Concept-Oriented Access to Longitudinal Multimedia Medical Records: A Case Study in Brain Tumor Patient Management

Shahram Ebadollahi[1], James W. Cooper[1], David Kaufman[2], Anthony Levas[1], Andrew F. Laine[3], Robert DeLaPaz[4], and Chalapathy Neti[1]

[1]IBM T. J. Watson Research Center, Hawthone, NY
[2] Dept. Medical Informatics, Columbia University, New York, NY
[3] Dept. Biomedical Engineering, Columbia University, New York, NY
[4] College of Physicians and Surgeons, Columbia University, New York, NY

Abstract. The current clinical practice requires physicians to gather, interpret and correlate information from multiple independent multimedia data sources to manage patients. Due to poor structuring and organization, it is too time-consuming to access the information snippets embedded in the various pieces of data in the longitudinal patient records. This becomes more of a problem when correlating the temporal progression of various factors obtained from patients clinical, laboratory, imaging and genomics studies. Making such correlations is an essential component of the prognosis and treatment planning tasks in patient care. In addition, the similarities in the disease progression pattern among different patients and their relationships to outcomes remain hidden from the clinicians in the ”piecemeal” use of the data. We believe that there is a gap between the decision-enabling information and insight required for efficient patient management and the heterogeneous data comprising the patient records that can be bridged with advanced multimodal content analytics, semantic information organization, summarization, and visualization tools. In this paper we present a case study in organizing, accessing, and visualizing information obtained through structuring the multimedia and multimodal data for brain tumor patient management and how such information map to the information needs of the clinicians. We report our early work on the analytics, user interface and the preliminary evaluation results which indicate that the presented approach caters well to the clinician needs for the task of brain tumor patient management.

1 Introduction

The current clinical practice of neuro-oncology requires physicians to correlate and interpret information from multiple independent data sources to diagnose, treat and manage patients with brain tumors. The heterogeneous sources of data needed for these complex tasks are multi-media and multi-modal [1] in nature. Each piece of data intends

[1] Here medium refers to a distinct type of communication channel for conveying semantic information such as text, audio, image, video, volumetric images, etc. Modality on the other hand refers to type variations for a given medium. For example, volumetric images can be Magnetic Resonance images of type T1, T2, FLAIR, or Computed Tomography images. Or there can be many types of text data such as radiology reports, clinical notes, discharge summaries, etc.

to capture specific information about the health status of the patient (primary information), or reflect the decisions made by the clinician at a given point in time (secondary information). These information when properly linked, correlated and summarized can form an overall picture of the patient's condition and provide insight on the effectiveness of the course of treatment or the natural progression of the disease, which can then be used as the basis for making further clinical decisions.

Today, access to these information resources is typically a ”manual” process; regardless of how technologically advanced a health-care center is in adopting the Electronic Patient Record (EPR) [1]; requiring time-consuming interrogation of each relevant piece of data to identify the information that best map to the information needs of the clinician. The current ”piecemeal” use of data is too time-consuming and leaves the task of threading and correlating the information entirely to the clinician. In addition correlations among pieces of information may remain hidden and never noticed by the clinician. This becomes more pronounced in the case of correlating the temporal progression of a patient's clinical status, laboratory and imaging studies, which are the key factors in brain tumor patient care and prognosis. Figure 1 demonstrates the co-evolution of the volume of edema region [2] associated with the brain tumor of a patient diagnosed with Glioblastoma Multiforme [2] and the dosage of a particular drug. Being able to access, visualize, and compare such temporal trends of different factors obtained from the contents of the multi-media and multi-modal data for different patients is essential for proper patient management tasks.

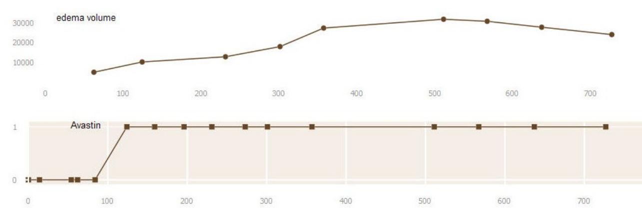

Fig. 1. Figures shows the co-evolution of the control variable (drug=Avastin) and the response variable (volume of edema region) for a given patient over time.

