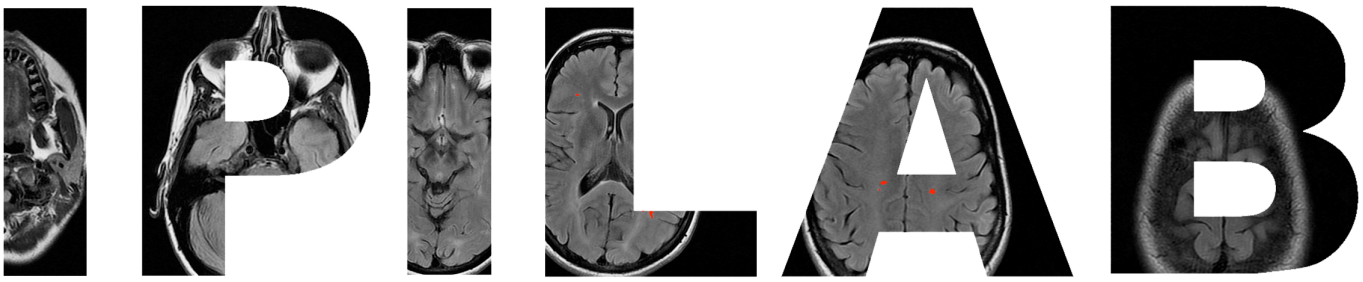


Annual Progress Report
February 2016

DEPARTMENT OF BIOMEDICAL ENGINEERING
VITERBI SCHOOL OF ENGINEERING
UNIVERSITY OF SOUTHERN CALIFORNIA

2016 Annual Report

Image Processing and Informatics Laboratory

Denney Research Center (DRB) 265, 1042 Downey Way, Los Angeles CA 90089-1111. Tel: (213) 821-8395

2015 – 2016 SUMMARY

The Image Processing and Informatics Laboratory (IPILab) has settled into the USC Biomedical Engineering Department in the Denney Research Center (DRB) 265 located on the University Park Campus. Dr. Brent Liu, IPILab Director, is also a full-time teaching professor within the Medical Imaging and Imaging Informatics track for both graduate and undergraduate courses. IPILab continues to thrive towards training a new generation of up and coming scientists and researchers interested in Imaging Informatics research. IPILab continues to bridge two schools through collaborations - Viterbi School of Engineering and the Keck School of Medicine – as well as hosting visitors interested in the Imaging Informatics research field.

Last year's milestones for past IPILab members include the following:

Ruchi Deshpande, AMI fellowship recipient and PhD student, completed her studies and final PhD defense in early February 2016. Former T32 trainee and PhD candidate Kevin Ma will be completing his studies and final defense this Spring Semester. Ximing Wang and Sneha Verma have continued their PhD research. We have two MS graduate students from the BME program who are participated in research training activities and plan to enroll in the USC BME PhD program. The USC Summer Undergraduate Research Program continues to fund our efforts to recruit and foster bright young undergraduate students searching for future academic research directions and we recruited two undergraduate researchers. IPILab continues to participate with the Engineering for Health Academy Program in conjunction with Bravo Medical Magnet High School in Los Angeles, CA to train two Senior High School students in Imaging Informatics research for the entire academic school year. IPILab hosted one Visiting Scholars, Professor Xue-Jun Zhang from the School of Computer, Electronics and Information at Guangzi University to foster collaborative research. Professor Zhang has returned to China and continues to collaborate with IPILab. The entire academic continuum from High School to Post-Graduate has been represented by the IPILab family.

We have continued in our areas of Medical Imaging Informatics research with a transition to new frontier areas of research: 1) The development of an eFolder System for Multiple Sclerosis Patients; 2) The development of imaging informatics core for large-scale stroke rehab clinical trials (eg, Interdisciplinary Comprehensive Arm Rehabilitation Evaluations – ICARE); 3) Continued development of data mining of DICOM-RT objects in conventional radiation therapy of Head and Neck cancer patients; 4) An ePR to provide decision support in evaluating does optimization in Stroke Rehabilitation (DOSE); 6) Integrating wearable sensors and imaging data in Wheelchair-bound patients. We attended the RSNA conference in December 2015 with one accepted presentation. We are continuing to transition to new areas in Rehabilitative Science and Physical Therapy since multi-media data is utilized in the research field in addition to patient-related imaging informatics data to form a new Rehab Informatics domain.

