

Overview

Computational analysis and decision support systems in oncology

FILLIA MAKEDON¹, VANGELIS KARKALETSIS² and ILIAS MAGLOGIANNIS³

¹Dartmouth College, Department of Computer Science, Hanover, NH 03755, USA; ²National Center for Scientific Research 'Demokritos,' Institute of Informatics and Telecommunications, 15310 Aghia Paraskevi, Athens; ³University of Aegean, Department of Information and Communication Systems Engineering, 83200 Karlovasi, Samos, Greece

Received September 6, 2005; Accepted October 26, 2005

Abstract. Computational analysis tools and decision support systems have increased their penetration in the support of clinical processes and management of medical data and knowledge. Applications range from adjunct tools for diagnosis and disease investigation to the treatment and monitoring of therapeutic procedures. As all medical fields, the field of oncology is affected. This special issue includes studies presenting research and applications of computational intelligence in oncology, covering four main areas: i) decision support systems (DSS) and artificial intelligence (AI) applications in oncology; ii) design and assessment of classification tools in oncology; iii) intelligent accessing, retrieving, and storing of medical images; and iv) intelligent telemedicine and telehealth applications in oncology.

Introduction

There has been a significant increase in the level of interest regarding the development and use of information technology tools in medicine and healthcare provision. As the computer-based patient record system expands to support more clinical activities, healthcare organizations are asking physicians and nurses to interact more often with computer systems to perform their duties. Several studies exist on the impact of information technology on the life of patients, practitioners, and researchers in the field.

One major area of accelerated growth is the field of computational analysis and artificial intelligence where much research has been performed to develop and design diagnostic tools to support the work of medical professionals. Expert systems and machine learning algorithms are used to provide second opinions in diagnosis, while intelligent telemedicine and telehealth applications cover the need of

constant medical supervision of inhabitants in remote, isolated and underserved locations. Medical and biomedical informatics utilizing data mining, computational analysis, visualization, and the simulation of data and experiments have advanced the discovery of new knowledge regarding biological and therapeutic procedures.

This special issue is the result of a session organized in the context of the 9th World Congress on Advances in Oncology (http://www.iit.demokritos.gr/skel/WCAO2004_workshop/). With the goal of illustrating recent advances in computational analysis and decision support systems in oncology, and taking into account the relevant fields, we included studies that present novel computational analysis techniques, methods, practices and systems.

Decision support systems and artificial intelligence

Anagnostopoulos *et al* undertook the problem of breast cancer diagnosis and prognosis by employing two neural network architectures over the Wisconsin Diagnostic and Prognostic Breast Cancer (WDBC/WPBC) datasets (1). A probabilistic approach is dedicated to solve the diagnosis problem by detecting malignancy among instances derived from fine needle aspirate (FNA) tests, while the second architecture estimates the time interval that possibly contains the right endpoint of the disease-free survival (DFS) time of the patient. The study also tests the robustness and real-time response of proposed classifiers over the web as a potential integrated web-based decision support system.

Argyropoulos *et al* propose a method for the normalization of raw microarray data that allows gene expression profiling at a global level (2). The study presents the development of a set of operational criteria for assessing the impact of various normalization algorithms in terms of accuracy (bias), precision (variance) and over-fitting (information reduction). The use of these criteria is illustrated by applying the three most widely used algorithms of global median normalization, spiked-in based normalization and lowess on a specifically designed, multi-controlled dataset.

Magoulas examines the potential of neuronal networks and textural feature extraction for recognizing suspicious regions in endoscopy (3). As the author notes, the experimental results provide numerical evidence that neuronal networks

Correspondence to: Dr Ilias Maglogiannis, Department of Information and Communication Systems Engineering, University of Aegean, 83200 Karlovasi, Greece
E-mail: imaglo@aegean.gr

Key words: computational analysis, decision support systems, diagnostic tool

are capable of classifying offline and online tissue samples extracted from standard images and VHS videotape recordings of colonoscopy procedures with satisfactory success. This type of technology may potentially be useful in assisting medical experts to automate minimally invasive diagnostic procedures.

Dounias *et al* compares various intelligent methodologies in the task of Pap smear diagnosis (4). The methodologies used for the construction of Pap smear classifiers are different clustering approaches, feature selection, neuro-fuzzy systems, inductive machine learning, genetic programming, and second order neural networks. The acquired results are discussed with respect to their comprehensibility and usefulness to medical staff, either for fault diagnosis tasks or the construction of automated computer-assisted classification of smears.

