czech Technical University in Prague
Faculty of Electrical Engineering
Czech Republic

Stanislav Tůma, Michal Polovinčák[‡]

Faculty Hospital, Motol
Department of Imaging Methods
Czech Republic

ABSTRACT

We are developing a complex computer aided diagnosis (CAD) system to detect small pulmonary nodules from helical CT scans. Here we present a classifier to reduce the number of false positive responses of the primary detector. Our approach is based on an asymmetric Adaboost which enables us to give different weights to missed nodules (false negatives, FNs) and incorrectly detected structures (false positives, FPs). This is useful because there are noticeably more negative examples in the nodule candidate set than real nodules—true positives (TPs). The whole system is meant as a second opinion for a human radiologist to speed up reading the examination. That is why we should detect as many true nodules as possible, while a certain number of FPs is acceptable. The system was tested on 147 cases (36559 slices) containing 357 nodules marked by an expert radiologist. The new classifier significantly reduced the number of false positives, while only a few nodules were incorrectly omitted.

Index Terms— computer aided diagnostic, nodule detection, CT, asymmetric Adaboost.

1. INTRODUCTION

Almost all nodule detecting systems employ a two step scheme. Nodule candidates are detected from 3D data (images) by a sensitive detector based on thresholding [1], region-growing [2], or template-matching [3]. Our approach uses multiscale filtration and morphology operations [4]. All mentioned detectors have high sensitivity, but a lot of non-nodule structures are detected too (low specificity). To suppress these FP detections further classification step is used, e.g. a linear discriminant [5], or artificial neural network (ANN) [6]. This paper describes a classifier based on

^{*}The author was supported by Czech Science Foundation Project 102/07/1317.

[†]The author was supported by Czech Ministry of Education project 1M0567.

[‡]The authors was supported by the grant agency of Czech Ministry of Health under Project NR8314-3/2005.

an asymmetric Adaboost [7].

The whole system we are developing is aimed as a second opinion for a human radiologist to speed up reading the examination, so we should detect as many true nodules as possible, while a certain number of FPs is acceptable.

1.1. Nodule taxonomy

For the purpose of this work we divide nodules into two classes. Parenchymal (non-pleural) nodules are round or egg-shaped lesions in the lung parenchyma; juxtapleural nodules are worm-shaped lesions connected to the pleura. These two classes are detected separately (by two different detectors), and are also treated separately during classification.

1.2. Pre-processing and candidate detection

first, we pre-process the input 2D DICOM data—construct a 3D volume and segment the lungs. For segmentation we use a simple thresholding followed by morphological operations to clean up the contours [4].

As a parenchymal nodule detector we use multiscale filtering by a Gaussian filter. This algorithm detects regular objects. For juxtapleural nodules we use methods of mathematical morphology [4]. Each of the two candidate detectors provides a set of points, the centers of juxtapleural or parenchymal nodule candidates. Finally each suspicious object is segmented by local thresholding.

The aim of this work is to reduce the number of FPs among nodule candidates by an asymmetric Adaboost classifier, while preserving as many “true” nodules as possible. Example of an input CT slice and a detected nodule is in Fig. 3.

2. METHODS

2.1. Symmetric Adaboost

Symmetric Adaboost [7] is an iterative greedy algorithm to construct a strong classifier from a predefined set of weak

ones. During learning all errors are penalized equally.

$$L_s(i) = \begin{cases} 1 & \text{for } C(\mathbf{x}_i) \neq y_i, \\ 0 & \text{otherwise,} \end{cases}$$

where \mathbf{x}_i is a feature vector of the i -th nodule candidate, y_i the label of the candidate (1 for nodule, -1 for non-nodule), $C(\mathbf{x}_i)$ is the output of the strong classifier. The resulting strong classifier aims to minimize the total classification error:

$$E_s = \sum_i L_s(i),$$

where the index s stands for symmetric.

2.2. Asymmetric Adaboost

The classifier can make two types of errors: either to miss a real nodule, or to classify a non-nodule structure as a nodule. If a nodule is missed (FN result), the patient is in danger; if some non-nodule is detected (FP result) physician only has to work longer when checking computer's results.

That is why we want to penalize missed nodules (FNs) more than FPs. We introduce a new weighted error:

$$E_a = \sum_i L_a(i), \quad (1)$$

with a loss function:

$$L_a(i) = \begin{cases} \sqrt{k} & \text{for } y_i = 1 \quad \text{and } C(x_i) = -1, \\ \frac{1}{\sqrt{k}} & \text{for } y_i = -1 \quad \text{and } C(x_i) = 1, \\ 0 & \text{otherwise} \end{cases}$$

and

$$k = \frac{N_n}{N_p} A,$$

where N_n is a number of negative samples (non-nodules) and N_p a number of positive samples (nodules). A is an asymmetry parameter. If $A = 1$, the total weight of positive and negative samples are equal, if $A > 1$ the group of positive examples weights more than the negatives and vice versa.

