A Demo of Multimodal Medical Retrieval

Ranveer Joyseeree
Dept. of Information Technology and Electrical Engineering
ETH Zürich, Zurich, Switzerland
ranveer.joyseeree@hevs.ch

Roger Schaer, Henning Müller
Institute of Information Systems
HES-SO Valais/Wallis, Sierre, Switzerland
roger.schaer@hevs.ch, henning.mueller@hevs.ch

Abstract—*Providing personalized medical care based on a patient's specific characteristics (diagnostic--image content, age, sex, weight, and so on) is an important aspect of modern medicine. This paper describes tools that aim to facilitate this process by providing clinicians with information regarding diagnosis, and treatment of past patients with similar characteristics. The additional information thus provided can help make better-informed decisions with regards to the diagnosis and treatment planning of new patients. Two existing tools: Shambala and Shangri--La can be combined for use within a clinical environment. Deployment inside healthcare facilities can become possible via the MD—Paedigree project.*

I. INTRODUCTION

Imaging data is currently being produced in large quantities on a daily basis in hospitals [1], where they are digitally stored in large repositories. For example, the Geneva University Hospitals produced over 300'000 images per day on average in 2013 [2]. Such data are usually accompanied by patient demographic information, such as age, sex, weight, occupation, and so on. Clinicians cannot decide on diagnoses using only image data [3]. Together, the multimodal data represent a wealth of information that directly link patients with disease diagnoses, and treatments.

By comparing the imaging and demographic data of past patients with those of new ones, there is potential for gaining insight into diagnoses and treatment planning for the latter by finding similar past cases. Such similarity search requires the enabling of retrieval in these large patient image and metadata collections.

This is exactly the purpose of the tool proposed in this paper. It aims to complement other tools in the **M**odel-**D**riven **Pa**ediatric **E**uropean **Dig**ital **Re**pository (MD-Paedigree)[1] to reach its objectives, namely of patient-specific medical care aided by knowledge of past clinical cases. At the beginning of the project, technical requirements were gathered and it became apparent that clinicians wish for insight into possible diagnoses and treatment options for a newly-arrived patient by studying similar patients that have been treated in the past. A study of the clinical workflows of medical personnel involved in the project further reinforced that notion. In other words, they are looking for a system that allows them to add the information (symptoms, demographic data, and images) pertaining to a new patient as input and get as output a list of similar patients with details of their respective diagnoses and treatments. Past cases need to be available for this scenario in a big multimodal data digital repository.

There are already several tools for search using text input and a combination of text and visual input. Examples include GoldMiner[2] (that allows filtering by image modality), RadMiner[3] (for text and semantic search aimed at radiologists), OpenI[4] (that allows searching based on text and image examples and also has filters, for example for automatically detected image modality), and Yottalook[5], which provides support for searching for text in the medical literature. Other systems, like IRMA (**I**mage **R**etrieval in **M**edical **A**pplications)[6] or img(Anaktisi)[7] focus entirely on the visual aspect. For multimodal search, the CITI (**C**enter of **I**nformatics and **I**nformation **T**echnology) proposes NovaMedSearch[8], a search engine for medical information able to search for similar images and connected cases.

Of particular interest in this paper are two existing tools: Shambala[9] [4] and Shangri-La[10]. Shambala combines CBIR (**C**ontent-**B**ased **I**mage **R**etrieval) with textual retrieval in the form of web search interface for clinicians to find images to fulfill an information need. It is a web-based multimodal search system for biomedical information and documents created initially in the Khresmoi project [5]. It is based on a retrieval system called ParaDISE (**Para**llel **D**istributed **I**mage **S**earch **E**ngine), described in [6]. Shangri-La is a modern web-based interface for medical case-based search. It is optimized for ease-of-use and is geared towards the needs of users.

Built for search in terms of two-dimensional (2D) images, Shambala and Shangri-La are, however, flexible enough to be adapted for other retrieval tasks, especially for retrieval of three-dimensional (3D) images. This possible through an adaptable and modular system design for each.

