Biomedical Image Processing Integration Through INBIOMED: A Web Services-Based Platform

David Pérez del Rey[1], José Crespo[1], Alberto Anguita[1], Juan Luis Pérez Ordóñez[2], Julián Dorado[2], Gloria Bueno[3], Vicente Feliú[3], Antonio Estruch[4], and José Antonio Heredia[4]

[1] Biomedical Informatics Group, Artificial Intelligence Department, School of Computer Science, Universidad Politécnica de Madrid Campus de Montegancedo, s/n. 28660 Boadilla del Monte, Madrid
dperez@infomed.dia.fi.upm.es
[2] RNASA-IMEDIR (Lab. Redes de Neuronas Artificiales y Sistemas Adaptativos Centro de Informática Médica y Diagnóstico Radiológico).Universidade da Coruña
julian@udc.es
[3] UCLM-ISA, E.T.S.I. Industriales, Ingeniería de Sistema y Automática Universidad de Castilla-La Mancha – Av. Camilo José Cela, s/n – 13071 Ciudad Real
gloria.bueno@uclm.es
[4] Departamento de Tecnología, Universitat Jaume I, Castellón

Abstract. New biomedical technologies need to be integrated for research on complex diseases. It is necessary to combine and analyze information coming from different sources: genetic-molecular, clinical data and environmental risks. This paper presents the work carried on by the INBIOMED research network within the field of biomedical image analysis. The overall objective is to respond to the growing demand of advanced information processing methods for: developing analysis tools, creating knowledge structure and validating them in pharmacogenetics, epidemiology, molecular and image based diagnosis research environments. All the image processing tools and data are integrated and work within a web services-based application, the so called INBIOMED platform. Finally, several biomedical research labs offered real data and validate the network tools and methods in the most prevalent pathologies: cancer, cardiovascular and neurological. This work provides a unique biomedical information processing platform, open to the incorporation of data coming from other feature disease networks.

1 Introduction

INBIOMED is a cooperative research network on biomedical informatics, founded by the "Fondo de Investigación Sanitaria" (FIS) of the Institute of Health Carlos III. The research network is integrated by 13 groups of biomedical research, belonging to 11 research centres. INBIOMED provides storage, integration and analysis of clinical, genetic, epidemiological and image data oriented to pathology research. Its main objective is to develop analysis tools and to evaluate them in their corresponding research environments [1].

This paper is centred on the image data treatment within INBIOMED, and how a new developed, and so called INBIOMED platform, has been used to tackle these

J.L. Oliveira et al. (Eds.): ISBMDA 2005, LNBI 3745, pp. 34–43, 2005.

challenges in the biomedical image processing field. This platform permits the integration and reuse of image processing methods developed by different INBIOMED partners the widely adopted technology of web services [2].

In the last years, the World Wide Web growth has been increasing more and more, with a big amount of research and efforts to create models and tools for new applications. A new application concept has appeared together with this growth, web services. Web services are applications executed in remote server computers and they may be accessed by clients from any computer located in other place. These users can access the web services through a minimum software client or even none, since many of these applications are accessible using a regular web browser. In the Internet there exist many applications available in different areas that may be accessed by users. It is not necessary to code again these applications when this functionality is required for a new application. Therefore, it is important to be able to reuse these services to be integrated in other applications (at a lower scale, this was already present in the object-oriented approach). These challenges are common to many working areas and also to Biomedicine where the INBIOMED project is enclosed.

Image processing methods are one of the main application fields within Biomedical Informatics. Biomedical image processing is widely needed by professionals and researchers and may require important computation capabilities to reach satisfactory results. Therefore, the use of distributed applications based on web services is especially appropriate for these functionalities. Firstly, these systems are used by professionals working in centres spread over the country, and they use computers with different operating systems, capabilities, etc. and web services overcome these heterogeneities. In addition it is easier to maintain the application if it is centralized than if each user have to update its software when a modification is introduced. Nowadays the research in the image processing field is intensive and new techniques or modifications of existing ones appear periodically. Adding new modifications or even new methods using web services is straight forward, providing the users with the most advanced techniques instantaneously.

