Medical Informatics Europe 2005 Workshop on Medical Image Processing

Henning Müller[1], Alexander Horsch[2], Thomas Lehmann[3], Andreas Schmidt[4]
[1]*Service of Medical Informatics, University and Hospitals of Geneva, Switzerland*
[2]*Munich University of Technology, Munich, Germany*
[3]*Medical Informatics, RWTH Aachen, Aachen, Germany*
[4]*Novartis Pharma AG, Basel, Switzerland*

Abstract

The workshop on medical image processing was organized and supported by the EFMI (European Federation for Medical informatics) working group on Medical Image Processing (MIP). Title of the workshop was "Medical Image Management" with the goal to talk about a variety of aspects concerning the management of medical images. Presentations included the industry perspective of medical image management of a large pharmaceutical company, the need for reference image databases for image processing tasks, content-based access methods to medical images, and the analysis of user behavior concerning the access and use of medical images.
The need for systematic management of images and the retrieval of medical images were highlighted. These two fields are the basis for the better use of images and the extraction of knowledge inherently stored in the images and the context in which they were taken. Image reference databases in the medical field were seen as the only way to properly evaluate medical image processing algorithms.

Introduction

The EFMI Working Group (WG) on Medical Image Processing (MIP[1]) that organized the workshop supports the mission of EFMI by fostering the discussion and activities in the member countries about theory and practice of medical image processing. This comprises, in particular, the discussion of how to integrate decision support by means of medical image processing into clinical practice, including the important topics of clinical evaluation, standardization and technology transfer.
The WG is organizing several meetings per year to address specific needs of the WG members and to support its goals:

- The working group will establish a reference image database for medical image processing R&D groups within the EFMI member countries. The images shall be used in order to make the capabilities of different methods comparable.
- In order to foster cooperation between R&D groups and to avoid redundant development, the working group will establish a Web-based information system on European image processing groups and their current activities. The working group will provide the technical infrastructure. The R&D groups enter and maintain their own data record.
- The working group will create and maintain a Website, which will provide all information about the work and upcoming events relevant for interested colleagues and the public.

Prior workshops were held on the integration of medical imaging into clinical workflow (St. Malo, 2003), biomedical informatics for clinical decision support (Bethesda, 2004), and medical reference databases (Berlin, 2005). The workshop at MIE in Geneva had the clear goal to combine the experience from the past workshops into new ideas for the management of medical image data on several levels and from several viewpoints.

Presentations at the workshop

The EFMI reference image database initiative

Research and approvals depend on evidence. Evidence in complex domains like medicine, pharmacy and medical image processing cannot be created without appropriate evaluation methods and validation platforms serving as environments for testing the performance of methods and systems in terms of absolute measurements (benchmarking) and by comparison with known and proven methods. In medical image processing a non-trivial problem exists with respect to validation environments. The development of new methods is typically based on images taken from one or few image acquisition

[1] http://www.efmi-wg-mip.net/

units. The algorithms tend to be optimized for these machines. Research groups use very different and usually incompatible datasets, which prevents comparability of new methods. Image datasets obtained within only one research center do never represent the medical variety desirable for sound clinical studies. In academic medical image processing research, emphasis is usually put on innovation in terms of algorithmic novelty. Instead of sound validation and evaluation on clinically relevant data, only small feasibility studies are conducted. Industry has problems with the acceptance of image processing applications as automation tools in approval procedures by the certification authorities.

In 2002, the WG MIP started an initiative with the aim to trigger the establishment of a reference image database for medical image processing research to support validation and comparability of methods. Close contact exists with other initiatives, especially with the NIH and the Insight Software Consortium, as well as with industry in order to make the concept practicable. The concept of the EFMI reference image database initiative consists of the following main points [1-4]:

- Creation of a sustainable framework for reference image datasets and tools for validation and quality control of academia and industry.
- Establishment of a board of experts to define criteria on relevance assessment of medical problems with respect to the importance of image processing.
- Use of the defined criteria to assess medical problems and identify the most relevant ones with a high potential for digital image processing methods.
- Specification of image needs for scientific validation and evaluation, including quality criteria and standardization of data structures for annotations (Gold standards).
- Collection of images from providers to create validated datasets to serve as references for research and development groups.
- Platform creation for dissemination of reference datasets, including copyright agreements between providers and users.
- Impact assessment in terms of outcome indicators such as number and quality of published results, or costs and time for approval processes using the datasets.

