

Analysis of medical images: the MAGIC-5 Project

Piergiorgio Cerello*

INFN, Sezione di Torino

E-mail: cerello@to.infn.it

The MAGIC-5 Project focuses on the development of analysis algorithms for the automated detection of anomalies in medical images, compatible with the use in a distributed environment. Presently, two main research subjects are being addressed: the detection of nodules in low-dose high-resolution computed tomographies of the lung and the analysis of brain MRIs for the segmentation and classification of the hippocampus as an early marker of the Alzheimer's disease. MAGIC-5 was started by a community of high energy physics software developers, still in constant contact with - some of them also involved in - HEP projects. Some of the MAGIC-5 results, obtained with a newly developed model based on virtual ant colonies for the segmentation of complex structures, will be presented and discussed. The project takes advantage of a close collaboration with representatives of the medical community, that set requirements in terms of performance and functionality. Future developments are focusing more and more on the analysis of neuroimages, trying to exploit at best the huge amount of morphological and functional information generated by MRI and PET imaging.

*13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research,
ACAT2010
February 22-27, 2010
Jaipur, India*

*Speaker.

1. Introduction

Tools and techniques developed for High Energy Physics applications are often suitable, after some adaptation and/or tuning, to tackle other problems, such as the automated search for anomalies in medical images of various kinds and the data management of medical information in a distributed environment, with GRID or WEB services.

MAGIC-5 (Medical Applications on a GRID Infrastructure Connection) [1] was started in 2003 with the goal of developing a GRID-based interface for the use of GRID services in handling and analysing medical data and images to be shared by a geographically distributed community of users. It delivered a prototype GRID interface for the analysis of distributed mammograms, based on Alien Services [2], as early as 2004 and it extended it with a PROOF interface in 2006. Over the years, driven by an increasing interest by the researchers in the team, the focus moved more and more to the development of algorithms for the automated search for anomalies in medical images.

In general, automated algorithms can be useful whenever a large amount of data must be analysed by radiologists: their use can reduce time and cost-related problems as well as increase the sensitivity in detecting diseases and eventually reduce the rate of false positive findings (when combined with the radiologist's reading). Several works have been published, in the past years, for what was the golden example: mammography [3, 4]. In that case, a very large population is subject to screening and it was natural to address the problem of a common infrastructure for the data analysis. However, a still unsolved problem prevents its successful use: different hardware generates mammograms with very different features (resolution, dynamic range, contrast): in such conditions, a common algorithm with a satisfactory and comparable performance over such different images is still far from being available. Therefore, commercially available Computer-Aided Detection (CAD) tools for the analysis of mammograms are usually distributed together with the hardware they were tuned for. However, breast cancer is not the only disease for which automated algorithms can be useful.

2. The MAGIC-5 Project: an overview

MAGIC-5 was launched in 2004 with the goal of developing a CAD system for the analysis of Computed Tomographies of the lung, to be used in a GRID-based distributed environment. A couple of years later, another line of research was started, focused on the objective of an early diagnosis of the Alzheimer's disease, to be obtained with the automated analysis of brain MRI and/or PET scans.

Over the years, the idea of deploying the algorithms on a GRID was dismissed, for practical and strategic reasons. GRID services, unfortunately, do not provide the required flexibility for such a system. By design, the GRID approach is more suitable for a batch than an interactive analysis. In addition, hospitals would only join their own GRID, which should not allow other applications to access computing resources, mostly for data security reasons. Also, in case a dedicated GRID were deployed, it would most likely miss the economic profitability, since it would involve many sites and the site maintenance is not yet optimised. For these reasons, our current view is that CAD analysis should be offered as a WEB-service and provide the results in a format compatible with the most used DICOM visualisation tools.

The MAGIC-5 scientific effort is now mostly on algorithm development, with the exception of a small team that is developing a prototype WEB portal for the access to CAD services.

The approach to any algorithm development has always been driven by the basic idea that the combination of several methods should produce the best overall performance. Both for the analysis of CTs of the lung and brain MRIs and PETs, several algorithms are being developed, to be used in parallel or in sequence.

The present work focuses on the Channeler Ant Model (CAM), a virtual ant model that is designed in a general way and is therefore suitable for the analysis of very different images: in the following the approach and the results for CTs on the lung and brain MRIs are discussed in detail.