The rest of this paper is organized as follows. First, Shambala, is described in detail. Then, the same is done with Shangri-La. That is followed by a discussion regarding what needs to be done to attain the goal of achieving patient-specific healthcare.

978-1-4673-8695-1/16/$31.00 ©2016 IEEE

[1] MD-Paedigree. http://www.md-paedigree.eu/ [Accessed: 29.02.2016]

[2] Goldminer. http://goldminer.arrs.org/home.php [Accessed: 29.02.2016]

[3] RadMiner. https://averbis.com/wp-content/uploads/2015/01/Averbis-RadMiner-Brochure-en.pdf [Accessed: 29.02.2016]

[4] OpenI. http://openi.nlm.nih.gov/ [Accessed: 29.02.2016]

[5] Yottalook. http://yottalook.com [Accessed: 29.02.2016]

[6] IRMA. http://ganymed.imib.rwth-aachen.de/irma/ [Accessed: 29.02.2016]

[7] img(Anaktisi). http://orpheus.ee.duth.gr/anaktisi/ [Accessed: 29.02.2016]

[8] NovaMedSearch. http://medical.novasearch.org/ [Accessed: 29.02.2016]

[9] Shambala. http://shambala.khresmoi.eu/ [Accessed: 29.02.2016]

[10] Shangri-La. http://shangrila.khresmoi.eu/ [Accessed: 29.02.2016]

II. TOOLS THAT ARE PRESENTED

A. Shambala

In this section, the characteristics of Shambala and the ParaDISE backend that it relies on are explored.

1) Concept and Technologies: Shambala is a web application built upon the ParaDISE backend with ease-of-use in mind. It presents users with a simple interface having few options and parameters. It lets the user to 'drag-and-drop' images onto the interface, thus providing a similar experience to desktop applications. Searching, therefore, becomes a more engaging experience. It reacts swiftly to user input and refreshes search hits on the fly while the user updates a query. Integrating web technologies like **H**yper**T**ext **M**arkup **L**anguage **5** (HTML5) File & History **A**pplication **P**rogramming **I**nterfaces (APIs) and HTML5's Local Storage feature, it is a modern client-side only application, based only on HTML5 and Javascript. All interactions with ParaDISE (which can be hosted on a different server) use AJAX (**A**synchronous **J**avaScript **A**nd **X**ML) calls to the Global ParaDISE web service. Its operation is independent of the type of images. It is flexible enough to be adapted for 3D images.

2) ParaDISE Backend: ParaDISE, called by Shambala, allows indexing and retrieval of images using image features and text. It is made up of independent components. This increases its flexibility and makes it easier to distribute. In addition, it is possible to add new features, representations and distance metrics to it. This is enabled through the use of object-oriented programming and having a plugin organization of the component libraries. The system is easy to integrate with other components or user interfaces such as the MD-Paedigree repository (through the use of the standard HTTP transmission protocol). The use of JSON (**J**ava**S**cript **O**bject **N**otation), a standardized and light-weight data transfer format, allows for interoperability with other retrieval systems. ParaDISE also aims to be scalable, both for offline operations (such as image indexing) and online operations (for instance, image search). For the offline part, scalability is ensured through the distributed computing framework: Hadoop[11]. Its use and benefits are discussed in [7]. For online operations, parallel computing and efficient indexing techniques for fast similarity search are used. This enables the system to cope with large amounts of images.

The architecture of the ParaDISE backend is made up of four components (Figure 1). The **extractor** computes local visual features from images. A global representation of the image is obtained through the **descriptor**. Visual indices are written and read by the **storer**. Finally, a **fusor** compiles results from different queries using a pre-determined fusion rule. This part of the system may be integrated with other systems like semantic search engines. The above-mentioned four core components are connected to the **indexer** and the **seeker** (Figure 1), which allows the main retrieval operations: indexing and similarity search.

