Performance of a combined CADe (Computer Aided Detection) system based on different approaches in the automatic search of pulmonary nodules in lung CT scans

Poster No.: C-0948
Congress: ECR 2011
Type: Scientific Exhibit
Authors: M. E. Fantacci\textsuperscript{1}, M. Barattini\textsuperscript{2}, R. bellotti\textsuperscript{3}, P. Bosco\textsuperscript{4}, N. Camarlinghi\textsuperscript{1}, R. Cataldo\textsuperscript{5}, P. Cerello\textsuperscript{6}, I. Gori\textsuperscript{7}, \textsuperscript{1}Pisa/IT, \textsuperscript{2}Carrara (MS)/IT, \textsuperscript{3}Bari/IT, \textsuperscript{4}Genova/IT, \textsuperscript{5}Lecce/IT, \textsuperscript{6}Torino/IT, \textsuperscript{7}SAN DONATO MILANESE/IT

Keywords: Computer applications, CT, CT-High Resolution, Computer Applications-Detection, diagnosis, Neoplasia

DOI: 10.1594/ecr2011/C-0948

Any information contained in this pdf file is automatically generated from digital material submitted to EPOS by third parties in the form of scientific presentations. References to any names, marks, products, or services of third parties or hypertext links to third-party sites or information are provided solely as a convenience to you and do not in any way constitute or imply ECR's endorsement, sponsorship or recommendation of the third party, information, product or service. ECR is not responsible for the content of these pages and does not make any representations regarding the content or accuracy of material in this file.

As per copyright regulations, any unauthorised use of the material or parts thereof as well as commercial reproduction or multiple distribution by any traditional or electronically based reproduction/publication method is strictly prohibited.

You agree to defend, indemnify, and hold ECR harmless from and against any and all claims, damages, costs, and expenses, including attorneys' fees, arising from or related to your use of these pages.

Please note: Links to movies, ppt slideshows and any other multimedia files are not available in the pdf version of presentations.
Purpose

The purpose of the work here presented consists in the evaluation of the performance of CADe (Computer Aided Detection) systems for automated lung nodule identification on multislice CT examinations based on different analysis approaches and of their combination, and to evaluate the impact of this system on radiologists’ performance in the usual reading conditions of clinical practice.

In fact, lung cancer is one of a relevant health issues in developed countries and the most common cause of cancer-related deaths, with about 28% and 19% of all cancer-related deaths in the United States [1] and in European Union [2], respectively. It most commonly manifests itself as non-calcified pulmonary nodules. Computed Tomography (CT) was shown to be the most sensitive imaging modality for the detection of such nodules [3]. Therefore, screening programs based on CT, realized with low dose settings, are regarded as a possible approach for detecting lung cancers at early stages [4] and reducing the number of lung cancer deaths, as recently confirmed by the U.S. National Cancer Institute in its publication of the first results obtained with the National Lung Screening Trial (NLST) [5]. The amount of single slices to be analyzed in multi-detector helical CT examinations can be very large and the images can be very noisy, especially screening settings with low dose and with thin collimation are used. It was indeed demonstrated that a large number of nodules (20-35%) can be missed in screening diagnoses [6].

Therefore, Computer Aided Detection (CAD) methods able to automatically search suspect pathological objects could be very useful for radiologists. It has already been demonstrated [7-9] that, in addition to a considerable time saving, the sensitivity of radiologists assisted by Computer Aided Detection (CAD) systems improves with respect to the performance of the same radiologists alone.