Thireou *et al* utilize independent component analysis (ICA), a statistical method used for feature extraction and signal separation, for dynamic PET studies to facilitate the initial interpretation and visual analysis of these large image sequences (5). ICA produces parametric images, where structures with different kinetic characteristics are assigned opposite values and readily discriminated, which improves the process of identifying lesions and facilitates posterior detailed kinetic analysis.

Tsakonas *et al* propose a computer-assisted methodology for the effective modelling of the diagnostic decision for breast tumor malignancy (6). The suggested computational intelligence approach is a hybrid that combines modern mathematical logic principles, neural computation and genetic programming. As the authors note, the approach proves promising either in terms of diagnostic accuracy and generalization capabilities, or comprehensibility and practical importance for the related medical staff.

Abbod *et al* compare the accuracies of neuro-fuzzy modelling (NFM), artificial neural networks (ANN) and traditional statistical methods (logistic regression, LR) for the prediction of bladder cancer (7). For all three methods, models were produced to predict the presence and timing of tumor progression. According to their experimental results, NFM can predict tumor behavior with greater accuracy than both ANN and LR. The features of NFM make it an important tool, with extensive clinical application not only in bladder cancer, but also in other human malignancies.

Design and assessment of classification tools

Cochand-Priollet *et al* compare various classifiers in the discrimination of benign from malignant thyroid lesions (8). Their study was performed on May Grunvald Giemsa-stained smears taken by fine needle aspiration (FNA). Using a custom image analysis system, 25 nuclear features describing the size, shape and texture of the nuclei were measured in each case. A portion of cases were used for the training of classifiers, and classification was performed at the nuclear and patient level. Their approach produces encouraging results and shows promise as a helpful tool in the daily cytological laboratory routine.

Maglogiannis *et al* review computational vision-based diagnostic systems for the detection of malignant melanoma (9). This study details the visual features used for skin lesion classification and how to extract these features through digital

image processing methods, i.e. segmentation, registration, border detection, and color and texture processing. It also discusses artificial intelligence methods, i.e. discriminant analysis, neural networks, support vector machines employed for skin lesion classification, and compares these techniques in discriminating malignant melanoma tumors from dysplastic naevi lesions.

Megalopoulou *et al* investigate the potential value of morphometry, feature selection and statistical classifier techniques, such as neural networks, for distinguishing benign from malignant gastric nuclei (10). According to their approach, a set of gastric smears are routinely processed and stained using the Papanicolaou technique, then analyzed with a customized image analysis system. A feature selection technique was first applied to identify the most important nuclear features, which were used in the second stage by statistical classifiers to classify a nucleus as benign or malignant. The authors discuss the experimental results, which they note are significant at the nuclear and patient level.

Panourgias *et al* discuss the findings obtained from the application of two computational intelligence methodologies for the detection of microcalcifications in screening mammography data (11). Genetic programming and inductive machine learning are applied to produce meaningful diagnostic rules for the medical staff. The diagnostic rules obtained from both methods are evaluated in detail, then analyzed and discussed by collaborative medical experts in parallel to findings from related literature. As the authors note, the results appear encouraging for further use and analysis by medical staff specializing in screening mammography.

Reif *et al* describe the exploratory visual analysis (EVA) software and database, which is freely available to academic users, to explore data analysis results in the context of biological information about each gene available in public databases such as Entrez Gene (12). Using a publicly available microarray dataset, the authors demonstrate the usefulness of the EVA system in interpreting statistical results of gene expression response to chemotherapeutic agents in human breast cancer cell lines.

Sakka *et al* review and evaluate some of the classification algorithms on microcalcifications in mammograms used in various CAD systems to improve breast cancer prognosis (13). The algorithms are separated into categories according to the method in use. Neural networks are used in applications where only a few decisions are required for an amount of data, while the k-nearest neighbor classifier distinguishes unknown patterns based on the similarity to known samples. The decision tree approach is simpler than neural networks and does not require extensive knowledge of the probability distribution of features. The study also gives an overview of the currently available literature on the characterization of malignant and benign microcalcifications.

Wang *et al* examine the high-density single nucleotide polymorphism (SNP) array, which is a recently introduced technology that genotypes more than 10,000 human SNPs on a single array (14). It is shown that SNP arrays can be used to determine not only SNP genotype calls, but also DNA copy number (DCN) aberrations, which are common in solid tumors. The study addresses the feasibility of cancer classification

based on DCN aberrations, determined using SNP arrays, by applying state-of-the-art classification algorithms and feature selection algorithms to the DCN aberration data derived from a public SNP array dataset. The results indicate that DCN aberration data derived from SNP arrays is useful for etiology-based tumor classification.