2 Background

Image processing usually require powerful systems with expensive hardware, because of the huge computation resources needed. In addition, many image processing problems had been solved and implemented before. Distributed image processing provides a framework to overcome both problems. On the one hand, time consuming processes can be performed using remote and powerful servers in a transparent way. On the other hand, these technologies may supply a system architecture to reuse image processing methods already implemented.

A study of the main technologies that support distributed processing has been accomplished within this work. Table 1 presents a list of some of these technologies together with a brief description, references and some advantages and drawbacks.

Although GRID follows a different approach than the rest of the technologies listed in Table 1, it has been included in the study due to its widespread among distributed computing. While GRID issues a task to a set of different computers (any node may run any part of the process), the other technologies can not perform that load balancing in a transparent way for the user.

Table 1. Distributed processing technologies

Technology	Description	Advantages	Drawbacks
Java RMI (Remote Method Invocation)	Method invocation of remote java objects. Enables an object on one Java virtual machine to invoke methods on an object in another Java virtual machine [3]	Object oriented and very high level. Hides communication details. No need to learn a new IDL language [4], [5]	Constrained to java language, slow for complex tasks
CORBA (Common Object Request Broker Architecture)	Technology for creation of distributed object architectures. The core of its architecture is the Object Request Broker (ORB), the bus over which objects transparently interact with other objects [6] [7]	Object oriented. Allows communication between different operating systems and programming languages	You need to know the intermediate language of definition service (IDL). High complexity. Variations in vendor implementations.
DCOM (Distributed Component Object Model)	Distributed extension to COM (Component Object Model). Builds an object remote procedure call (ORPC) layer on top of DCE RPC [8] to support remote objects	Object oriented. Allows communication between different programming languages	Fully supported only in Microsoft Windows platforms. Architecture is not well-partitioned [9]
RPC (Remote Procedure Calls)	Protocol that enables programs to invoke procedures located in remote processes	Highly efficient	Not object oriented, low level. Different RPC systems can not communicate with each other [10]
Web Services	Service with a well defined and known interface that can be accessed through Internet [11]	Uses HTTP protocol for message transport. Communication based in XML and SOAP standards, avoids many incompatibilities	API is not fully standardized yet
Mobile Agents	Software entities capable of autonomous movement between hosts in a computer network [12]	Suitable for working with large amounts of data located in different places	Possible security issues. Code has to move to others computers. All computers involved must use the same language
GRID Computing	Distributed set of hardware and software resources that provided coordinated computing support [13]	Ideal for working with large amounts of data that require complex computing. Shows a set of computer as a single resource	Less useful working with small data. Requires complex software infrastructure. Big amount of bandwidth can be required [14]

Distributed image processing systems may be implemented using any of these technologies, because all of them allow remote communication between clients and processing methods in different computers – e.g. systems using CORBA [15], or Web-Based Medical Image Retrieval and Analysis systems using Mobile Agents [16].

In this paper, we present an integrated approach, based on Web Services, for distributed image processing within the INBIOMED network. The reasons for this approach are based on the heterogeneity and different features of image data, computational resources needed and functionalities that users may require for this kind of project. Using this model, different users can benefit from integrated image processing tools remotely located.

3 The INBIOMED Platform

The INBIOMED platform [17] is a web services oriented architecture to provide a framework for sharing resources such as data sources or data processing algorithms. These resources can be located at different nodes that may be accessed through web services published in the platform catalogs.

The philosophy of the platform is to minimize the load of processing in complex client applications to convert the algorithms in data processing web services. Consequently, the platform applications only have to implement the user interface with their input forms and the presentation results, delegating in the platform the responsibility to query, store and process the information.

The communication between the client applications and the INBIOMED platform must be accomplished using the SOAP protocol, so the applications have to be implemented using a programming language that is able to be published using web services. There are not any more restrictions imposed to the development of INBIOMED platform client applications.