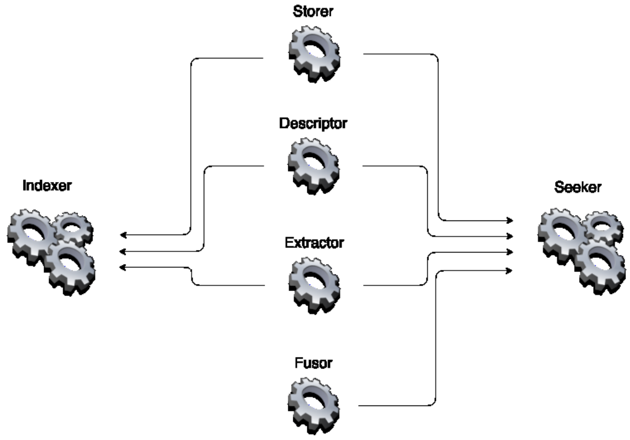

Fig. 1. Structure of the ParaDISE backend

3) ParaDISE Web Service Layer: Built on top of the above backend, the service layer of ParaDISE consists of several web services that use a REST-style (REpresentational State Transfer) architecture (Figure 2). It also provides features such as text search. To communicate with web services, standard HTTP GET and POST requests are used. This simplifies interaction between the system and various client applications (Web-based/desktop applications that can be written in any language making HTTP requests).

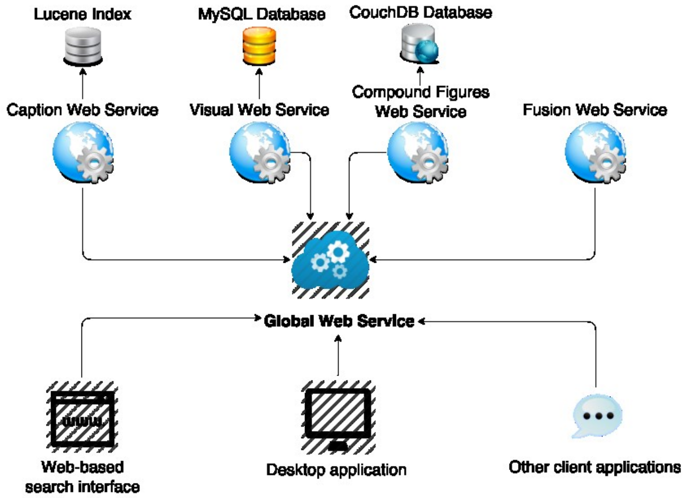

Fig. 2. Structure of the ParaDISE web service layer

The services of Figure 2 are:

- **Caption web service –** Supporting keyword search, this service uses the Lucene[12] search engine library.
- **Visual web service –** Backed by ParaDISE's seeker and indexer, this uses the visual information of the images in order to either retrieve similar matches using CBIR or index new images.
- **Compound figure web service –** This is used to obtain links between compound figures (multiple images in a single block) and their children (obtained by separating the compound figure into its subparts).

[11] Hadoop. http://hadoop.apache.org/ [Accessed: 29.02.2016]

[12] Lucene. http://lucene.apache.org/ [Accessed: 29.02.2016]

- **Fusion web service –** Using the fusion component of ParaDISE, this is responsible for compiling results from different sources (text or image search) into a single list (based on a given fusion rule).
- **Global web service –** This provides a front-end for client applications, thereby hiding the intricacies of calling the individual web services in succession. It provides general methods such as "search images", which use text and/or image examples.

4) Datasets: Currently, there are two datasets that can be accessed through the Shambala interface.

- **ImageCLEFmed 2012 dataset** - ImageCLEF[13] is an information retrieval competition and is a part of the Cross-Language Evaluation Forum (CLEF). Over 300'000 images from the biomedical open access literature are available.
- **PubMedCentral dataset** - Containing over 1.5 million images from the biomedical open access literature, the PubMedCentral[14] repository of 2014.

5) Interface: Shambala has a simple front-end and that is shown in Figure 3.