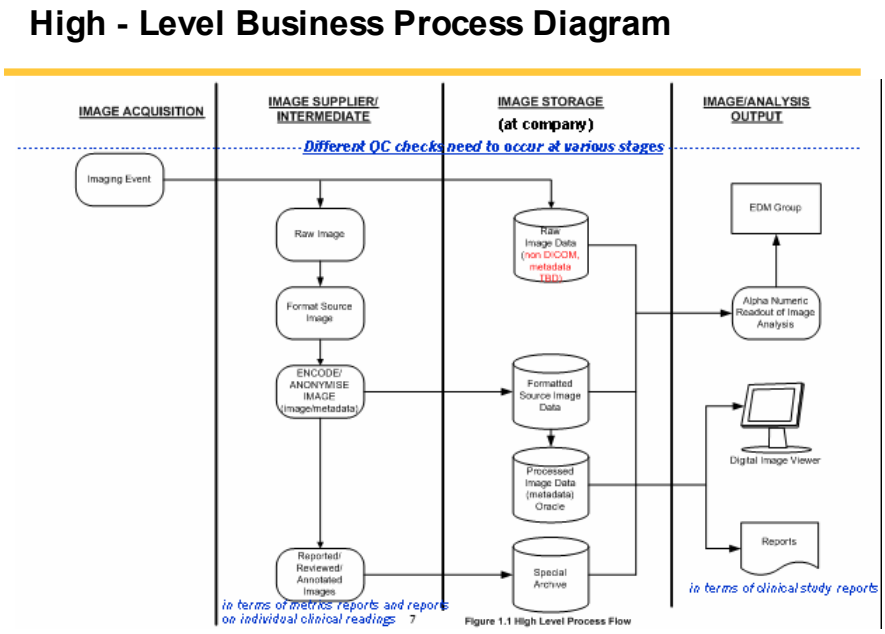
Conceptual and promotional work has been done by the WG since 2002. Awareness of the usefulness and necessity of the initiative exist but the concrete commitment is still limited. Academic institutions do not have the resources to set up such a framework and therefore they have to focus on the outcomes of their own research projects including image data acquisition and management. Industry concentrates on procedures required by the regulatory authorities and struggles with the threatening of their economical benefits by long development and approval cycles. Both sides would benefit from an approved common platform for validation. Since setting up such a platform needs efforts from the public and industry, the WG will form a strong alliance of academia and industry to create a sustainable platform to be implemented in a public-private effort.

Managing Medical Image Data in a Pharmaceutical company

The need for pharmaceutical companies to manage their medical image data sets generated in clinical and preclinical studies is increasing. Any Pharma image management (IM) system supports the loading, aggregating, viewing, storing, and reporting of processed digital image data sets, related annotations, measurements and metadata. The business process can be described by the following components: Clinical Image Acquisition/Storage, Clinical Image Viewing, Clinical Image Analysis, Clinical Image Data Export to Electronic Data Management (EDM), and Exception and Error handling. The goals of image data analysis include disease detection for patient inclusion into a studies, disease modification in response to drug treatment, correlation of images with clinical data, and clinical endpoints. Retrospective data mining can be used to identify patterns and model efforts for outcome prediction. Multiple challenges in data management occur such as a lack of harmonization of image acquisition, infrastructure problems in processing and storing large data volumes, lacking of diagnostic aid including images, and the lack of content-based data access.

Key functions for image management include import/processing of DICOM images and metadata as well as quality control, population of an image repository with images metadata and analytical results, provision of an image viewer, export of data, and query tools. Key differences between Pharma IM versus a PACS (Picture Archival and Communication System) system are related to the workflow, user groups, and regulatory compliance in drug development PACS may not need an automatic QC

functionality or an export capability. On the other hand, Pharma IMs do not need dedicated workstations for the users or the RIS workflow and a high level of system redundancy.



Currently, no product satisfies all needs for a Pharma IM system. Vendors concentrate rather on the PACS/RIS market and closer communication with the Pharma industry is needed.

Automatic classification and content-based retrieval of medical images

Automatic classification describes the decision to assign semantic concepts to an image. Applications for automatic classification are manifold in radiology, for PACS, for computer-aided diagnosis (CAD), and for content-based image retrieval (CBIR).

Usually, images are represented by a few numerical values describing features such as color, texture, and shape. Several publications exist using this concept to distinguish a small number of classes. Arimura et al. report a correctness of about 100 % when determining the view detection on chest radiographs [5]. Similarly, Pinhas & Greenspan reached 99 % correctness for automatic identification of body regions [6]. However, the question is whether such global features are still applicable for large number of categories. In a recent study, Güld et al. reported 85% correctness when categorizing 6283 images into 81 classes covering anatomy, biosystem, imaging modality, and direction between patient and imaging device [7]. A similar problem is currently under investigation in the ImageCLEFmed 2005 competition [8]. Here, a set of 9.000 training images were released to 24 research groups from 15 nations worldwide (12 submitted results). Training images were from 57 classes. 1.000 images without class identification where used to evaluate the different approaches. The best algorithm reached a correctness of 87,4 %.

The study of Güld et al. shows that the amount of training data in each category is important for the classification result. Classes represented by hundreds of reference images reach a recognition rate of more than 99 %. For classes with only 5 reference images the recognition rate can drop to 0 %. Since global features are used for categorization, collimation fields and shutters are difficult to handle. Some classes have high intra class variability and others (such as elbow and knee) show low inter class variability. More details on this can be found on the web page of the IRMA[2] project.