Cooperative Model

Every INBIOMED network node collaborates with the other nodes in, at least, one of the following ways:

- Developing its own INBIOMED services
  In this way, it facilitates the sharing data access. Every node keeps its own site independently. Every node provides a user interface to the other nodes; that is used to send requests and will have the proper rights or privileges to access the data sources, data process service catalog and data repository.
  The client applications send all requests to the platform through this INBIOMED platform web service. This web service receives requests written in a language oriented to the batch execution called IQL (INBIOMED Query Language) [17]. The syntax of this language includes sentences for every function that can be performed within the Platform
  At this level is necessary to prepare the data to be accessible through the platform. The data warehouse technology plays a fundamental role, guaranteeing that the information stored in the databases is well consolidated and oriented to be queried, and is obtained from the internal or public data sources.
- Developing data processing services
  A node in the INBIOMED network, may not have information by itself, but can provide the development of a data processing web service implementing algorithms to analyze information for the other nodes.

The data processing web services are responsible to implement the methods necessary to analyze the data, process them and generate the results. The implementation of the algorithms can be made in the most suitable programming language for the resolution of the problem.

A common data processing web service needs to be built at the top of these algorithms to be included in the platform. The goal is to make available a wide catalog of services of this type that implement the processes needed for the different research groups.

- Developing client applications.

  Using the resources available in the platform it is possible to implement client applications that solve a problem. These client applications may also be used by other members of the INBIOMED network.

Therefore, the topology of the INBIOMED network can vary from models completely centralized – where there is only one web service in a node, and all the other nodes receive and send data from this node – to federated models, where every node keeps its own information independently.

Among all the applications and data treated by the INBIOMED network, biomedical images and their corresponding processing algorithms are the main topic of this work.

4 Type of Images

The types of data that the INBIOMED Platform handles are molecular, functional and genetic images. These images group into three different medical applications: Pharmacogenetics, Cancer Epidemiology, and Molecular and Genetic Epidemiology. Table 2 shows the relationship between the image data and the medical application.

The data were collected from the user groups and an entity-relationship model, explaining the structures and characteristic for each image was created An *in-situ* observation of the tasks carry on by these groups was done in order to identify the image processing functions required. To this end, different image processing tools have been tailored to integrate them into INBIOMED's platform as explained in next section.

Table 2. Relationship between the image data and the medical application

Image Type	Group
Electrophoresis Gels	Pharmacogenetics
ADN Gels	Pharmacogenetics
	Cancer Epidemiology
	Molecular and Genetic Epidemiology
Inmunofluorescency	Molecular and Genetic Epidemiology
Inmuno-histochemistry	
TMA	Epidemiology
Samples Sequencer	Cancer Epidemiology
CT	Cancer Epidemiology

5 Image Processing Integration

The integrated image processing system included within the INBIOMED platform is based on distributed architecture with three different modules:

1. The image processing manager, responsible for image processing methods. This module is also responsible for publishing these methods and making them available to users.
2. A set of web services that implement the image processing methods.
3. The INBIOMED integration platform. It introduces and adapts the web services for remote use of data, provided by the shared databases within or outside the network.

The manager system contains a web application, used as user interface – that allows accessing the system from any computer connected to Internet and a web browser. It also includes an application server that is in charge of the image processing management methods and the communications with the remote applications that offer them. Thus, the application server incorporates a communication kernel that obtains the required information from local databases or from databases integrated in the INBIOMED platform – in the latter case, the use of a Data Access Object (DAO) is required. The application server included in the system also provides access to computation servers. Communication between user interface and application servers is accomplished through servlets, while communication with remote web services is performed using SOAP messages.

Together with web capabilities, the architecture offers the possibility of integrating algorithms developed in different computer languages, enabling the integration and linkage of already developed libraries. Some examples of image processing algorithms, integrated using the INBIOMED platform are described below.

