Medical Sieve: A Cognitive Assistant for Radiologists and Cardiologists

T. Syeda-Mahmood\textsuperscript{a}, E. Walach\textsuperscript{b}, D. Beymer\textsuperscript{a}, F. Gilboa-Solomon\textsuperscript{b}, M. Moradi\textsuperscript{a}, P. Kisilev\textsuperscript{b}, D. Kakrania, C. Compas, H. Wang, R. Negahdar, Y. Cao, T. Baldwin, Y. Guo, Y. Gur, D. Rajan\textsuperscript{a}, A. Zlotnick, S. Cohen, R. Ben-Ari, Amit Guy\textsuperscript{b}, P. Prasanna, M.D., J. Morey, M.D., O. Bokyo, M.D.\textsuperscript{a}, and S. Hashoul, M.D.\textsuperscript{b}

\textsuperscript{a}IBM Research - Almaden Labs, 650 Harry Road, San Jose, CA 95120.
\textsuperscript{b}IBM Research - Haifa Labs, Haifa University Campus, Mount Carmel, Haifa, 3498825, Israel.

ABSTRACT

Radiologists and cardiologists today have to view large amounts of imaging data relatively quickly leading to eye fatigue. Further, they have only limited access to clinical information relying mostly on their visual interpretation of imaging studies for their diagnostic decisions. In this paper, we present Medical Sieve, an automated cognitive assistant for radiologists and cardiologists designed to help in their clinical decision-making. The sieve is a clinical informatics system that collects clinical, textual and imaging data of patients from electronic health records systems. It then analyzes multimodal content to detect anomalies if any, and summarizes the patient record collecting all relevant information pertinent to a chief complaint. The results of anomaly detection are then fed into a reasoning engine which uses evidence from both patient-independent clinical knowledge and large-scale patient-driven similar patient statistics to arrive at potential differential diagnosis to help in clinical decision making. In compactly summarizing all relevant information to the clinician per chief complaint, the system still retains links to the raw data for detailed review providing holistic summaries of patient conditions. Results of clinical studies in the domains of cardiology and breast radiology have already shown the promise of the system in differential diagnosis and imaging studies summarization.

Keywords: Clinical Decision Support, Precision Medicine, Cognitive Assistant, Data Summarization, Differential Diagnosis

1. INTRODUCTION

With the growth in digital imaging such as MR, CT and perfusion imaging, radiologists and cardiologists have to examine a large number of images a day. A typical emergency room radiologist may look at as many 200 imaging study cases a day containing thousands of images, such as in lower body CT angiography where there can be as many as 3000 images per study. Eye fatigue is therefore a common problem facing radiologists and cardiologists. Moreover, due to the volume of imagery to examine, these specialists have little time left to assimilate all pertinent clinical information from patient records before they interpret the images. As a result, overlooking of diagnosis, and misinterpretation is a common problem with diagnosis error rates exceeding 30%.

In this paper we present Medical Sieve, a cognitive assistant system for radiologists and cardiologists that acts as a medical sieve retaining the essential clinical information about the patient from their electronic health records and imaging studies. The creation of a summary and differential diagnosis involves sophisticated textual and image processing, anatomical localization and disease understanding many of which use machine learning algorithms to learn the various patterns. The resulting cognitive assistant system can be used in conjunction with existing PACS and EMR (electronic medical record) systems to provide a unified reasoning forum for the radiologists and cardiologists.

Further author information: (Send correspondence to Tanveer Syeda-Mahmood)
Tanveer Syeda-Mahmood.: E-mail: stf@us.ibm.com.
2. METHODS

The overall architecture of the Medical Sieve cognitive assistant system is illustrated in Figure 1. The system takes two major sources of input for its analysis, namely, patient data from enterprise health systems in hospitals, and knowledge from clinical sources. It then analyzes the incoming patient data using a core analysis engine that extracts anomalies from raw imaging, and disease-specific clinical measurements. These measurements are combined with clinical knowledge to infer additional details about the patient’s condition. The resulting information extracted is summarized and presented in a compact way to clinicians. Finally, the system leverages the accumulated knowledge about patients to offer cross-patient informatics. The result of the analysis produced by Medical Sieve can be utilized by PACS and EMR systems alike for further uses in clinical decision support.

2.1 Patient data ingestion

The system has built-in adapters to bring in data from hospital exchanges and vendor neutral archives (VNA) or PACS systems using standard protocols such as HL7 and DICOM. Once the data is ingested, the system processes incoming HL7 messages and separates clinical information present in these messages. A proprietary patient data model is used to assemble a longitudinal patient record from the extracted clinical information. This model records clinically relevant information about the patient including chief complaint, history of present illness, past medical history, medications, problem lists, etc.

2.2 Incorporation of clinical knowledge

As shown in Figure 1, Medical Sieve utilizes patient-independent knowledge. In particular, it has a large clinical knowledge base derived from UMLS ontologies,\(^1\) and semi-automatically generated assertions or facts about diseases and their relationships with signs and symptoms, medications, lab measurements, etc. The semi-automatic assembly process uses algorithms for concept and relationship extraction from textual sources as described in.\(^2\) This knowledge is weighted with evidence from the actual co-occurrence of clinical features in large scale electronic health records.