3. Lung CT Analysis: goals

Lung cancer is one of the most lethal kinds of cancer worldwide. The overall 5-year survival rate is only 10-15% [5, 6] and no significant improvement has occurred in the last 20 years [7]. Non-calcified small pulmonary nodules are considered as the primary signs of early-stage lung cancers. Computed Tomography (CT) has been shown to be the most sensitive imaging modality for detecting small pulmonary nodules, particularly since the introduction of the multi-detector-row and helical CT technologies [8]. The efficacy of screening trial protocols based on low-dose CT with thin reconstructed slice thickness in reducing the lung cancer mortality rate are currently under investigation in many developed countries [9, 10, 11, 12]. Depending on the screening trial protocol, radiologists may have to identify even very small nodules, carrying out an extremely difficult and time-consuming task. Nodules are spherical or hemispherical objects that can be characterized by CT number values in the same range of those of blood vessels and airway walls and may be strongly connected to them or to the pleura surface. The nodule identification in screening CT is particularly difficult as low-dose CT images show a noisier appearance with respect to the standard-dose ones and an amount of image data as large as 500 2D slices per scan may be generated in case thin reconstructed slice thickness is used. To support radiologists in the challenging task of interpreting screening lung CT scans, researchers focused on computer-aided detection (CAD) methods devoted to the automated identification of pathological objects in the CT scans.

4. Brain MRI Analysis: goals

The incidence increase of the Alzheimer's Disease (AD) and other neurodegenerative diseases, related to the progressive increase in life expectancy, demands more and more efforts for its early diagnosis [13]: although, as of now, there are no available drugs to treat patients, some correlation between the lifestyle and the speed of the disease progression is acknowledged. In order to diagnose AD, it is required to analyze increasing amounts of neuro-images obtained by magnetic resonance (MRI) [14, 15, 16, 17] and/or positron emission tomography (PET). The evaluation of the atrophy degree of some interesting anatomical structures and the study of the connections between different brain regions can provide lots of useful information. In order to do so, an assessment of the relative contribution of white matter (WM), grey matter (GM) and the cerebro-spinal fluid (CSF) to each discretised volume of the MRI (PET) is important: its comparison to templates will provide a numerical indicator of the likeness of the AD diagnosis. In particular, several research efforts

point at the hippocampal region as a brain volume that is very sensitive to the AD presence and development.

5. The Channeler Ant Model (CAM)

The Channeler Ant Model (CAM) [18], based on the natural ant capabilities of dealing with 3-D environments through self-organization and emergent behaviours, is briefly described. Ant colonies, defined in terms of moving, pheromone laying, reproduction, death and deviating behaviours rules, is able to segment artificially generated objects of different shape, intensity, background. The model depends on few parameters and provides an elegant solution for the segmentation of 3-D structures in noisy environments with unknown range of image intensities: even when there is a partial overlap between the intensity and noise range, it provides a satisfactory segmentation with negligible contamination (i.e., fraction of segmented voxels¹ that do not belong to the object).

The deployment of ant colonies in 3-D images is in principle very effective whenever complex connected structures, with several ramifications of different size and intensity, must be identified and reconstructed, as long as a general model with few requirements on parameter tuning is designed and validated on images with known properties (different signal to background ratio and intensity range). Such conditions occur with lung CTs and brain MRIs, where complex structures must be segmented in a noisy environment.

The behaviour of ants, partially derived from [19, 20], is described in terms of four modules: the moving rules, the pheromone laying rules, the reproduction/death rules and the description of the ant response to anomalies (deviating behaviour). Ants explore (i.e., live in) a 3-D spatial environment described in terms of the properties (position, intensity) of discretized VOLUME ELEMENTS (voxels) and their life cycle is defined in terms of atomic time steps, during which ants move from one voxel to a neighbour. So, at time t an ant k is in voxel v_i ; after one life cycle (time t_1) it will move to voxel v_j . In a 3-D environment, a voxel has 26 first order neighbours according to Moore's neighbouring law [21].

Like in nature, in the CAM the random component associated to the way an ant walks is relevant to the choice of the future destination. However, the choice of the direction for a step must also take into account the colony global knowledge of the environment, which is provided by the amount of pheromone already released in a given position. The meaning of a pheromone message is the same as in nature: a large amount of pheromone in a candidate destination must correspond to a high probability of becoming the actual destination.