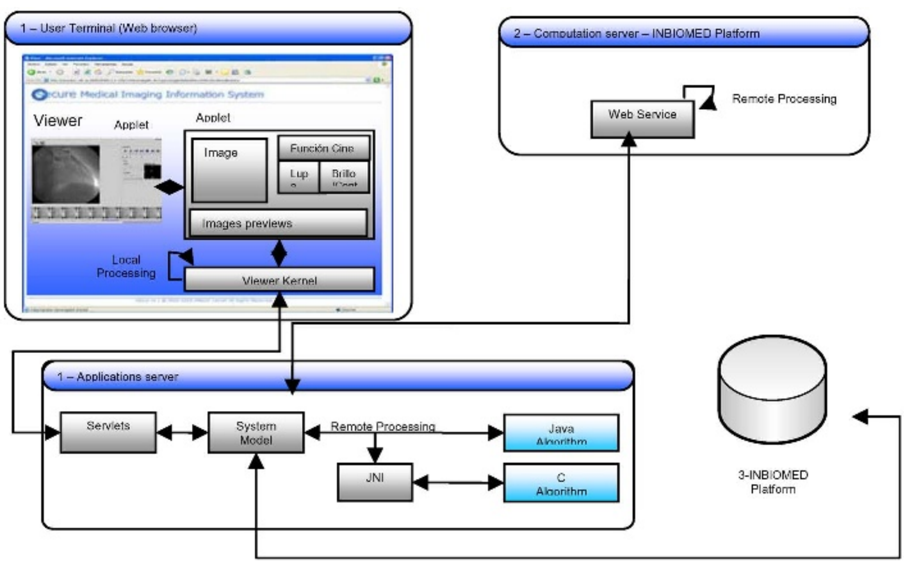

Fig. 1. Image processing within the INBIOMED platform

In the context of the INBIOMED research network different partners had been previously working on several biomedical image processing tools based on different methods. Some of these algorithms currently include:

- Morphological processing and analysis of images
  Morphological processing and analysis of images focus especially on the input image shapes, both in the filtering stage and in the segmentation stage [18] [19]. The Web services currently integrated in the platform provide morphological operators and filters, such as erosions, dilations, openings, closings, and openings and closings by reconstruction [20] [21]. In addition, analysis techniques are also provided, concretely, segmentation methods such as the watershed approach (with markers) and flat zone region merging techniques [22]. Such analysis techniques use a marker or mask image to indicate an initial set of particularly significant regions. This set can be computed in an automatic or semi-automatic manner, or previously by hand by a human user or specialist.
- Partial differential based models (PDE)
  PDE models are based on active contour models (ACM) or snakes [23]. The basic idea in the model is to evolve a curve, via minimization of an energy functional, subject to constraints from a given image, in order to detect region of interest within the image. Different models may be obtained, constrained to the resolution of the minimization problem. Thus, a geometrical and geodesic models where introduced by Caselles [24] and Malladi [25] to handle changes in the topology of the segmented shapes.
- Edge and region based detector
  Together with the previous image processing tools several pre-processing algorithms are included through the INBIOMED platform. These algorithms go from filters to edge and region based detectors.

In order to allow more compatibility, all the provided methods accept the most used image formats, including bmp, jpg, dic y pgm. Besides, although all methods return their results as grey scale images, input images can have any colour depth, avoiding the transformation in client application.

6 Use Cases

The integration image processing approach presented in this paper may have many applications within the biomedical area. Some of them, performed using the INBIOMED platform, are presented below.

Fig. 2 shows an example of the utilization of an application within the INBIOMED platform. In particular, the first two images correspond to a morphological erosion, and the second two visualizes a connected morphological filtering of the red band of a micro-array colour image.