Error minimization using this loss function can be achieved using asymmetric Adaboost based on pre-weighting of each sample [7].

2.3. Features

Input of the classifier are feature vectors \mathbf{x}_i computed for all detected candidates. We are using 38 features derived from shape (Table 1), intensity (Table 2), covariance matrix (Table 3), and different Hessian matrices (Table 4). Volume, shape and intensity features are measured on the segmented candidate (using thresholding and connected component analysis), for Hessian features computation the 10mm cube neighborhood $c(x_c, y_c, z_c)$ of nodule center x_c, y_c, z_c is used. $c(x_c, y_c, z_c) = \{(x, y, z) : |x - x_c| + |y - y_c| + |z - z_c| \leq 10mm\}$

Shape features	References
$V = N_a V_{vox}$	[8, 9, 3]
$S = N_s S_{vox}/2$	[9]
$F_1 = S/V$	[9]
$F_2 = \text{var}_j \{ \ \boldsymbol{\mu} - \mathbf{e}_j\ \}$	[3]

Table 1. Four features based on the shape of the suspicious region. N_a is the number of region voxels, V_{vox} the volume of a voxel, N_s is the number of surface voxels, S_{vox} the surface of voxel, $\boldsymbol{\mu}$ the coordinates of the region center of gravity, and \mathbf{e}_j the candidate edge points.

Intensity features	References
T_{hr} —threshold from candidate segmentation	[10]
$I_m = \text{mean}(\{I(\mathbf{x}); I(\mathbf{x}) > T_{hr}\})$	[3]
$I_v = \text{var}(\{I(\mathbf{x}); I(\mathbf{x}) > T_{hr}\})$	[3]
$I_{max} = \max(\{I(\mathbf{x}); I(\mathbf{x}) > T_{hr}\})$	[3]
$I_{mb} = \text{mean}(\{I(\mathbf{x}); I(\mathbf{x}) < T_{hr}\})$	[10]
$K = I_{mb}/I_m$	[3]

Table 2. Six features based on the intensity (in Hounsfield units) of the foreground and background in the suspicious region.

Covariance matrix features	References
λ_{C3}	[9]
$M_1 = \lambda_{C1}/\lambda_{C3}$	[10]
$M_2 = \lambda_{C2}/\lambda_{C3}$	[10]
$R_{ef} = \sqrt[3]{\prod_{i=1}^3 \lambda_{Ci}}$	[10]
$V_{ef} = (4/3\pi R_{ef}^3)$	[9]
$F_3 = V_{ef}/V$	[9]
$P_1 = R_{ef}/\lambda_3$	[9]
$P_2 = R_{ef}/\sqrt{S/4\pi}$	[9]

Table 3. Eight features based on the covariance matrix of the suspicious region. $\lambda_{C1} < \lambda_{C2} < \lambda_{C3}$ are eigenvalues of the covariance matrix $\Sigma = E[(\mathbf{x} - \boldsymbol{\mu})^T (\mathbf{x} - \boldsymbol{\mu})]$.

Hessian matrix features	References
$B_\sigma = \lambda_{H1}/(\lambda_{H2}\lambda_{H3})$	[11]
$T_\sigma = (\lambda_{H3} - \lambda_{H1})^2/\lambda_{H1}^2$	[11]

Table 4. Twenty features based on the eigenvalues of the Hessian matrix on different scales. $[H]_{ij} = \frac{\partial^2}{\partial x_i \partial x_j} G_\sigma * f$, where $G_\sigma * f$ is a convolution of the image and a 3D Gaussian filter with $\sigma = 0.4i, i = 1, 2, \dots, 10$.

3. EXPERIMENTS

3.1. Input data

The system (detector + classifier) was tested on real data consisting of 147 scans of adolescent lungs (66 patients, 36559 slices), xy resolution 0.416mm, slice thickness 1mm, acquired by Somatom AR Star CT machine. There were 357 true nodules marked by a human radiologist expert in the dataset. Data was processed by candidate detectors (juxtapleural, parenchymal) and the 38 features were computed for each nodule candidate. In total there were 323390 candidates detected (1210 candidates were in positions of real nodules, i.e. more than 1 candidate per nodule detected).