However, it is not guaranteed that a single CAD scheme would be the best solution. It is indeed more likely that by means of different approaches better results can be obtained [10]. Following this idea, the MAGIC-5 Italian Collaboration [11,12] adopted a development strategy that led to different CADe algorithms for automated lung nodule identification. In the work here presented we discuss how the different approaches were combined to generate a common result and describe a new plugin (called M5) for the OsiriX DICOM viewer to visualize single and combined CADe findings and allow the clinical test of the CADe system as radiologists’ support.
Methods and Materials

The three CADe system prototypes, described more in detail in the following, are:

- the CAM\textsuperscript{CAD} (Channeler Ant Model CAD), in which, after the lung volume segmentation [13], the Channeler Ant Model is used as a segmentation method for the vessel tree and the nodule candidates and a feed-forward artificial neural network is implemented in order to classify the segmented objects [14-16];

- the RGVP\textsuperscript{CAD}, in which, after the lung volume segmentation [13], a region growing algorithm is iteratively applied to the lung volume to detect nodule candidates, a thresholding on the candidate volume and sphericity and a neural network classifier are applied to reject the false positive findings [16,17];

- the VBNA\textsuperscript{CAD}, in which, after the lung volume segmentation, a 3D dot-enhancement algorithm identifies the internal nodule candidates [18,19], a procedure enhancing regions where many pleura surface normals intersect provides the juxtapleural nodule candidates [20] and a SVM (Support Vector Machine) classifier working at the voxel level reduces the amount of false positives [16, 21].

The public research database LIDC [22] was used for the training and the validation of all the CAD systems.

These systems have been both trained and tested by means of the public research database LIDC [22] and a dedicated OsiriX plugin to be used for the clinical tests has been developed.

The dataset

The dataset used consists of 138 CT scans from the LIDC [22] database, the biggest publicly available collection of annotated CTs. LIDC is a multi-centre and multi manufacturer database, currently under development, with cases of different collimation, kVp, tube current and reconstructed slice thickness. It therefore provides a general sample which is likely to realistically represent the input from a large scale multi-centre screening program. Presently, several hundred cases of the LIDC database are available; for this study 138 CT "thin slice" scans (i.e., scans with a slice thickness in the 0.5 to 2.0 mm range), were selected.
In order to capture the inter-reader variability the LIDC consortium provides four different annotations made by four expert radiologists for each case, obtained in a two phase reading modality. The LIDC annotations contain three kinds of objects [23]: nodules with diameters > 3mm, nodules with diameters < 3mm and "false positives" with diameters > 3mm. The contours of the objects marked as nodules with a diameter > 3mm were provided by every reader together with eight subjective characteristics in a 1 # 5 scale: subtlety, internal structure, calcification, sphericity, margin, spiculation, texture, malignancy. The CT scans used for the analysis were acquired at a 120 kVp voltage, with a current varying from 40 mA to 172 mA, a pixel spacing in the 0.434 mm to 0.762 mm range and a number of single slices ranging from 154 to 730.

The selected dataset was randomly divided in two homogeneous subsets, each containing 69 CTs, which were used for the training/optimization and for the validation procedures of the three CAD systems, respectively.

The Gold Standard reference was defined as the group of nodules with diameters > 3mm annotated by at least two radiologists. The nodules with diameters > 3mm annotated by only one radiologist were not considered as false findings in the evaluation of the FROC curves. According to these criteria, the training dataset contains a total of 138 (96 internal and 42 juxtapleural) nodules, while the validation dataset contains a total of 114 (95 internal and 19 juxtapleural) nodules.

**The Channeler Ant Model CAD**

The full CT analysis is a sequence of four functional modules: lung segmentation, nodule hunter, filtering stage and neural network classification.

*Lung segmentation*

The lung parenchyma in the CT is identified by means of a 3D region growing method and a wavefront algorithm for the definition of the lung surface on the inner side, followed by a morphological closing with a cylinder from the outside [13]. The "a posteriori" check on the training/testing and validation datasets confirmed that none of the radiological findings is rejected at this stage. The lung segmentation module is also used for the region growing-based (RGVP) approach.

*Nodule Hunting*

The Channeler Ant Model [14,15] is used as a segmentation method for the vessel tree and the nodule candidates. The approach, a sequence of two independent deployments for the right and left lungs, is iterative.