This paper reports our preliminary work on bridging the gap between the wealth of information embedded in the contents of heterogeneous data in the longitudinal patient records and the information needs of the clinicians for better manaing brain tumor patients. We present the architecture and plan for the Concept-Oriented Structuring of Multimedia medical records (COSMus) system, which enables concept-oriented structuring of the contents of multimedia medical records. In addition, we elaborate on our work on image content analytics and the user interface for summarizing and presenting the information to the clinician. We also conducted limited clinical evaluation, which the results indicate that summarizing, threading and correlating information obtained from various sources do have a positive effect in better understanding the patient condition and enabling better decision making by the clinicians.

2 Concept-Based Organization

The purpose of the longitudinal medical records is to document the various aspects of the patient's condition over time. The patient condition and overall status is assessed from the collection of information snippets embedded in the various pieces of data at any given time. For every given sub-speciality, there are certain concepts of interest along with their attributes that are being used to assess the patient condition. There are many existing ontologies used today to represent concepts relevant to human anatomy, disease, treatment plans and procedures, such as FMA [3], RadLex [4], SNOMED [5], and MeSH [6]. These ontologies provide a common vocabulary with an agreed upon meaning (semantics).

In this work, we aim at representing and linking instances of concepts of interest and their attributes as manifested in the heterogeneous data sources. We define the notion of a *ConceptFrame* that affords knowledge-guided extraction of medical concepts (along with their associated attributes) from heterogeneous medical artifacts. Figure 2 illustrates the idea of the *ConceptFrame* for the concept edema. As shown in this figure, there is a mention (instance) of edema in the MRI of type FLAIR, which visually captures the region corresponding to the concept of *edema* [2]. Edema has also been mentioned in the text of the oncology note, where the physician is expressing her understanding of the effect of edema. Based on the concept, there are certain attributes that describe the concept. For example, it is important to know the volume of the edema region or what anatomical entities it affects. In addition there are preferred methods for accessing and visualizing a concepts and its attributes. These are all captured in the knowledge model, which we refer to as the *ConceptFrame*.

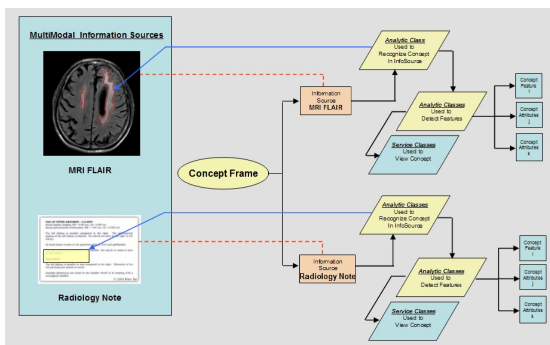

Fig. 2. Figures shows the *ConceptFrame* for the concept edema.

The task of instantiating ConceptFrames and populating them is peformed by analytic engines that specialize in finding instances of the concepts in various sources of multimedia/multimodal data. For the type of data we have in this study, i.e. MR images and oncology notes, we need image analytics that can find instance of the concept edema in the volumetric images of the brain and text analytics that find concepts such as drug, diagnosis, etc. Figure below shows the architecture of the COSMus system, which is responsible for applying the appropriate analytics to the right data source and instantiating the *ConceptFrames*. This architecture illustrates our final goal. In the current work we have a limited proof-of-concept implementation to let us study the effect of organized and linked information in the clinical practice. The following section we present our work on using transductive and inductive mechanisms for concept identification in multi-protocol MR images. For text analytics, we refer the reader to our prior work in medical text analytics [7].

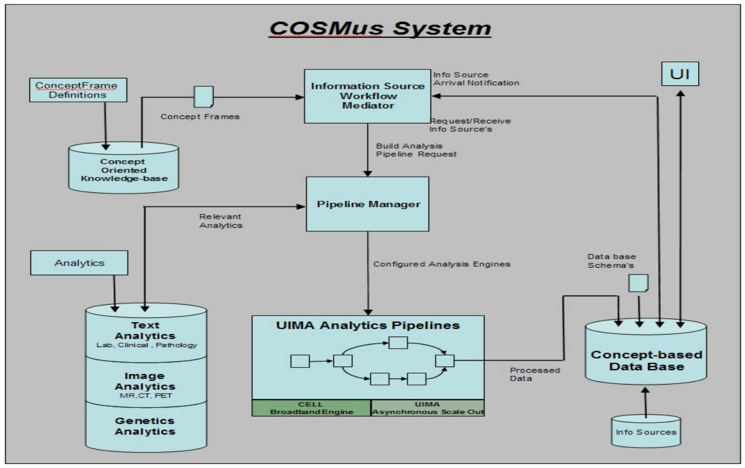

Fig. 3. The architecture of the COSMus system.