Intelligent accessing, retrieving and storing of medical images

Arfanakis *et al* discuss the problem of accurate localization of white matter fiber tracts in relation to brain tumors (15). White matter fiber tractography by means of diffusion tensor magnetic resonance imaging (DTI) is the only non-invasive method that can provide estimates of brain connectivity. However, conventional tractography methods suffer from image distortions and artifacts. The authors evaluate fiber-tracking results obtained from Turboprop-DTI, a technique that provides images with significantly fewer distortions and artifacts than conventional methods. The results show that Turboprop may be a more appropriate technique for tracing white matter fibers than conventional DTI methods, especially in applications such as pre-surgical planning and image-guided surgery.

Astrakas *et al* propose the development of local research databases as an infrastructure for improved data management and detailed data analysis (16). As an example, they developed a pediatric brain tumor database of magnetic resonance imaging data, including conventional MRI imaging, hemodynamic MRI, diffusion weighted MRI and MR spectroscopic imaging, combined with neuropathological and neurological evaluation data. The authors claim that such a database will significantly enhance the assessment and treatment of pediatric brain tumor patients.

Athanasiadis *et al* discuss an integrated approach for the location of salient regions in a CT scan, focusing on the area of the liver (17). A number of image processing tasks are performed in successive scans to extract areas with a texture different than that of the greater part of the organ. In general, these areas do not always correspond to pathological patterns, but may be the result of noise in the scanned image or related to veins passing through the tissue. The result of the algorithm is an original image with a mask indicating these regions, so the attention of the medical practitioner is drawn to them for further examination.

Demidenko presents a statistical method for the comparison of color cancer cell images, which is applied to discriminate cancer histology images from four treatment groups (18). The study describes a well-established multivariate analysis of variances (MANOVA) on the logit scale in which image comparison is reduced to the classical Chi-square test.

Huang *et al* discuss the assessment of tumor angiogenesis by perfusion-sensitive MRI, a method for early and accurate identification of lung cancer (19). Compensating for the motion of respiratory and anatomical structures is a challenge in this method. To use perfusion MRI in a signal imaging session to define the metabolic and vascular characteristics of tumors, the authors examine a novel method that can register points of interest by tracking the image intensity changes around the target point. The registration results are used to

generate accurate time-intensity curves of different regions of interest. The authors note that their method is computationally efficient and performs well in perfusion MRI sequence analysis.

Shen *et al* utilize perfusion magnetic resonance imaging (pMRI), an important tool in assessing tumor angiogenesis for the early detection of lung cancer (20). Their study demonstrates a novel integrated framework for spatio-temporal modeling of pulmonary nodules in pMRI image sequences. The authors describe the segmentation procedure in the region to extract the nodule boundary, then use thin-plate spline interpolation for nodule registration along the temporal dimension. The resulting spatio-temporal model can lead to many types of nodule characterization, e.g. a time-intensity profile of a nodule region, and be used to capture important angiogenic patterns in the lung that can distinguish between cancer and benign nodules and assist early detection.

Intelligent telemedicine and telehealth applications in oncology

Kontaxakis *et al* describe the TENPET (trans-European network for positron emission tomography) system, which aims to the provision of integrated teleconsultation and intelligent computer-supported cooperative work services for clinical positron emission tomography (PET) studies (21). The study addresses technological challenges by linking PET Centres and developing support services that permit remote consultation between professionals in the field. TENPET, the authors claim, is the first telemedicine service specifically designed for oncological PET. Technical innovations incorporated in the TENPET platform and the initial pilot studies at real and diverse clinical environments in the field of oncology are presented.

Karacapilides *et al* present a web-based system that aims at the exploitation of personalized knowledge through a structured process of collaborative and argumentative resolution of the issues raised in such settings (22). By using an illustrative example concerning collaboration between a group of experts deciding on the most appropriate treatment for a particular patient with breast cancer, the authors present the design rationale, features and functionality of their system. The study draws attention to associated knowledge management issues.

We hope that the studies included in this special issue will assist the medical audience in learning about diagnostic tools, expert systems and machine learning algorithms designed to support medical decision making and provide second opinions in diagnosis, while contributing to the roadmap of the emerging field of computational analysis and artificial intelligence in medicine.