The first ant colony segments the vessel tree, starting from an anthill in the vicinity of the root of the tree.
Ants explore (i.e. live in) a 3D environment described in terms of positions and intensities of voxels. Their life cycle is a sequence of atomic time steps, during which they behave according to a predefined set of rules [20]: they release pheromone while moving in the 3D environment defined by the lung volume; they also change their energy, so as to be able to reproduce or die depending on its value. The environment is defined by the voxel image intensities, which can be thought of as the amount of available food for the colony: therefore, voxel intensities should be progressively consumed when the number of visits increases. This mechanism, required to make the colony evolve and explore the environment, is implemented in a complementary way: whenever the limit to the maximum number of visits in a voxel is reached, the voxel is no more available as a destination. When all the ants in the colony have died, the process stops, the segmented object is removed from the original image and its coordinates are added to a list.

In the remaining image, any voxel with intensity above a predefined threshold is a new anthill and a new ant colony is deployed. If the size of an object is large with respect to the maximum expected size of a nodule, as it happens with the bronchial and vascular trees, the object is processed and smaller connected objects are looked for. The procedure is repeated by trying as anthill each voxel in the lung volume with an intensity larger than -700 Hounsfield Units: when no more voxels meet the conditions to become anthills, the information provided by the pheromone map is analyzed.

An iterative analysis is carried on: each voxel with a pheromone content above the minimum accepted value (8,000 units) is used as a seed for a region growing with an adaptive threshold. The threshold value is lowered iteratively for each seed and the selected value is the one corresponding to the minimum growth of the region when the hypothetical threshold is lowered by a quantum.

Whenever a region is larger than a preset value (50 voxels), it is further analyzed in search of nodule candidates connected to it. In order to do so, a rolling sphere scans the finding and disentangles spherical-like structures. The procedure is repeated three times, with spheres of increasing initial radius (1.5, 2.5, 3.5 mm).

In short, a full sequence of ant colony deployments generates a pheromone release map that is analyzed by a dedicated filter, which turns it into a list of candidate findings, each defined by a list of voxels and the values of a set of features related to their geometrical properties, their intensity pattern, their location in the lung.

Filtering

The number of candidates per CT, although depending on the number of slices, ranges between several hundreds to a few thousands per scan, a number far too large to be used as input for a neural network classifier. However, the vast majority of findings is easily rejected with some selections that make use of the correlation between few of the evaluated features: the radius, the sphericity, the fraction of voxels connected to the cage, the so called attach flag (AF), which identifies whether the finding is isolated (AF =
0) or not (AF > 0). If the finding is attached to a larger structure (i.e., the vessel tree), AF is related to the size of the rolling sphere and can range from 1 to 3.

The filtering is performed with a cut function on the histogram that correlates the sphericity to the radius: findings with a sphericity below the cut value at any given radius are rejected. Given the different way in which findings are extracted from the pheromone map, for each AF value the function parameters are different. The filtering level is defined as a compromise between the requirement of maintaining a high sensitivity and the goal of forwarding as less as possible findings to the classification stage.

An additional filter requires the fraction of voxels connected to the cage to be smaller than 0.6, in order to get rid of elongated artifacts attached to the cage.

After the filtering stage, the average number of findings in the training/testing and validation set is 27 and 25, respectively.

For the training/testing set, the sensitivity and the number of false positives/scan, defining the first point on the FROC curve obtained with the classification module, are 0.81 and 0.83, respectively.

**Classification**

Until the end of the filtering stage, very few of the nodule candidate features are used. In particular, no direct information about the image intensities in the candidate voxels is taken into account. The selected set of features for the classification stage makes use of properties that describe the finding size, shape, location, intensity (inside and on the border), as well as the above-defined AF value, which corresponds to different parameters of the nodule hunting algorithm.

The classification was carried on with a four layer feed-forward neural-network: 11 neurons in the input layer, 25 and 7 in the intermediate layers and - obviously - one in the output layer.

The full list of features for the input layer follows: sphericity, radius, Shannon entropy of the inner and the border voxels, skewness, kurtosis, average and standard deviation of the voxel intensities, fraction of voxels connected to the cage, AF value. The classification was optimized on the training/testing sample of 69 CTs and 138 true findings, with a cross validation procedure: 30 sub-list of true findings and false findings were classified as testing sample against all the other true and false findings used as training sample.