3 Image Analytics

A number of methods to detect and track changes in MR images have been proposed [8, 9]. Most of the existing systems use inductive learning techniques to create a model capable of distinguishing and categorizing different classes. An inductive learning method such as support vector machines (SVMs) [10] uses a set of labeled input data for training purposes and produce a generic model which can be used to automatically label new images. The primary limitation of inductive techniques is the training data. In the medical domain the labeling process requires expert knowledge and often tedious editing effort to obtain accurate label information for the object of interest.

Recently, semi-supervised learning methods such as transductive inference have been getting a significant amount of attention given their effectiveness on quickly labeling a given set of input data. A transductive method minimize the human interaction by inferring the labels of a complete dataset from a small initial expert input [11]. The primary limitation of transductive techniques is that the training (data along with provided labels) and the test data should be avilable at the time of training.

For locating and characterizing the concepts of interest which are manifested in longitudinal MR images we combine the transductive and inductive learning techniques. After the registration phase, which aligns the multi-modal set of images at a given timepoint into a common coordinate system, we capture minimal input provided by domain expert for identifying the concept of interest (in this case the edema region). We then use a Bayesian transductive learning approach [12] to account for non-identically distributed data domains as well as integration of expert knowledge through adaptive probabilistic modeling. The classification obtained from the transductive inference are used as pseudo-ground truth to train the inductive model. For each training point, a combination of first- and second-order statistics are estimated to create a multi-dimensional descriptor. In particular, histogram features including mean, skewness, and standard deviation are extracted from each training point in conjunction with textural features such as energy, contrast, and correlation. Those set of features are combined and used as the characteristic descriptor for each training point under consideration. SVMs are used to learn an inductive and more generic classification model capable of automatically identifying the pathological concept under consideration within new data. The learned inductive models are then employed to automatically identify the medical concept of interest for the new input data.

The transductive approach for obtaining the pseudo ground truth shows promising results with a sensitivity and specificity of 90.37% and 99.74% respectively (see [12]). The technique is computationally efficient and takes about 1-3 seconds on $256\times256\times30\times9$ multi-modal datasets using a dual core 2.4 GHz machine showing its suitability to be used within an interactive environment. Figure 4 displays the initial input by the domain expert and the propagation of the labels obtained by using the transductive approach.

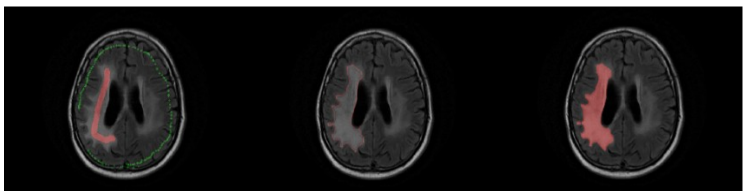

Fig. 4. The brush strokes provided by the expert to identify a portion of edema and non-edema regions are shown on the left. The right-hand side images show the result of label propagation on the original image shown in the middle using the transductive mechanism.

Leveraging the pseudo-ground truth resulting from using the transductive method on a few images and creating inductive (SVM) classifiers for the concept of *edema*, produced a classifier which was able to classify other images in the longitudinal patient records with a minimum accuracy of 80%, when the SVM classifier was trained using pseudo-ground truth for 3 time points. This indicates that one can capture limited amount of expert annotations and train a classifier that can generalize well to other images, without overwhelming the domain experts.

4 User Interaction with Multimedia Temporal Information

We have built and tested a prototype system for analyzing and displaying patient images and text-based oncology reports of brain tumor patients accumulated over time. The text analysis includes the extraction of neurological symptoms, drug and dosage information, lab results, vital measurements and assessments. This is coupled with the ability to scroll through 3D MRI images, plot and highlight edema volumes and plot any recurring numerical value, such as weight, platelet count and blood pressure. Limited clinical evaluation has been performed to validate the usability of the prototype system in clinical decision making.