Analysis of the use of medical images

Before evaluating medical image retrieval systems [9] or image management solutions it is important to figure out the real needs of users. For the context of medical image information retrieval we conducted a survey among medical professionals at OHSU in Portland, Oregon [10] and the Geneva University hospitals. Goal was to find out important image information needs for the context of the ImageCLEFmed image retrieval evaluation campaign [8,11] and to advance the medGIFT[3] project. A total of 18 participants were interviewed in Geneva and 13 in Portland. Questions of the survey

[2] http://www.irma-project.org/§

[3] http://www.sim.hcuge.ch/medgift/

include a list of *tasks* where images are useful, the *types* of images that are most common search for the task, *where* and *how* the image are search for and how *relevance* is judged. The interviewed persons could also tell us about search methods that would be useful for them but do not exist as of yet. Results were collected separately for the functions as clinician, researcher, lecturer, student and librarian, where several persons had more than one function.
Results show that search is along the three axes related to patient care, research (presentations, etc.) and teaching. Researchers often need illustrations and they need as well representative as abnormal images. Search is mostly by text in the PACS and sometimes at google. Clinicians basically search directly in the patient record or sometimes in the PACS and everything is patient care-related, so by patient name, and mostly x-ray, followed by CT images. Many clinicians for like to search for similar or the same pathologies and also visual search was mentioned. Lecturers often look for illustrations of a certain system and scans of schemas in books. Search is mostly by text and very frequently on the web via google or specialized sites for teaching. Search for visually similar images was mentioned as a very useful extension. In general the quality is judged based on personal experience, but in most cases quality could not be influenced.
Little is known about image use and image search behavior of medical professionals and before developing complex systems for diagnostic aid, it seems important to find out which parts of the image use process could be enhanced.

Future directions and ideas

The WG concluded in the session that the management of images in medical institutions (including hospitals and the pharmaceutical industry) is still neglected. Much knowledge is store in images and its metadata such as annotations and marked regions of interest. To avoid the double assessment of ground truth for certain tasks and widen the spectrum of image processing and analysis systems to more than a single imaging unit, the creation of reference databases is necessary. The WG will prepare the creation of an infrastructure to manage such reference images for several tasks that still need to be defined. Image retrieval, particularly based on the visual content was also regarded as a key component to better manage the current, large image repositories in medical institutions.

References

[1] A. Horsch, M. Prinz, S. Schneider, et al., Establishing an International Reference Image Database for Research and Development in Medical Image Processing. *Methods of Information in Medicine* 43 (2004), pages 409-12.
[2] M. Prinz, A. Horsch, S. Schneider, et al., A Reference Image Database for Medical Image Processing. *Proceedings of the 2nd Conference of the Österreichische Wissenschaftliche Gesellschaft für Telemedizin, a-telmed 2002, OCG-Schriftenreihe*, Vienna, 2002, pages 45-51.
[3] A. Horsch, R. Thurmayr, How to Identify and Assess Tasks and Challenges of Medical Image Processing. *Studies in Health Technology Informatics* 95 (2003), pages 281-285.
[4] A. Horsch, T. Wittenberg, K. Spinnler, Concept and Roadmap for Establishing an International Reference Image Database for Medical Image Processing R&D Groups. *The Sixth Korea-Germany Joint Workshop on Advanced Medical Image Processing*, 2002, Heidelberg, Germany.
[5] H. Arimura, S. Katsuragawa, T. Ishida, et al, Performance evaluation of an advanced method for automated identification of view positions of chest radiographs by use of a large database. *Proceedings SPIE Medical Imaging 2002, 4684*, pages 308-315, 2002.
[6] A. T. Pinhas, H. Greenspan, A continuous and probabilistic framework for medical image representation and categorization, *Proceedings SPIE Medical Imaging 2004, 5371*, pages 230-238.
[7] M. O. Güld, D. Keysers, M. Leisten, H. Schubert, T. M. Lehmann, Comparison of global features for categorization of medical images. *Proceedings SPIE Medical Imaging 2004, 5371*, pages 211-222, 2004.
[8] P. Clough, H. Müller, T. Deselaers, M. Grubinger, T. Lehmann, J. Jensen, W. Hersh, The CLEF 2005 Cross-Language Image Retrieval Track, *Springer Lecture Notes in Computer Science, Proceedings of the 2005 CLEF workshop*, Vienna, 2005 – to appear.
[9] H. Müller, N. Michoux, D. Bandon, A. Geissbuhler, A review of content-based image retrieval systems in medicine – clinical benefits and future directions, *International Journal of Medical Informatics*, volume 73, pages 1-23, 2004.
[10] W. Hersh, H. Müller, P. Gorman, J. Jensen, Task Analysis for Evaluating Image Retrieval Systems in the ImageCLEF Biomedical Image Retrieval Task, *Slice of Life, conference on multimedia in medical education*, Portland, OR, USA, 2005.
[11] H. Müller, A. Rosset, J.-P. Vallée, F. Terrier, A. Geissbuhler, A reference data set for the evaluation of medical image retrieval systems, *Journal on Computerized Medical Imaging and Graphics*, volume 28, pages 295-305, 2004.