**The Region Growing Volume Plateau CAD**

The RGVP-CAD system is an upgrade of the system presented in a previous paper [17] and consists of four main modules:
i) lung parenchyma segmentation via wave front region growing and rolling ball algorithm [13];

ii) nodule candidate detection performed through an iterative region growing with inclusion rule that is the logical AND of the following two rules:

- Simple Bottom Threshold (SBT) rule: a voxel is included in the growing region if its Hounsfield value is greater than a threshold #1;

- Mean Bottom Threshold (MBT) rule: a voxel is included in the growing region if the average of the Hounsfield values of the voxel and its 26 neighbours is greater than a threshold #2.

The #1 threshold is set to the fixed value #1fx, while the #2 threshold is varied in the range [#b, #t] until an optimal value is found for every nodule candidate on the basis of some size-related properties: a graph of the volume of the growing nodule versus the #2 value is generated and #2 is set to a value corresponding to a region were the volume is stable (plateau). Using this procedure with a value of #b close to the Hounsfield unit of air and appropriate values for #1 and #t, it is possible to obtain a good segmentation of nodules in the lung parenchyma without any prior knowledge of their mean Hounsfield value.

iii) statistical and morphological features extraction: a set of seventeen statistical and morphological features are computed: {Volume, Radius Variance, Over Radius Variance, Radius Standard Deviation, Over Radius Standard Deviation, Radius Skewness, Radius Kurtosis, Over Radius Kurtosis, Radius Compactness, Over Radius Compactness, Maximum Distance, Roundness, Ellipticity, Maximum Hounsfield intensity, Standard deviation of Hounsfield intensity, Shannon's Entropy of Hounsfield intensity distribution}. From these features the best discriminating variables according to the sequential backward selection (SBS) procedure [24] are selected;

iv) classification by means of two linear filters on Volume and Roundness followed by a Feed Forward Neural Network (FFNN).

In order to optimize the parameters and to train the RGVP-CAD system a cross-validation technique [25] is used.

**The Voxel Based Neural Approach CAD**

Lung nodules are partitioned in two main classes, depending on their location in the lung. A nodule is labeled either as internal if fully contained in the lung parenchyma or as juxtapleural if connected to the lung volume border, as identified by the dedicated
module. The system deals differently with internal and juxtapleural nodules, by means of two dedicated procedures: CAD\(_{I}\) for internal and CAD\(_{JP}\) for juxtapleural nodules. Both are three-step procedures [16, 18-21]. The first step consists in the lung segmentation; the second step consists in the ROI (Region Of Interest) hunter and performs the candidate nodule selection; the third step consists in the FP reduction. For this last step, an original procedure, the Voxel-Based Neural Approach, aims at reducing the number of FPs in the lists of internal and juxtapleural candidate nodules. The final list of findings is simply obtained by merging the output lists of findings generated by CAD\(_{I}\) and CAD\(_{JP}\).

**Lung segmentation**

An approach based on thresholding, region growing and morphological operators is implemented. In order to outline the shape of the pleura irregularities (including juxtapleural nodules), the lung boundaries are not smoothed. The identified lung mask, including vessels and airway walls, is used for CAD\(_{I}\), whereas its boundary is used for CAD\(_{JP}\).

**ROI hunter for internal nodules**

In the CAD\(_{I}\) the internal nodules are modelled as spherical objects with a Gaussian profile, following the approach proposed in [26]; the 3D matrix of data is filtered with a multi-scale filter function built to discriminate between spherical objects and objects with planar or elongated shapes. The local maxima of the 3D filtered matrix are the internal candidate nodule locations. A large number of false positives is included at this stage, above all crossings between blood vessels.

**ROI hunter for juxtapleural nodules**

In the CAD\(_{JP}\), in order to identify juxtapleural candidate nodules, pleura surface normals are constructed and each voxel is assigned a score proportional to the number of normals intersecting in it. Normals are evaluated using the triangular mesh representing the pleura surface. In particular, the normal to each triangle is calculated by using the vector product between the triangle edges; then the normals to each mesh vertex are evaluated averaging all the triangle normals of the neighbouring triangles.