Figure 6 shows the components of the user interface designed to access and interact with structured information derived from multimedia sources. The system was

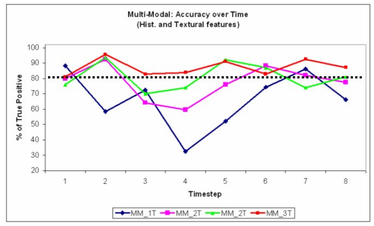

Fig. 5. Results of applying the inductive classifier trained using the pseudo-ground truths obtained from applying the trnasductive mechanism to 1, 2, and 3 images for patients with data for 10 time points on average.

initially designed and developed based on several rounds of discussion between computer scientists, radiologists and oncologists with a goal of providing the most useful information on each screen. Then it was demonstrated to other members of the team. We conducted a cognitive evaluation with two residents (one in oncology and one in radiology) to determine the relative ease of use in obtain, coordinating and intergrating patient information to make treatment decisions. The subjects were asked to look at combined text and image data for two patients and answer several questions: Did the patient experience seizures at any time; when did they have the most edema; when did they have the lowest platelet count. They were asked at the point of maximum edema whether the patient had any other significant symptoms and wheter the images indicated anything significant abou the patient's prognosis. Finally, they were asked whether the drug Avastin effected the amount of edema. During the study, they were asked to think aloud and comment on the utility and usability of the system. At the conclusion of the study, they completed a brief 6 item Likert survey.

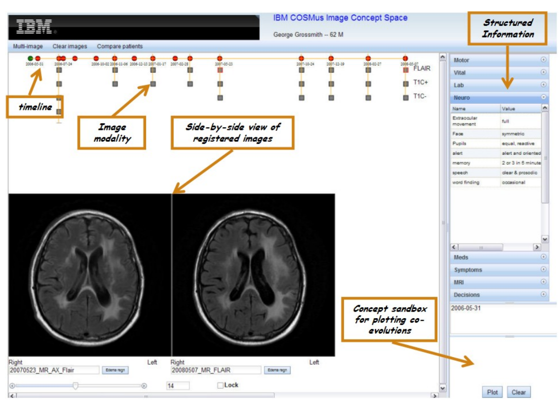

Fig. 6. Figures shows the co-evolution of the control variable (drug=Avastin) and the response variable (volume of edema region) for a given patient over time.

4.1 Clinical Evaluation

During the initial presentation to members of the team, several remarked that the ability to plot data such as edema volume and drug dosage led to immediate and unexpected insights. Both subjects were able to use the system effectively with minimal training. The experimenters provided assistance only when the subjects were stumped. The complete session with each subject was recorded, and notes were taken by the investigators. Although the participants were able to employ most system functions, they experienced a range of problems. The semantic mappings used to label buttons that subsume findings were not always intuitive. For example, both subjects expected that the ”seizures” findings would be subsumed under the Neurology tab when in fact they were listed under ”Symptoms” as it was in the oncology notes. Both subjects had some difficulty discerning the date of each time point they were examining as that date was displayed just under the accordion control rather than closer to the actual time data point. If they were placed in such proximity, they would have overlapped each other.

Initally, the neuro-oncology resident experienced some difficulty selecting and juxtaposing images in the 2 panels. However after a short period of time, she had no difficulty finding the edema information and was able without any assistance to make a comparative plot of edema volume and Avastin dosage. She was also able draw appropriate inferences about changes in the patient’s condition over time.

In summarizing the patient’s status in question 7, she readily grasped that the current status would be available by clicking on the last data point. She noted that the best part of the interface was the way it integrated views of text and image data, and that it was most useful to be able to look at any pairs of images together. She remarked ”Being able to determine your thrombocytopenia at that point and the edema volume feature is the best aspect of this whole thing.”