In the Table of Contents, this 2016 Annual Report includes materials related to the IPILab, IPILab R D plans and current results, selected published and in-press peer-reviewed papers during the year, as well as preprints to appear in the Proceedings of the International Society for Optical Engineering (SPIE) in Medical Imaging, San Diego, CA, February 2 -2 , 2016.

Our research has been supported by:

- NIH/NINDS/NICHD U01NS05625 (ICARE)
- NIH/NICHD R01HD065438 (DOSE)
- DOD/Loma Linda University Subcontract No. W81XWH-11-2-0151
- USC Undergraduate Research Award No. 22-1508-1030
- Northrop Grumman Enterprise Student Research Program
- ImageNation, LLC, USA

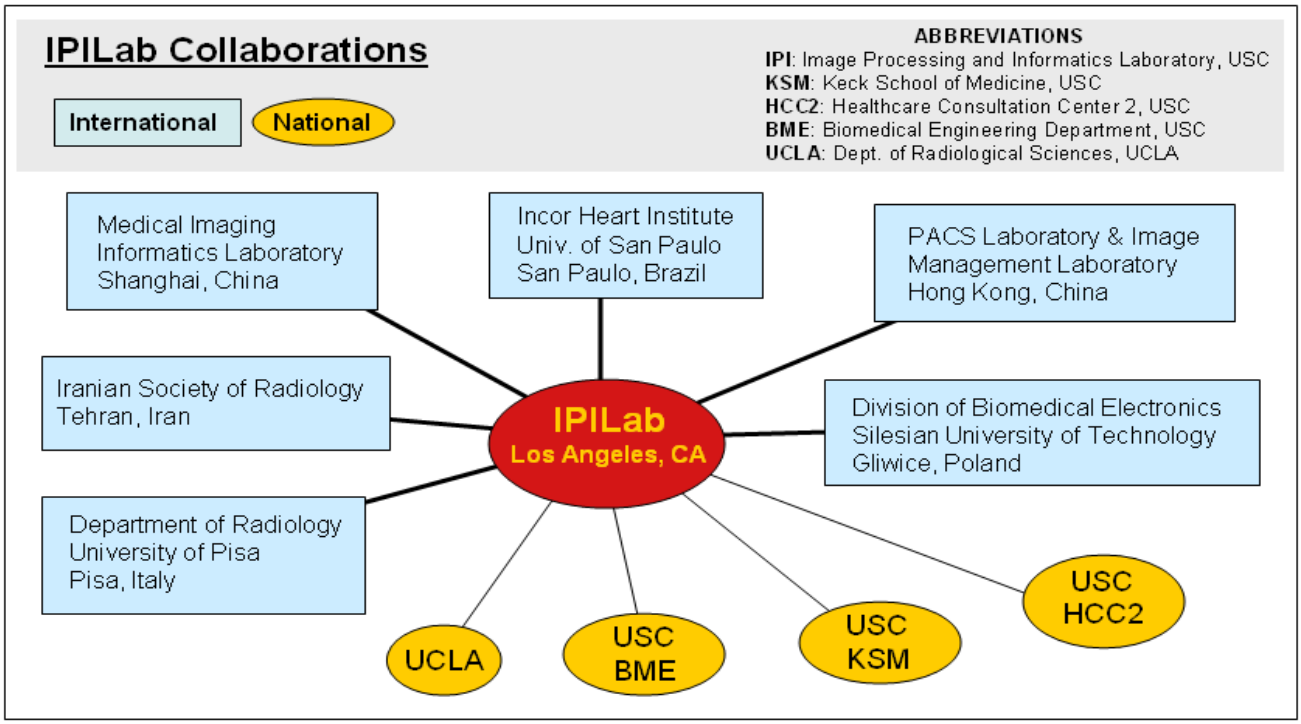
TABLE OF CONTENTS

SPIE 2016

STAFF AND COLLABORATORS

Faculty and Administration	
Michael Khoo, PhD Professor and Chairman, Department of Biomedical Engineering (BME)	**Carolee Winstein, PhD** Professor, Biokinesiology and Physical Therapy
Brent J. Liu, PhD Associate Professor of BME Director, IPILab	**Ewa Pietka, PhD, DSc** Professor, Technical University of Silesia, Poland *Visiting Professor of Radiology*
H.K. Huang, DSc, FRCR (Hon.), FAIMBE Professor Emeritus of BME & Radiology	**Jianguo Zhang, PhD** Professor, Shanghai Institute of Technical Physics, The Chinese Academy of Science *Visiting Professor of Radiology*
James Sayre, PhD Professor of Biostatistics and Radiological Science, UCLA *Consultant*	**Edward V. Grant, MD, FACR** Professor and Chairman, Department of Radiology
Maria YY Law, MPhil, BRS, PhD Consultant (Medical Imaging and Radiotherapy) Hong Kong Sanatorium and Hospital *Visiting Associate Professor of Radiology*	**Heinz U. Lemke, Professor** Technical University Berlin
Dr. Jill McNitt-Gray, PhD Professor, Departments of Biological Sciences and BME	**Dr. Philip Requejo, PhD** Adjunct Assistant Professor, BME and Kinesiology
Lilyana Amezcua, MD Assistant Professor of Neurology, Keck Hospital of USC	**Sophia Chun, MD** Chief, SCI at Veterans Health Association
Meng Law, MD Director of Neuroradiology, Keck Hospital of USC	**James Slater, MD** Radiation Medicine, Loma Linda University
	Alexander Lerner, MS Visiting Assistant Professor of Clinical Radiology, Keck Hospital of USC
Visiting Fellows and Collaborators	