The probability P_{ij} that a candidate destination is chosen depends on the amount of pheromone in voxel v_j , so it is not directly related to the primary properties of the environment, allowing for the identification of different structures when different pheromone laying rules are defined.

The pheromone-related term in P_{ij} , taken from [19], depends on the osmotropotaxic sensitivity β (the larger it is, the larger the influence of the pheromone trail in deciding the ant's future destination) and on the sensory capacity $1/\delta$ (if the pheromone concentration is too high, it will determine the decrease of the ant's capability of sensing it).

¹a voxel is the smallest defined shape in a 3D image

According to the biological laws of ant colonies, before moving to the future destination an ant k deposits in the voxel it is about to leave a quantity of pheromone T , defined as [20]

$$T = \eta + \Delta_{ph} \quad (5.1)$$

where η is a small quantity of pheromone that an ant would leave anyway and Δ_{ph} , the differential quantity of the pheromone, links the image properties to the pheromone habitat in which the ants live. Its value is a voxel-intensity dependent function:

$$\Delta_{ph} = \Psi(I) \quad (5.2)$$

The choice of the depositing rule is the only point in the model where the original image properties (i.e., the intensity) come in: depending on the Ψ function choice, ants will segment different structures. In the analysis of lung CTs and brain MRIs, the two developments discussed in the following, the regions of interest are high intensity voxels: therefore a simple proportional deposition rule was selected:

$$\Psi(I) = I - I_{min} \quad (5.3)$$

where I and I_{min} are the current voxel intensity and the minimum intensity in the whole image, respectively.

Ants, like all the living creatures, live for a finite amount of time. The life cycle is regulated by a parameter called energy [22], which is assigned at birth with a default value:

$$\varepsilon_0 = 1 + \alpha$$

The energy variation for ant k must take into account the properties of the environment, which are defined by the deposited amount of pheromone Δ_{ph}^k for the current cycle and by the average amount of pheromone per step the colony has deposited since the beginning of its evolution, used as a normalization factor ($\langle \Delta_{ph} \rangle$): the integrated average, used for the normalization, asymptotically converges to a constant factor and is far more stable than the average per cycle.

The energy range is defined by a lower limit, the death energy ε_D and an upper limit, the reproduction energy ε_R : an ant with energy ε_t^k will die whenever $\varepsilon_t^k < \varepsilon_D$ and give birth whenever $\varepsilon_t^k > \varepsilon_R$. In that case, the ant energy is reset to the default starting value ε_0 . The ant life cycle duration is therefore a function of the ratio between the rate of the energy variation (α) and the amplitude of the allowed energy range ($\varepsilon_R - \varepsilon_D$), all of it modulated by the properties of the environment.

The Channeler Ant Model performance was successfully validated on artificial images with a wide selection of shapes, dynamic range, noise level. No parameter tuning was required when the image or the environment properties changed.

6. Lung CT Analysis with the CAM: status and results

Since the Channeler Ant Model is designed as a tool for the segmentation of generic 3D objects of unknown shape, it can be used for the automated search for lung nodules in low-dose CT scans.

The validation was carried on with three samples from different sources: 5 CT scans provided by the ANODE09 challenge organisers [23], 20 CT scans provided by the ITALUNG_CT project,

with similar characteristics to the development and training set (GS20 in the following), 37 CT scans downloaded from the LIDC [24] site, together with their annotations. The validation scans are high resolution, with a reconstructed slice thickness below 2 mm.

The annotation, containing the list of true findings, are provided according to different protocols: in order to compare them to the CAD results, the so-called relevant nodules for the ANODE09 scans and the subsample of nodules highlighted by at least 3 radiologists for the LIDC sample were selected (the 4 radiologists consensus did not provide a statistically acceptable sample, just 19 nodules). The GS20 scans were annotated with the protocol used by the ITALUNG_CT trial and all the findings are considered relevant. It is important to remark that the minimum size required by the three protocols to consider a nodule as relevant for a follow-up was different: 3, 4, 5 mm diameter for LIDC, ANODE09, GS20, corresponding to 39, 39, 51 true findings, respectively.

The lung parenchymal segmentation [25] starts with a simple-threshold 3D region growing applied to the CT volume. The result is a binary mask of the respiratory system, containing the trachea, the bronchi, and the lungs. The next step is the segmentation of the external airways (trachea and bronchi) by a 3D region growing with wave-front simulation and suitable stop conditions, allowing a proper handling of the hilar region. Particular attention is given to detecting and solving the problem of the apparent fusion between the lungs, caused by partial-volume effects. 3D morphology closing with a thin cylinder ensures the accurate inclusion of all the nodules (internal, pleural, and vascular) in the segmented volume.