Fig. 3. User interface of Shambala. Zone (1) is for 'drag-and-drop' of images relevant for search and for keyword input. Zone (2) is for irrelevant images and keywords. Zone (3) is provides a results view. Zone (4) allows for uploading images. Zone (5) allows the configuration of search options.

The interface is split up into 5 main zones:

- **Relevant zone** - The user can add relevant keywords and provide examples of relevant images here.
- **Irrelevant zone** - The user can filter out keywords and provide examples of irrelevant images here
- **Results zone** - The search hits are displayed here.
- **Image sample zone** - This menu allows the user to get random sample images or upload his images that can be added to the relevant or irrelevant zones.
- **Search option zone** - This pull-down menu allows the user to configure the search options.

Extending this to retrieval of 3D images is straightforward thanks to the modular and extensible structure.

B. Shangri-La

In this section, the basic concept behind and the various technologies underpinning Shangri-La are explored.

1) Concept and Technologies: Shangri-La is similar to Shambala in terms of simplicity and ease-of-use. It was developed to help clinicians intuitively find medical cases and articles from the literature. Shangri-La, is also a client-side application based on HTML5 and Javascript. It also uses the ParaDISE backend and interacts with it using AJAX calls.

2) Shangri-La-specific ParaDISE alteration: Shangri-La interacts with ParaDISE mostly in the same way as Shambala. A difference is that support for case-based retrieval was added to ParaDISE during the development of Shambala, thus allowing for search through medical article and case repositories. The technologies supporting case-based retrieval are described in [8]. The web service layer of Shangri-La, shown in Figure 4, is also different from that of Shambala.

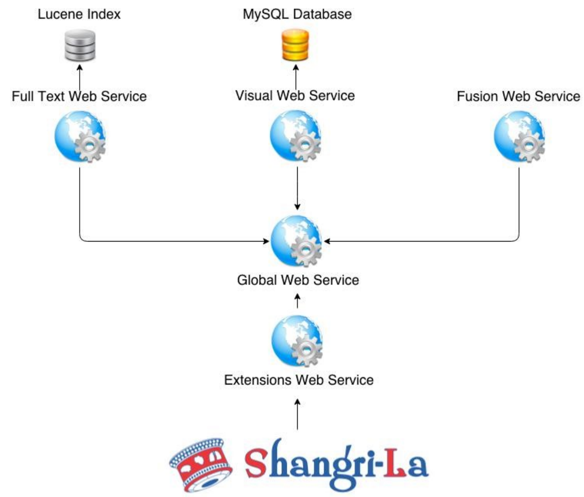

Fig. 4. Structure of web services within ParaDISE for use with Shangri-La.

3) Datasets: As with Shambala, the ImageCLEFmed 2012 and PubMedCentral datasets are available with Shangri-La.

4) Interface: Shangri-La also has a simple, intuitive front-end with three main functionalities.

- **Build case** – This allows the user to build a record of a query case. A case can be named in a text box (Figure 5) for later consultation. Images and a description of his case can be added onto Shangri-La.
- **Results** – This page displays a ranked list of similar cases/articles to the query from the literature (Figure 6). The layout was influenced by a study on search behavior of users [9]. Each displayed case can be expanded to show more details (Figure 7). Articles thus found can be bookmarked [10].
- **Recalling previous cases** – Shangri-La allows the user to consult previous cases created using the Build

[13] ImageCLEF. http://www.imageclef.org/ [Accessed: 29.02.2016]
[14] PubMedCentral digital repository. http://www.ncbi.nlm.nih.gov/pmc/tools/openftlist/ [Accessed: 29.02.2016]

Case page. A list of previously saved articles can be consulted again, as shown in Figure 8.

Fig. 5. Screenshot of the *Build Case* page of the Shangri-La interface. A text box allows the user to specify a case name (here: "Lung consolidation"). Images can then be uploaded (here: two chest X-rays). Finally, a case description may be added via another text box (here: "A 47-year old man…").

Fig. 6. The results page of Shangri-La. A list of articles is presented in order of relevance to the query presented previously by the user.