Since the evaluation of the normal intersections in the real 3D space is a complex and computationally intensive operation, it is implemented in the voxel space. This means that each voxel is associated a score proportional to the number of normals passing through it. To deal with noise, cylinders with Gaussian profile are considered instead of segments [27]. This information is collected in the score matrix S(x,y,z). The local maxima of the 3D matrix S(x,y,z) are the juxtapleural candidate nodule locations. Also in this case a
A large number of FPs is found, mostly due to irregularities in the pleura surface (e.g. apical scars, pleural thickening and plaques) and movement artefacts.

**Classification**

In order to classify the candidate nodule findings obtained in the previous step, an original procedure, the Voxel-Based Neural Approach (VBNA) [19-21] performs the reduction of the number of FPs in the lists of internal and juxtapleural candidate nodules. First, a region of interest (ROI) including voxels belonging to the candidate nodule is defined from each location provided by the previous step. The basic idea of the VBNA is to associate to each voxel of a ROI a feature vector defined by the intensity values of its 3D neighbours and the eigenvalues of the gradient matrix and of the Hessian matrix. In the original VBNA method the feature vectors are then classified by a three-layer feed-forward neural network which is trained to assign each voxel either to the nodule or normal tissue target class. In this paper a different classifier is implemented at this stage: Support Vector Machines (SVM) are used instead of neural networks. At the end of the procedure, each ROI is assigned a degree of suspicion, defined as the percentage of voxels tagged as nodule by the classifier.

**The combination procedure**

As there is no reason to assume that a single CAD scheme would be optimal for nodule detection but it is more likely that different methods have complementary strengths, the outputs of the three CAD sub-systems already described are evaluated and combined following the same procedure adopted for the ANODE09 competition [28]. The findings of each CAD sub-system must be considered in terms of their degree of suspicion (i.e. likelihood or probability to be a true nodule) $p$, which corresponds to the final output of the procedure of candidate nodules classification for the three sub-systems. These probabilities are normalized by associating to each $p$ of a value $f(p)$, calculated as: $f(p)=\frac{TP}{FP+TP+1}$, where TP (FP) are the number of true (false) positives obtained by considering all the CAD findings with $p_i \geq p$ as positive.

The $f(p)$ values can then be considered to be the probability that a finding in the validation set with likelihood $p$ or higher represents a true nodule.

$f(p)$ is computed for every finding from every sub-system. All findings are then sorted so that $f_i$, $i = 1 \ldots n$ and $f_i \geq f_j$ if $i < j$. Starting at $f_i$ with $i = 1$, all findings $f_j$, $j = i + 1 \ldots n$ are checked against a "matching condition" defined by a preselected clustering distance. If two findings $f_i$ and $f_j$ match, $f_i = f_i + f_j$, is set and $f_j$ is removed from the list of findings.

**The OsiriX plugin**
A dedicated plugin (M5) for the OsiriX DICOM viewer was implemented with the Cocoa environment of MacOSX in order to show the findings of the three CAD sub-systems, their combination and the radiologists annotations. OsiriX [28] is an open source DICOM image processing and visualization software for different imaging types (MRI, CT, PET, PET-CT, SPECT-CT, Ultrasounds, etc.), which allows a simple mechanism for writing plugins that expand its functionalities.

The M5 plugin allows annotating nodules and visualizing in different colors the separate CAD sub-system results, their combination and radiologists' annotations. A screenshot of the plugin is reported in Fig. 1. With the M5 plugin, radiologists can visualize and process the CT data using all the tools included in the OsiriX software (zooming, changing the intensity windowing, browsing slices, MIP projections, etc.), identify nodules and annotate them. The annotations consist of simple circular ROIs, identified by their center and radius. It is also possible to select, using the sliders, a different working point for each CAD, to visualize only findings above the selected thresholds. In Fig. 2 an example of nodule visualization is reported: this nodule was detected by two CAD sub-systems (with the settings of Fig. 5), and the corresponding outputs are shown together with their combination and the radiologist annotation. The sliders by which the operative points for each CAD can be set in order to visualize only findings above the selected thresholds are shown in Fig. 3.
Images for this section:

Fig. 0: Example of a screenshot of the OsiriX plugin.