The Channeler Ant Model is used, in the first stage, as a segmentation method for the vessel tree and the nodules candidates. The ant colony is released in a selected position of the 3D matrix, i.e. the anthill. Each ant behaves according to the CAM set of rules described earlier. When the colony is extinct the generated pheromone map is analysed with a region growing algorithm: the connected region with pheromone release above the saturation threshold (20000 units) that includes the anthill is removed. Moreover, all the voxels with pheromone below threshold are inhibited for further colony deployment. Regions with pheromone release above threshold that are not connected to the anthill are considered as nodule candidates and stored for neural network classification. At the end of the first colony analysis, an iterative procedure scans the whole lung volume: whenever a voxel meets the required conditions to become an anthill, a colony is deployed until extinction and its pheromone map analysed: in this case, however, all the connected regions above threshold become nodule candidates. The procedure terminates when no voxels meet the conditions to become an anthill and generates a list of nodule candidates. For each nodule a set of features are calculated: volume, distance from the cage, maximum, minimum and average intensity, etc.: these values are used to feed the first layer of a feed-forward neural network classifier, with one hidden layer and one neuron in the exit layer. The neural network output is considered to be the probability that the finding is true.

The neural network training/testing is done in cross validation mode for the three samples separately (as the protocol requirements are different). The preliminary results (fig. 1) show that the performance is reasonably good, although likely to be improved. It is particularly interesting to remark that the CAM was used on lung CTs without any optimisation, just "as it is" after the validation on artificial objects.

However, it is likely that the evaluation and definition of new features would improve the performance. As an example, the distance from the cage will be added to the classification: it is

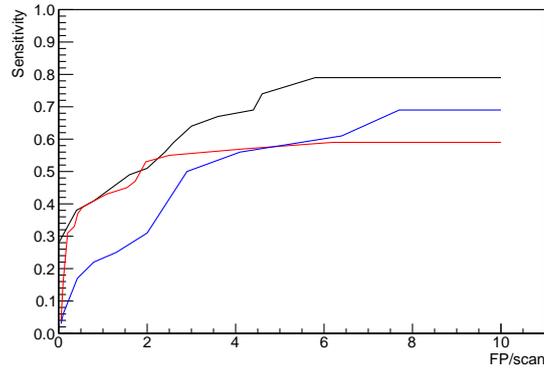


Figure 1: Sensitivity as a function of the number of false findings/scan (FROC curve) for the ANODE09 (black), GS20 (blue), LIDC (red) databases.

expected to be related to the nodule sphericity, since the findings connected to the cage are not spherical, by definition. Therefore, it should help in distinguishing non-spherical findings inside the lung volume (which should be rejected) or attached to the pleura (which should be analysed making use of other features).

A reliable working point could be set at about 8-10 false findings/scan, where the sensitivity is already pretty good. It is also important to point out that, for the LIDC sample, there is an intrinsic disagreement between radiologists of about 20% (38/51 findings for 4/3 radiologists consensus, respectively). Therefore, a 60% sensitivity for a 3 radiologists consensus is compatible with a sensitivity of about 80% for a 4 radiologists consensus.

7. Brain MRI Analysis with the CAM: status and results

The Simulated Brain Database, provided by the Brain Imaging Centre, Montreal Neurological Institute, McGill University and available on-line [26], contains a large number of brain MRI images with different noise levels from 0% to 9% and different bias-field levels from 0% to 40%. The database also includes the images of the phantom used to obtain the simulated ones, in which the voxel the intensity is proportional to the percentage of white matter. The phantom represents a standard segmentation and can be used as a gold standard to evaluate the CAM segmentation result.

The CAM parameters are the same used for the artificial objects and lung CT analysis, once again a proof of the CAM adaptability to images with different dynamic range and noise levels.

For the preliminary analysis, a portion of 15 slices in the middle of the stack of T1 weighted images with $1 \times 1 \times 1 \text{ mm}^3$ voxel dimension was considered, as in [27].