![Figure 1. Illustration of the Medical Sieve Cognitive Assistant System.](image)

2.3 Analysis of modality imaging

The imaging studies from PACS or VNAs are extracted using a DICOM query-retrieve mechanism. The DICOM studies are then analyzed to segment by mode and viewpoint as described in our earlier work.\(^3, 4\) In each viewpoint image study sequence, the organs of interest are extracted using multi-atlas label segmentation algorithms as...
described in. Once anatomical regions are localized, anomalies are detected within regions as deviations from the normal appearances described in the atlas. Finally the diseases are classified and measurements from the disease-specific regions are extracted. The multimodal analytics engine is designed for operation over large scale imaging datasets starting from raw DICOM format in an unattended fashion. The major stage of processing are illustrated in Figure 2.

2.4 Patient Data Summarization

The system offers clinical summaries of both textual and imaging data. For textual summaries, the system analyzes the underlying longitudinal patient record to record major events and episodes. The current encounter’s chief complaint is then used to retrieve relevant elements from the patient record. The system classifies each clinical concept extracted from the patient record (structured as well as unstructured reports) as belonging to one of 13 categories of chief complaint elements summarized in Figure 3 using machine learning on a large corpus of trained collections and their associations with chief complaints. The summarization of images is dependent on the modality as well as the disease depicted. Different forms of summaries are offered ranging from salient viewpoints to salient anomalies. For example, a summary of a coronary angiography study may pick salient keyframes as those that have a high vessel visibility measure. That is, the images picked depict the coronary arteries in their best possible visibility. A measure of anomaly-driven saliency can similarly be defined based on diseases detected and used to rank images from the study. The top K images across the study are then assembled to form a study summary with retained links to the entire study underneath for faster browsing by the clinician. Figure 4 shows an example of automatic summary produced for the coronary angiography study iconified in Figure 5. The system determines the salient key frames within the angiography study using a vesselness measure as described in. Next, it recognizes that the problem is with the right coronary artery (as also indicated by the EKG results of ST elevation in the V1-V2 leads), and analyzes the right coronary artery to detect the anomaly as shown in the inset in Figure 4. The details of algorithms for coronary artery stenosis detection are described elsewhere.

2.5 Cross-patient informatics

Among the features of Medical Sieve is the ability to derive information from cross-patient analysis. One such algorithm used in the system is based on multimodal patient similarity. Here disease-specific similarity of images as well as textual reports are used to retrieve similar cases from pre-diagnosed datasets. The patient similarity algorithms for various diseases and modalities have been reported in a number of our earlier papers. These algorithms allow the system to benefit from the consensus opinions of other physicians who have looked at similar data sets.
2.6 multimodal reasoning

Utilizing the extracted summary information from both textual and imaging data, the system then reasons using the clinical knowledge to arrive at a differential diagnosis for the patient given all available facts derived from provided data using goal-directed Artificial Intelligence algorithms based on branch-and-bound search. These algorithms essentially traverse the clinical knowledge graph using evidence gathered from the patient data to reach the goals of differential diagnosis. Besides knowledge-guided reasoning, the system also provides data-driven reasoning using patient similarity. Here data similarity is used to infer similarity in patients. Their associated diagnosis are then pooled together to form an overall diagnosis distribution.

3. RESULTS

Studies were conducted to evaluate various features of the system ranging from accuracy of view point recognition, to anomaly detection for various diseases. Here we describe the experiments relating to summarization and clinical decision support.

Figure 5 illustrates the overall summary produced by Medical Sieve in the case of a cardiac patient who entered the hospital with a chief complaint of chest pain. As can be seen, all pertinent information about the patient in terms of past medical history, recent lab results showing elevated troponin, and results of EKG test showing ST elevation were all selected by the system as relevant to creating a summary. An evolving differential diagnosis based on current evidence is also shown in Figure 5.

3.1 Evaluation of reduction in visual examination burden through summarization

We conducted a study on the effectiveness of summary generation for over 250 coronary angiography studies, each of which contained over 3000 images in the sequences across viewpoints. Using our summarization methods, 95% reduction in data was achieved with the keyframes selected by our algorithm containing a physician-identified frame in 90% of the cases. These experiments involved clinicians marking relevant frames in the whole angiogram study for purpose of ground truth generation. In case of misses, our key frames were within 2 frames of the physician-identified frames, thus establishing the utility of summarization for these studies.
Figure 5. Illustration of the web interface of the Medical Sieve Cognitive Assistant Clinical Decision Support.

Figure 6. Illustration of anomaly-driven summarization and differential diagnosis through reasoning in breast radiology.
3.2 Evaluation of time to diagnose reduction through patient record summarization

To measure the time to diagnose reduction, we compared the performance of clinicians for paper chart review versus review of patient in an earlier version of our summarization system. The results conducted on 230 patient cases showed that the average chart browsing time reduced from 0.5 hours (paper chart) to 7 minutes using our system without changing the overall diagnosis.

3.3 Evaluation of clinical reasoning

To evaluate the quality of reasoning for differential diagnosis, we conducted experiments focusing on relevant case information derived from 100 cases who have undergone mammography and/or breast ultrasound and the diseases were tumors or calcifications as seen in mammography or breast ultrasound. Figure 6 shows the result of anomaly-driven summarization and diagnostic reasoning for a patient with a chief complaint of mastodynia. The differential diagnosis before and after applying clinical reasoning is also shown in this figure on the right. Finally, the diagnosis conclusion accuracy using the clinical rules in the system and the branch-and-bound reasoning algorithm for inference is shown on the bottom right in Figure 6.

4. CONCLUSIONS

In this paper, we have presented Medical Sieve a cognitive assistant system that allows radiologists to reduce their visual examination burden through intelligent summarization, while still offering them additional clinical information about the patient to make more informed judgments about their patients. Future work will study the benefits of such a system for clinical decision support.

REFERENCES