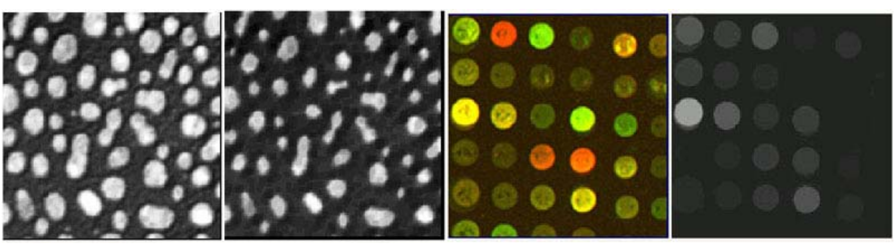

Fig. 2. Morphological processing through the INBIOMED platform

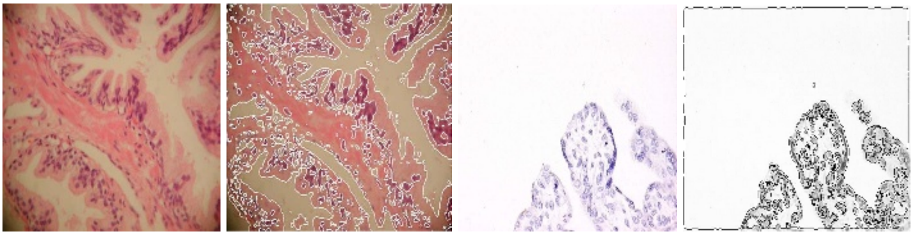

Fig. 3. Fuzzy C-Means clustering and PDE segmentation

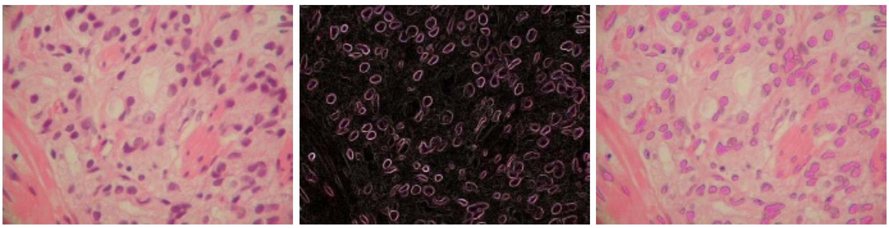

Fig. 4. Edge and region based detectors image processing

Fig. 3 and 4 display the results of some image processing tools applied to histological images. Figure 3 shows an example of fuzzy c-mean clustering, a PDE segmentation [26]. Fig. 4 shows an edge and region based detectors to highlight the contour and the nucleolus of the cells into the images.

7 Conclusions

Every biomedical image process presented in the previous section are integrated within the INBIOMED platform and can be remotely executed. They are available to be used by any node of the INBIOMED network, but also open to contributions outside the network, speacially intended for academic training purposes [27].

This paper has presented a new web service-based approach to integrate and publish image processing algorithms within the INBIOMED. Previous studies have shown that web services features matched many needs of image processing algorithms in the field of biomedicine. These methods must be independently developed using any programming language, operating system, etc. They also needed to be available to reuse previous methods and to be easily used by others in the future. The integration of these algorithms into a common platform and format was a priority as well. And finally, the communication should be clear, passing through firewalls and methods must be remotely accessible with a regular web browser.

All these characteristics are supported by web services. As a result, this technology has proven to be especially suited for image processing integration within the biomedical field. In addition, the INBIOMED platform and its further experiments have confirmed the suitability of this approach to integrate biomedical image processing tools.

Acknowledgements

This work was funded by the INBIOMED research network. We also would like to thank José Antonio López for its work within this project.