3.2. Cross-validation

We divided the 323390 nodule candidates into 10 disjoint subsets at random, and learned 10 classifiers using 9 subsets as training data, one left for subsequent testing [12]. We evaluated the mean sensitivity (Fig 1, top) and mean number of FPs/slice (Fig 1, middle). We chose the number of Adaboost steps empirically, according to the mean weighted classification error E_a (Eq. 1). The chosen number was typically between 50 and 100 steps, close to the minimum of E_a . The choice is not critical, as the error E_a usually does not vary significantly after the initial decrease (Fig. 1, bottom). The sensitivity and the number of FPs correspond to one point of the ROC curve (Fig. 2). This process was repeated for 13 different values of asymmetry A from 0.75 to 100.

4. CONCLUSIONS

We have implemented an asymmetric Adaboost classifier to reduce the number of FPs results for an automatic lung nodule detection system. We have also introduced new descriptors for nodule classification and tested the whole system on a large dataset by cross-validation. Our results indicate that this classifier works better than Fisher linear discriminant function used in our previous attempt (whole system sensitivity 74,3% and 2,6 FPs/slice) [4] and also better than results reported in [13] and [14]. The work of Zhang has very small number of FPs, but also low sensitivity for screening programmes. Desired ratio of sensitivity and FPs/slice can be set by choosing a particular asymmetry A during the learning process.

The detection system is implemented in Matlab. Processing one CT examination takes about 15 minutes (candidate detection+classification), classifier learning takes about 7 hours for the whole dataset on 64-bit dual-core machine. Future work should concentrate on improving the candidate detectors to further increase the sensitivity.

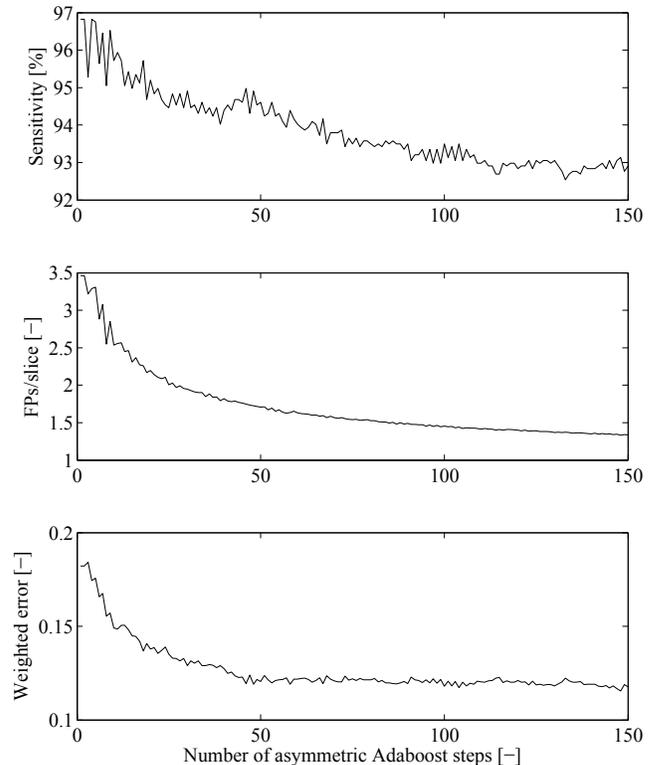


Fig. 1. Sensitivity, FPs/slice, and weighted error (Eq. 1) of the strong classifier as function of number of asymmetric Adaboost steps. Example for $A = 4$.

5. REFERENCES

- [1] J. P. Ko and M. Betke, "Chest CT: Automated nodule detection and assessment of change over time—preliminary experience," *Radiology*, no. 218, pp. 267–273, 2001.
- [2] M. Fiebich, C. Wietholt, B. C. Renger, S. Armato, K. Hoffmann, D. Wormanns, and S. Diederich, "Automatic detection of pulmonary nodules in low-dose screening thoracic CT examinations," in *Proceedings of SPIE*, Bellingham, Washington, USA, 1999, pp. 1434–1439, SPIE.
- [3] Yongbum Lee, Takeshi Hara, Hiroshi Fujita, Shigeki Itoh, and Takeo Ishigaki, "Automated detection of pulmonary nodules in helical CT images based on an improved template-matching technique," *IEEE Transactions on Medical Imaging*, vol. 20, no. 7, pp. 595–604, July 2001.
- [4] Martin Dolejší and Jan Kybic, "Automatic two-step detection of pulmonary nodules," in *Proceedings of SPIE*, Maryllen L. Giger and Nico Karssemeijer, Eds., Bellingham, Washington, USA, February 2007, SPIE.