The sensitivity S and contamination C , defined as follows:

$$S = N_R / N_T \quad (7.1)$$

$$C = (N_{TR} - N_R) / N_T \quad (7.2)$$

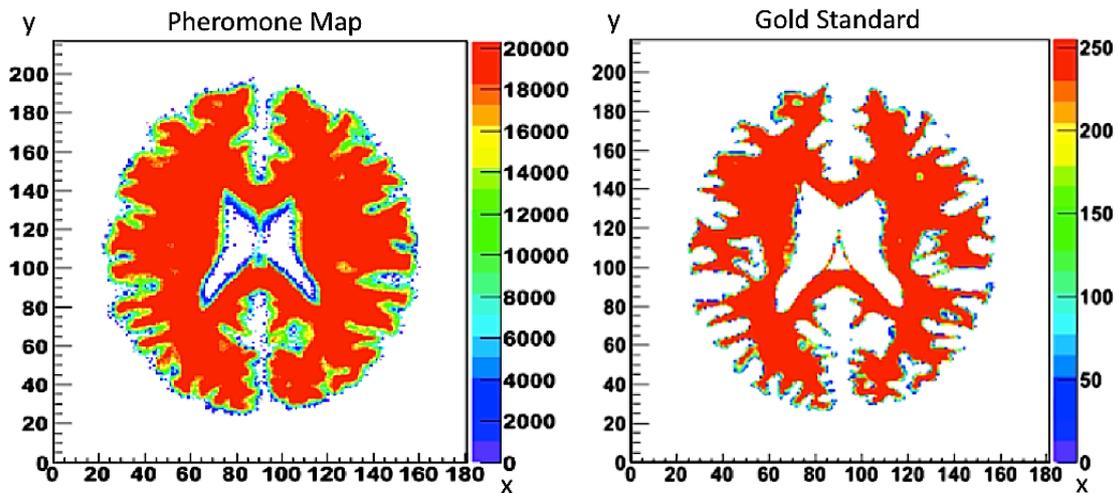


Figure 2: Pheromone map (left) and the corresponding Gold Standard (right).

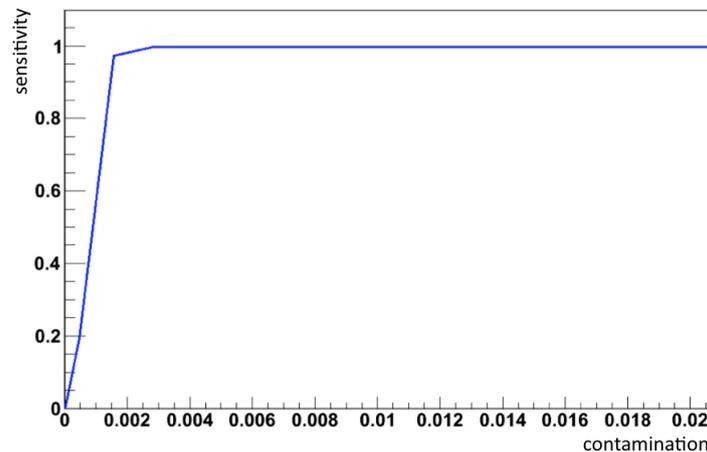


Figure 3: Figure 3. Sensitivity as a function of the contamination for a T1 image with noise=0% and RF-biasfield=0%.

N_R is the number of voxels correctly segmented by the CAM, N_{TR} the total number of segmented voxels and N_T the number of voxels that belong to the gold standard structure, were used as performance indicators.

Figure 2 2 shows an example of comparison between the pheromone map obtained by CAM segmentation and the gold standard. Voxels with saturated pheromone (i.e., 100% white matter) are represented in red. The pheromone map analysis is divided in two steps: first the voxels at intensity 0 and 255 are considered.

Figure 3 2 shows the ROC curve: the CAM reaches very high sensitivity values (i.e. $S = 0.997$) with a negligible contamination (i.e. $C = 0.003$).