References

1. Anagnostopoulos I, Anagnostopoulos C, Vergados D, Rouskas A and Kormentzas G: The Wisconsin Breast Cancer Problem: diagnosis and TTR/DFS time prognosis using probabilistic and generalised regression neural classifiers. *Oncol Rep* 15: 975-981, 2006.
2. Argyropoulos C, Chatziioannou AA, Nikiforidis G, Moustakas A, Kollias G and Aidinis V: Operational criteria for selecting a cDNA microarray data normalization algorithm. *Oncol Rep* 15: 983-996, 2006.

3. Magoulas G: Neuronal networks and textural descriptors for automated tissue classification in endoscopy. *Oncol Rep* 15: 997-1000, 2006.
4. Dounias G, Bjerregaard B, Jantzen J, Tsakonas A, Ampazis N, Panagi G and Panourgias E: Automated identification of cancerous smears using various competitive intelligent techniques. *Oncol Rep* 15: 1001-1006, 2006.
5. Thireou T, Pavlopoulos S, Kontaxakis G and Santos A: Blind source separation for the computational analysis of dynamic oncological PET studies. *Oncol Rep* 15: 1007-1012, 2006.
6. Tsakonas A, Dounias G, Panagi G and Panourgias E: Breast tumor malignancy modelling using evolutionary neural logic networks. *Oncol Rep* 15: 1013-1017, 2006.
7. Abbod MF, Linkens DA, Catto JWF and Hamdy FC: Comparative study of intelligent models for the prediction of bladder cancer progression. *Oncol Rep* 15: 1019-1022, 2006.
8. Cochand-Priollet B, Koutroumbas K, Megalopoulou TM, Pouliakis A, Sivolapenko G and Karakitsos P: Discriminating benign from malignant thyroid lesions using artificial intelligence and statistical selection of morphometric features. *Oncol Rep* 15: 1023-1026, 2006.
9. Maglogiannis I and Kosmopoulos D: Computational vision systems for the detection of malignant melanoma. *Oncol Rep* 15: 1027-1032, 2006.
10. Megalopoulou TM, Koutroumbas K, Pouliakis A, Sivolapenko G and Karakitsos P: The potential of feature selection by statistical techniques and the use of statistical classifiers in the discrimination of benign from malignant gastric lesions. *Oncol Rep* 15: 1033-1036, 2006.
11. Panourgias E, Tsakonas A, Dounias G and Panagi G: Computational intelligence for the detection and classification of malignant lesions in screening mammography. *Oncol Rep* 15: 1037-1041, 2006.
12. Reif DM and Moore JH: Visual analysis of statistical results from microarray studies of human breast cancer. *Oncol Rep* 15: 1043-1047, 2006.
13. Sakka E, Prentza A and Koutsouris D: Classification algorithms for microcalcifications in mammograms (Review). *Oncol Rep* 15: 1049-1055, 2006.
14. Wang Y, Makedon F and Pearlman J: Tumor classification based on DNA copy number aberrations determined using SNP arrays. *Oncol Rep* 15: 1057-1059, 2006.
15. Arfanakis K, Gui M and Lazar M: Optimization of white matter tractography for pre-surgical planning and image-guided surgery. *Oncol Rep* 15: 1061-1064, 2006.
16. Astrakas L, Ye S, Zarifi M, Makedon F and Aria Tzika A: The clinical perspective of large scale projects: a case study of multiparametric MR imaging of pediatric brain tumors. *Oncol Rep* 15: 1065-1069, 2006.
17. Athanasiadis T, Wallace M, Karpouzis K, Nikolakopoulos Y and Kollias S: Utilization of evidence theory in the detection of salient regions in successive CT images. *Oncol Rep* 15: 1071-1076, 2006.
18. Demidenko E: Statistical comparison of color cancer cell images. *Oncol Rep* 15: 1077-1079, 2006.
19. Huang H, Shen L, Ford J, Gao L and Pearlman J: Early lung cancer detection based on registered perfusion MRI. *Oncol Rep* 15: 1081-1084, 2006.
20. Shen L, Zheng W, Gao L, Huang H, Makedon F and Pearlman J: Spatio-temporal modeling of lung images for cancer detection. *Oncol Rep* 15: 1085-1089, 2006.
21. Kontaxakis G, Visvikis D, Ohl R, Sachpazidis I, Suarez JP, Selby P, Cheze-Le Rest C, Santos A, Ortega F, Diaz J, Pan L, Strauss L, Dimitrakopoulou-Strauss A, Sakas G and Pozo MA: Integrated telemedicine applications and services for oncological positron emission tomography. *Oncol Rep* 15: 1091-1100, 2006.
22. Karacapilidis N and Koukouras D: A web-based system for supporting collaboration towards resolving oncology issues. *Oncol Rep* 15: 1101-1107, 2006.