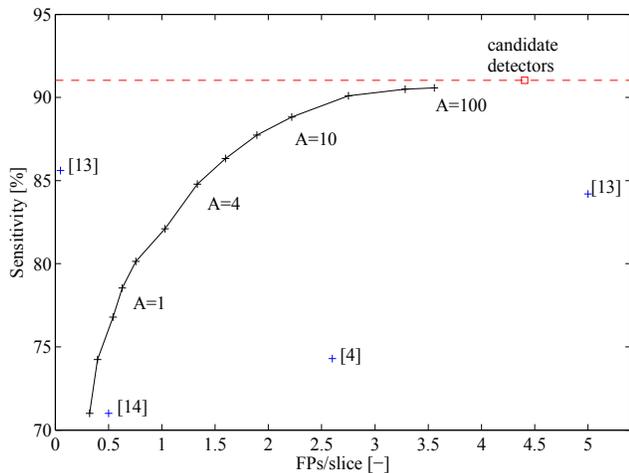


Fig. 2. ROC curve obtained by cross-validation on the complete dataset. The dashed line shows the sensitivity of the candidate detectors, the box corresponds to the performance of the system a without classifier, circles show results of other works with references. Asymmetry parameters A are also shown.



Fig. 3. Example of an input slice, detected nodule marked by a circle.

- vol. 6514 of *Medical Imaging 2007: Computer-Aided Diagnosis*, pp. 3j–1–3j–12, SPIE.
- [5] H. Suzuki and N. Inaoka, “An experimental system for detecting lung nodules by chest X-ray image processing,” in *Proceedings of SPIE*, Bellingham, Washington, USA, 1991, pp. 99–107, SPIE.
- [6] D. Jr. Catarious, A. Baydush, and C. Jr. Floyd, “Initial development of a computer-aided diagnosis tool for solitary pulmonary nodules,” in *Proceedings of SPIE*, Bellingham, Washington, USA, 2001, vol. 4322, pp. 710–717, SPIE.
- [7] Paul Viola and Michael Jones, “Fast and robust classification using asymmetric adaboost and a detector cascade,” in *Advances in Neural Information Processing Systems 14*, T. G. Dietterich, S. Becker, and Z. Ghahramani, Eds., Cambridge, MA, 2002, MIT Press.
- [8] Kazunori Okada, Dorin Comaniciu, and Arun Krishan, “Robust anisotropic gaussian fitting for volumetric characterization of pulmonary nodules in multislice CT,” *IEEE Transactions on Medical Imaging*, vol. 24, no. 3, pp. 409–423, March 2005.
- [9] Xiangwei Zhang, Geoffrey McLennan, Eric A. Hoffman, and Milan Šonka, “Automated detection of small-size pulmonary nodules based on helical CT images,” in *Proceedings of 19th International Conference, Information Processing in Medical Imaging*, G.E. Christensen and Milan Šonka, Eds., Berlin, Germany, July 2005, pp. 664–676, Springer-Verlag.
- [10] Martin Dolejší, “Detection of pulmonary nodules from CT scans,” MSc Thesis CTU–CMP–2007–05, Center for Machine Perception, K13133 FEE Czech Technical University, Prague, Czech Republic, January 2007.
- [11] Berkman Sahiner, Zhanyu Ge, Heang-Ping Chan, Lubomir M. Hadjiiski, Naama Bogot, Philip N. Cascade, and Ella A. Kazerooni, “False-positive reduction using hessian features in computer-aided detection of pulmonary nodules on thoracic CT images,” in *Proceedings of SPIE*, J. Michael Fitzpatrick and Joseph M. Reinhardt, Eds., Bellingham, Washington, USA, 2005, vol. 5747, pp. 790–795, SPIE.
- [12] Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern classification*, Wiley Interscience Publication. John Wiley, New York, 2nd edition, 2001.
- [13] Binsheng Zhao, Gordon Gamsu, Michelle S. Ginsberg, Li Jiang, and Lawrence H. Schwartz, “Automatic detection of small lung nodules on CT utilizing a local density maximum algorithm,” *Journal of Applied Clinical Medical Physics*, vol. 4, no. 3, pp. 248–260, Summer 2003.
- [14] Samuel G. Armato III, Maryellen L. Giger, and Heber MacMahon, “Automatic detection of small nodules in CT scans: Preliminary results,” *Medical Physics*, vol. 28, no. 8, pp. 1552–1561, August 2001.