In the second step, voxels with Gold Standard intensity between 0 and 255 are analyzed, in order to study the correlation between amount of pheromone deposition and the percentage of white matter, so as to be able to estimate it. The CAM is then required, for the first time, to provide a

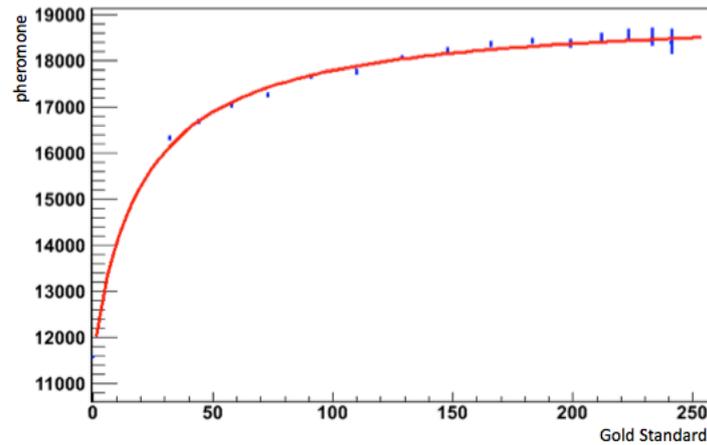


Figure 4: The correlation between the average pheromone deposition and the percentage of white matter in the Gold Standard.

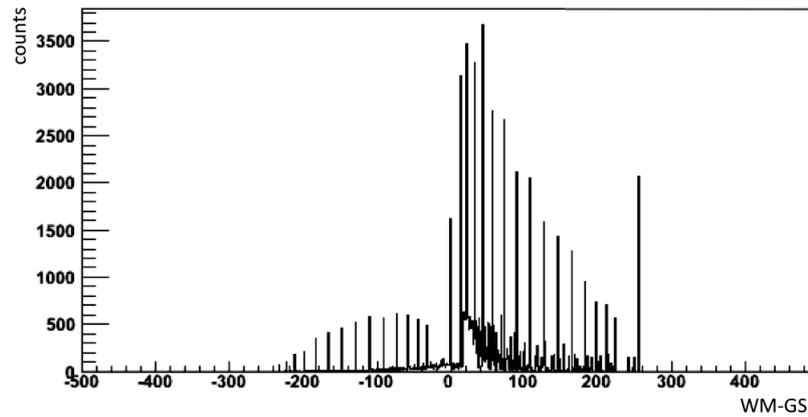


Figure 5: Difference between the CAM white matter estimate (WM) and the gold standard value (GS).

quantitative result, not just a binary answer.

In Figure 4 the average correlation between the released pheromone and the white matter amount, as defined by the gold standard, is shown. The red line is the result of an hyperbolic fit to the data with the following function:

$$GS = p0 + \frac{p1}{Pheromone + p2}$$

Starting from the fit result and the pheromone quantity deposit in each voxel, the CAM can estimate the white matter fraction. The difference between the CAM result and the gold standard value, shown in Figure 5, provides the resolution of the method.

For these voxels, the preliminary CAM result overestimates, on average, the white matter percentage, with a peak at about 10% and a RMS of about 20%.

8. Conclusions and Prospects

The most interesting feature of the CAM seems its adaptability: the results shown above were obtained with the same parameter values used for the analysis of artificial images, although the dynamic range of lung CTs and brain MRIs are remarkably different. In view of optimising the CAM performance, however, some fine tuning will likely be required.

Improvements in the analysis of lung CTs are also expected by introducing new relevant features in the neural network classification stage. The asymptotical values of the sensitivity, however, are comparable to the best results in literature, with the exception of the GS20 database. Therefore, the focus will be more on better defining the boundaries and features of the findings, so as to reduce the number of FP/scan at a given sensitivity.

Concerning the brain MRI analysis, which was started very recently, the source of the systematic error will be investigated. Moreover, it is planned to carry out a double colony deployment on the same image, with two different pheromone deposition rules: in addition to the one used to segment the white matter, the

$$\psi(I) = I_{max} - I$$

rule will segment the MRI volume corresponding to the cerebro-spinal fluid (CSF). The grey matter fraction will be obtained by complementing to 1 the values of each voxel for the white matter and the CSF. The optimisation on simulated images will be followed by the analysis of real brain MRI.