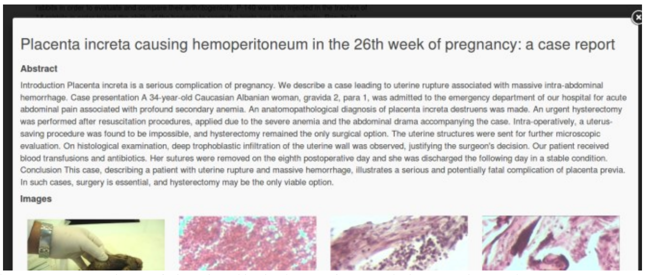

Fig. 7. Screenshot of the detailed view of an article on Shangri-La.

Fig. 8. The interface allows users to view previously saved articles.

III. Discussion and Conclusion

Shambala and Shangri-La are both flexible, extensible, and modular. To implement retrieval for the MD-Paedigree project, support for 3D images is needed. This will require changes in the underlying ParaDISE backend. A modern browser supporting HTML5 is also essential. Older browsers need to be updated to avoid security issues in any case. Some other additions already exist, such as the Leap Motion Controller allowing hands-free control of on-screen objects through a small motion detector [11]. Retrieval with the Google Glass [12] was also implemented. Such hands-free technology would allow the use of the tool in sterile environments. In the demo for this work, an example of retrieval of similar images and cases to a case at hand will be presented.

Acknowledgment

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 600932.

References

[1] C. Akgül, D. Rubin, S. Napel, C. Beaulieu, H. Greenspan and B. Acar, "Content--Based Image Retrieval in Radiology: Current Status and Future Directions," Digital Imaging, pp. 208-222, 2011.

[2] A. Widmer, R. Schaer, D. Markonis, and H. Müller, "Gesture Interaction for Content–based Medical Image Retrieval," in *International Conference on Multimedia Retrieval*, Glasgow, UK, 2014.

[3] G. Quellec, M. Lamard, L. Bekri, G. Cazuguel, C. Roux, and B. Cochener, "Medical Case Retrieval from a Committee of Decision Trees," IEEE Trans. Inf. Technol. Biomed., vol. 14, no. 5, pp. 1227–1235, 2010.

[4] R. Schaer, and H. Müller. "A modern web interface for medical image retrieval." Swiss Medical Informatics 30, 2014.

[5] A. Hanbury, C. Boyer, M. Gschwandtner and H. Müller, "KHRESMOI: Towards a Multi--Lingual Search and Access System for Biomedical Information," Luxembourg, 2011.

[6] R. Schaer, D. Markonis, and H. Müller. "Architecture and applications of the Parallel Distributed Image Search Engine (ParaDISE)." In GI-Jahrestagung, pp. 661-666. 2014.

[7] D. Markonis, R. Schaer, I. Eggel, H. Müller and A. Depeursinge, "Using MapReduce for Large-scale Medical Image Analysis," in HISB, La Jolla, California, 2012.

[8] A. García Seco de Herrera, "Use Case Oriented Medical Visual Information Retrieval & System Evaluation," Univ. of Geneva, 2015.

[9] O. Drori, "Improving Display of Search Results in Information Retrieval Systems–Users' Study," Center for Research in Computer Science of the Leibniz, 2000.

[10] S. Shen and S. D. Prior, "My Favorites (Bookmarks) Schema: One Solution to Online Information Storage and Retrieval," in *Proceedings of the 2013 International Conference on Information Systems and Design of Communication*, pp. 33–40, 2013.

[11] F. Weichert, D. Bachmann, B. Rudak, and D. Fisseler, "Analysis of the Accuracy and Robustness of the Leap Motion Controller," *Sensors*, vol. 13, no. 5, pp. 6380–6393, 2013.

[12] A. Widmer, R. Schaer, D. Markonis, and H. Müller, "Facilitating Medical Information Search using Google Glass connected to a Content-based Medical Image Retrieval System," *Conference of the IEEE Engineering in Medicine and Biology Society*, Chicago, IL, USA, 2014.