References

1. Lopez-Alonso, V., Sanchez, J.P., Liebana, I., Hermosilla, I., Martin-Sanchez, F.: INBIOMED: a platform for the integration and sharing of genetic, clinical and epidemiological data oriented to biomedical research. Proceedings of Bioinformatics and Bioengineering, BIBE 2004. pp 222- 226 (2004)
2. Erl, T.: Service-Oriented Architecture: A Field Guide to Integrating XML and Web Services. Prentice Hall, New Jersey (2004)
3. Grosso, W.: Java RMI, O'Reilly and Associates, Inc., Sebastopol, CA, USA (2001)
4. Waldo, J.: "Remote procedure calls and Java Remote Method Invocation", IEEE Concurrency, pages 5--7, July (1998)
5. Vinoski, S.: Where is Middleware? IEEE Internet Computing 6(2): 83-85 (2002)
6. Felber, P., Garbinato, B., Guerraoui, R.: The Design of a CORBA Group Communication Service. Proceedings of 15th Symposium on Reliable Distributed Systems (1996)
7. Vinoski S. "CORBA: Integrating Diverse Applications Within Distributed Heterogeneous Environments", 1997, IEEE Communications Magazine, vol. 14, no. 2, February 1997.
8. Chung, Y. Huang, S. Yajnik, D. Liang, J. Shin, C. Y. Wang, and Y. M. Wang.: DCOM and CORBA: Side by Side, Step by Step, and Layer by Layer. C++ *Report,* 18-29 (1998)
9. Adamopoulos, D.X., Pavlou, G., Papandreou, C.A., Manolessos E.: Distributed Object Platforms in Telecommunications: A Comparison Between DCOM and CORBA. British Telecommunications Engineering 18, 43-49 (1999)
10. Liang, K.C., Chu, W.C., Yuan, S.M., Lo, W.: From Legacy RPC Services to Distributed Objects, Proceedings of the Sixth IEEE Computer Society Workshop on Future Trends of Distributed Computing Systems (1997)
11. Snell, J., Tidwell, D., Kulchenko, P.: Programming Web Services with Soap, O'Reilly and Associates, Inc., Sebastopol, CA, USA (2001)
12. D. Kotz, R. S. Gray.: Mobile agents and the future of the Internet. ACM Operating Systems Review, 33(3):7–13, 1999
13. Bote-Lorenzo, M., Dimitriadis, A., Gómez-Sánchez, E.: Grid Characteristics and Uses: a Grid Definition. Proceedings of the First European Across Grids Conference (ACG'03), Springer-Verlag LNCS 2970, pp. 291-298, Santiago de Compostela, Spain, Feb. 2004.
14. Ernemann, C., Hamscher, V., Schwiegelshohn, U., Yahyapour, R., Streit, A.: On Advantages of Grid Computing for Parallel Job Scheduling. Proceedings of the 2nd IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGRID.02) (2002)
15. ImageProcess0.4 – Image Processing Tool, developed by sougeforge.net. Available in http://imageprocess.sourceforge.net/index.shtml (2004)
16. Smith, K., Paranjape, R.: Mobile Agents for Web-Based Medical Image Retrieval. Proceedings of the 1999 IEEE Canadian Conference on Electrical and Computer Engineering, Shaw Conference Center, Edmonton, Alberta, Canada May 9-12 1999.
17. Estruch, A., Heredia J.A., "Technological platform to aid the exchange of information and applications using web services" Lecture Notes in Computer Science, 3337, 458-468, Nov. 2004.
18. Serra, J.: Mathematical Morphology. Volume II: theoretical advances. Academic Press, London (1988)
19. Soille, P.: Morphological Image Analysis, Principles and Applications. Springer-Verlag, Heidelberg (2003)

20. Salembier P., Serra J.: Flat zones filtering, connected operators, and filters by reconstruction. IEEE Transactions on Image Processing, Vol 4, Num 8, pp 1153-1160 (1995)
21. Crespo, J., Maojo, V.: New Results on the Theory of Morphological Filters by Reconstruction. Pattern Recognition, Vol 31, Num 4, pp 419-429, Apr 1998
22. Meyer, F., Beucher, S.: Morphological segmentation. J. Visual Commun. Image Repres, Vol 1, Num 1, pp 21-45 (1990)
23. M. Kass, A. Witkin, D. Terzopoulos: Snakes: Active contour models. Int.J. Comput. Vis., 14 (26), pp. 321-331 (1988)
24. Malladi, R., Sethian, J.A., Vemuri, B.C.: Shape Modeling with Front Propagation: A Level Set Approach. IEEE Trans. on PAMI 17 (1995) 158–175
25. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic Active Contours. Int. J. Comput. Vis. 22 (1) (1997) 61–79
26. Bueno G., Martínez A., Adán A.: Fuzzy-Snake Segmentation of Anatomical Structures. Lecture Notes in Computer Science, LNCS, Springer Pb. Volume 3212, pp. 33-42 (2004)
27. Maojo V, Kulikowski C.A.: Bioinformatics and medical informatics: collaborations on the road to genomic medicine. Journal of the American Medical Informatics Association. 10 (6) 515-22 (2003).