References

- [1] P. Cerello, "The MAGIC-5 Project: Medical Applications on a Grid Infrastructure Connection", *Int. J. Sci. Res.*, Vol 16(2006) 413-417.
- [2] P. Cerello *et al*, "GPCALMA: a Grid-based tool for Mammographic Screening" *Methods of Information in Medicine* vol. 44 (2005)244-248.
- [3] R. Warren *et al*, "MammoGrid - a prototype distributed mammographic database for Europe", *Clinical Radiology*, Volume 62, Issue 11, Pages 1044-1051.
- [4] R. Bellotti *et al*, "Distributed Medical Images Analysis on a Grid Infrastructure", *Future Generation in Grid System*, Special Issue on Life Science Grids for Biomedicine and Bioinformatics, Vol. 23/3, 2006, pp. 475-484.
- [5] A.Jemal, T.Murray, E.Ward, A.Samuels, R.C.Tiwari, A.Ghafoor, E.J.Feuer, M.J. Thun *CA Cancer J.Clin.*55(2005)10.
- [6] A.Micheli, P.Baili, M.Quinn, E.Mugno, R.Capocaccia, P.Grosclaude, *Ann. Oncol.* 14(5)(2003)V28.
- [7] G.K. Singh, B.A. Miller, B.F. Hankey, *J. Natl. Cancer Inst.* 94(2002)916.
- [8] S.Diederich, M.G.Lentschig, T.R.Overbeck, D.Wormanns, W.Heindel, *Eur. Radiol.* 11(2001)1345.
- [9] S. Itoh, M. Ikeda, S. Arahata, T. Kodaira, T. Isomura, T. Kato, K. Yamakawa, K. Maruyama, T. Ishigaki, *Radiology* 215(2000)175.
- [10] C.I. Henschke, D.I. McCauley, D.F. Yankelevitz, D.P. Naidich, G. McGuinness, O.S. Miettinen, D.M. Libby, M.W. Pasmantier, J. Koizumi, N.K. Altorki, J.P. Smith, *The Lancet* 354(9173)(1999)99.
- [11] S.J. Swensen, *BMJ* 326(2003)894.

- [12] E.F. Patz, S.J. Swensen, J.E. Herndon, *J. Clin. Oncol.* 22(2004)2202.
- [13] K. Kantarci and C. R. Jack, "Quantitative magnetic resonance techniques as surrogate markers of Alzheimer's disease", *NeuroRx: The Journal of the American Society for Experimental NeuroTherapeutics* 112(2004)196-205.
- [14] M. Atiya, B. T. Hyman, M. S. Albert, and R. Killiany, "Structural magnetic resonance imaging in established and prodromal Alzheimer disease: A review", *Alzheimer Dis. Assoc. Disord.* 173(2003)177-195.
- [15] G. Chetelat and J.-C. Baron, "Early diagnosis of Alzheimer's disease: Contribution of structural neuroimaging", *Neuroimage* 182(2003)525-541.
- [16] K. Kantarci and C. R. Jack, "Neuroimaging in Alzheimer disease: An evidence-based review", *Neuroimaging Clin. N. Am.* 132(2003)197-209.
- [17] V. C. Anderson, Z. N. Litvack, and J. A. Kaye, "Magnetic resonance approaches to brain aging and Alzheimer disease-associated neuropathology", *Top. Magn. Reson Imaging* 166(2005)439-452.
- [18] P. Cerello *et al.*, "3-D Object Segmentation using Ant Colonies", *Pattern Recognition* 43 (2010), pp. 1476-1490.
- [19] D. Chialvo, M. Millonas, "How swarms build cognitive maps", *The Biology and Technology of Intelligent Autonomous Agents* 144 (1995) 439-450.
- [20] V. Ramos, F. Almeida, "Artificial ant colonies in digital image habitats - a mass behaviour effect study on pattern recognition", in: *Proceedings of ANTS 2000-2nd International Workshop on Ant Algorithms (From Ant Colonies to Artificial Ants)*, Belgium, 2000, pp. 39-43.
- [21] E.F. Moore, "Gedanken experiments on sequential machines", *Automata Studies*, 1956.
- [22] V. Ramos, C. Fernandes, A.C. Rosa, "Varying the population size of artificial foraging swarms on time varying landscapes", in: *Proceedings of the ICANN05: 15th International Conference, Lecture Notes in Computer Science*, vol. 3696, Springer, Berlin, 2005, pp. 311-316.
- [23] <http://anode09.isi.uu.nlS>.
- [24] <http://imaging.cancer.gov/programsandresources/InformationSystems/LIDC>.
- [25] G. De Nunzio *et al.*, "Automatic Lung Segmentation in CT Images with Accurate Handling of the Hilar Region", *Journal of Digital Imaging*, DOI:10.1007/s10278-009-9229-1.
- [26] <http://www.bic.mni.mcgillca/brainweb>.
- [27] Huang *et al.*, "An artificial ant colonies approach to medical image segmentation", *Computer Methods and Programs in Biomedicine* 92 (2008) 267-273.