© INFN - Pisa/IT
Fig. 0: An example of nodule visualization.

© INFN - Pisa/IT
**Fig. 0:** A detail of the sliders by which operative points for each CAD are selected.

© INFN - Pisa/IT
Results

In Fig. 1 are reported, in terms of FROC (free-response receiver operating characteristic) curves, the results obtained by means of the separate systems with the 69 CT of the validation dataset (114 nodules in total annotated by at least two radiologists).

The FROC curves have been evaluated using the matching criterion [10] that a CAD finding is considered a true positive if its euclidean distance from the center of the lesion annotated by the radiologists is less than 1.5 times the radius of the lesion annotated by the radiologists.

In Fig. 2 are reported the results obtained by means of the CADs combination in the range 0-10 false positive findings/CT, which is the most interesting for radiologists. Such combination has been computed considering a clusterisation distance of 3 mm. The overall sensitivity is 80% at about 3 FP/scan (considering as Gold Standard reference sample the nodules annotated by at least two radiologists out of four).
Fig. 0: FROC curves obtained with the three CAD systems.

© INFN - Pisa/IT
**Fig. 0**: FROC curve obtained with the CAD combination compared with those obtained with the single systems in the range of FPs of radiological interest.

© INFN - Pisa/IT
Conclusion

We have evaluated the performance of different CAD approaches developed in the framework of the MAGIC-5 Italian experiment. The systems have been trained and tested in the same working conditions by using the same train dataset and the same validation dataset. Such datasets have been extracted from a public research database (LIDC). Such performance was evaluated separately and then has been evaluated the performance obtained by means of their combination. It was already demonstrated [10] that with different systems tested on the same database can be obtained different results, and that the combination of the systems can allow relevant improvements, thanks to the peculiarities in detecting different kinds of nodules.

Here we have shown that also using both the same train dataset and the same validation dataset with different CAD analysis approaches have been obtained different results, which are complementary, as demonstrated by the results obtained with their combination, in particular in the range most interesting for radiologists (i.e. with a number of false positive findings up to 10/CT).

The overall sensitivity is 80% at about 3 FP/scan, which is a very good result, especially taking into account the loose criterion of selection of the Gold Standard reference sample (nodules annotated by at least two radiologists out of four).

In order to verify if and how the use of the combined M5 CAD system is actually useful in helping radiologists, has been already started by means of the M5 OsiriX plugin a clinical validation following a second reader protocol.


[16] Fantacci ME et al., Computer aided detection of nodules in low dose and thin slice lung CT, DOI: 10.1594/ecr2010/C-1053.


Personal Information

Maria Evelina Fantacci, University of Pisa and INFN Sezione di Pisa, Pisa, Italy, fantacci@df.unipi.it

Matteo Barattini, School of Specialization in Diagnostic Radiology, University of Pisa, Pisa, Italy, matteobarattini@gmail.com

Roberto Bellotti, University of Bari and INFN Sezione di Bari, Bari, Italy, roberto.bellotti@ba.infn.it

Paolo Bosco, University of Genova and INFN Sezione di Genova, Genova, Italy, paolo.bosco@ge.infn.it

Niccolo’ Camarlinghi, University of Pisa and INFN Sezione di Pisa, Pisa, Italy, niccolo.camarlinghi@df.unipi.it

Rosella Cataldo, University of Salento and INFN Sezione di Lecce, Lecce, Italy, rosella.cataldo@unisalento.it

Piergiorgio Cerello, INFN Sezione di Torino, Torino, Italy, cerello@to.infn.it

Ilaria Gori, Bracco Imaging SpA, Milano, Italy, ilariagori@gmail.com

Contact of presenting author (M.E. Fantacci): fantacci@df.unipi.it

Contact of MAGIC-5 spokeperson (P. Cerello): cerello@to.infn.it