The radiology fellow was more interested in examining the images. He showed less interest in the symptoms reported as text in the accordion tabs. He readily selected among the FLAIR, T1C+ and T1C- images and found them easy to study. In general, it took him more time and mouse clicks to find what he was looking for and longer to develop a basic mastery of the system. The system works by placing the image from the first thumbnail you select in the left image box and the second on the right. Then if you select a third image, it replaces the first one. The system provided insufficient feedback and guidance and he found it a source of confusion. He also noted that the ”Lock” checkbox which decouples one image from the other worked exactly the opposite of the ”Link” checkbox provided in the GE PACS system. The fellow was more conservative in drawing inferences from the images. In this system, the total edema volume was computed semi-automatically as described above. He commented ”everything that this machine is calling edema is not necessarily relevant to the patient’s survival because that could just be the radiation if that’s in the radiation field. The part of FLAIR signal abnormality which is in the field of the tumor or on the tumor may be more relevant, but once it becomes confluent, you can’t tell which one is tumor and which one is radiation change.” He scrutinized the images more carefully and commented on the resolution and potential noise that made clinical inferences less certain. He was also less sure about the value of integrating the text and image modalities because his work was almost entirely with the images. Despite his critical comments, he remarked that the system offered great potential as a clinical tool. His comments offered many excellent insights into the improvement of the interface.

On the Likert survey, both subjects rated the system very highly in terms of ease of use, learnability and were especially appreciative of the ability to integrate disparate sources of data. They were less convinced that this would be an effective tool to discover novel dimensions of the patients' problems. Both residents agreed emphatically that this was a tool with great clinical potential.

5 Conclusion and Future Work

The initial results of our work on extracting information from multimedia and multi-modal medical data and organizing them around a set of concepts of interest shows promise in the value of such approaches for aiding the clinicians in understanding the patient data better, which evenntually result in making better clinical decisions. We demonstrated that one can acheive a balance with obtaining limited annotations from the domain experts on the concepts of interest and leveraging them to desing a classifier for concept identification in patient's unseen longitudinal data sets. In addition the display and communication of the concepts and their attributes, rather than raw data, was helpful in expediting the time to information and aiding the clinicians see co-evolutions of clinical factors derived from the data. We plan to fully implement the COSMus system along with the supporting analytics for the domain of brain tumor data management.

6 Acknowledgements

The authors would like to acknowledge the contributions of Jesus Caban and Noah Lee for developing the image analytics and conducting the experiments. We would also like to acknowledge the advice and support provided by Dr. Steven Rosenfeld for conducting the clinical evaluations.

References

1. Lehmann, H.P., Abbott, P.A., Roderer, N.K., Rothschild, A., Mandell, S., Ferrer, J.A., Miller, R.E., Ball, M.J.: Aspects of Electronic Health Record System. Springer (2006)
2. Berger, M.S., Prados, M.: Textbook of Neuro-Oncology. Saunders (2004)
3. Structural Informatics Group, U.o.W.: "fma: Foundational model of anatomy"
4. RSNA: "radlex - a lexicon for uniform indexing and retrieval of radiology information resources"
5. of Medicine", N.L.: "snomed ct: Systematized nomenclature of medicine - clinical terms"
6. of Medicine", N.L.: "mesh - medical subject headings"
7. Mack, R.L., Mukherjea, S., Soffer, A., Uramoto, N., Brown, E.W., Coden, A., Cooper, J.W., Inokuch, A., Iyer, B., Mass, Y., Matsuzawa, H., Subramaniam, L.V.: Text analytics for life science using the unstructured information management architecture. IBM Systems Journal **43**(3) (2004)
8. Bosc, M., Heitz, F., Armspach, J.P., Namer, I., Gounot, D., Rubach, L.: Automatic change detection in multimodal serial mri: application to multiple sclerosis lesion evolution. NeuroImage **20** (2003) 643–656
9. Angelini, E., Atif, J., Delon, J., Mandonnet, E., Duffau, H., Capelle, L.: Detection of glioma evolution on longitudinal mri studies. In: Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on, Arlington, VA (2007)
10. Vapnik, V.N.: Statistical learning theory. Wiley (1998)

11. Duchenne, O., Audibert, J., Keriven, R., Ponce, J., Segonne, F.: Segmentation by transduction. In: IEEE Computer Vision and Pattern Recognition (CVPR08). (2008)
12. Lee, N., Caban, J., Ebadollahi, S., Laine, A.F.: Interactive segmentation in multi-modal brain imagery using a bayesian transductive learning approach. In: SPIE Medical Imaging (submitted). (2009)