Xue-Jun Zhang, PhD Visting Scholar, Guangzi University	**Jorge Documet, PhD** Software Engineer, MedQIA
Paymann Moin, MD Radiologist, Advanced Imaging Center, Valencia, CA	**Richard Lee, MD** Radiology Resident
Anh Le, PhD Instructor, UT Southwestern Medical Center	**Jasper Lee, PhD** R&D Manager, SCImage, Los Altos. CA
James Fernandez, MD Radiology Resident, USC	**Hiroshi Arai** Patent Examiner, Japan Patent Office
Graduate Student Assistants	
Kevin Ma, MS (PhD Candidate) **Ruchi Deshpande, PhD** **Ximing Wang, MS** **Sneha Verma, MS**	**Jeff Tse** **Mike Kwon** **Joseph Liu**
Undergraduate Research Interns	
Ly Pham	**Nikhil Kotha**

IPILAB ENVIRONMENT AND COLLABORATIONS



IPILAB WEBSITE

University of Southern California

USC Viterbi School of Engineering

Image Processing and Informatics Lab

Keck School of Medicine USC

Home | About | People | Research | Clinical Service | Education | Links | Contacts

About Us

The Image Processing and Informatics Laboratory (IPI) is located at 1042 Downey Way. Denney Research Center (DRB) 264, Los Angeles, CA 90089-1111

Our research facility includes PACS Simulator, Fault-tolerance Server, Data Grid, PACS workstations, CAD servers, and connections to two clinical PACS.

Reseach topics include:

- Computer Aided Detection and Diagnosis
- Data Grid and Image Archival
- Imaging Informatics Technology
- PDA Application in Clinical Environment
- Radiation Therapy Informatics
- Clinical Workflow Model
- CAD - PACS Integration Toolkit
- EPR for a surgical environment
- Multimedia ePR for Rehabilitation
- eFolder for Multiple Sclerosis Decision Support
- Spinal Cord Injury Pain Classification

NEWS AND EVENTS

IPILab has moved to USC Park Campus

October 16th, 2014

We have moved from our previous location in Annenberg Research Park to within the Department of Biomedical Engineering on the USC University Park Campus, in the Denney Research Center Building. Our new location helps to create a more convenient research and learning environment, which encourages more collaborations and sharing of ideas with other research groups in our BME department. We welcome you to visit our new laboratory and offices at:

Denney Research Center (DRB) 264
1042 Downey Way
Los Angeles, CA 90089-1111
Telephone: 213.821.8395

IPILab Update: RSNA 2013

August 20th, 2013

The IPILab have 4 abstracts accepted to RSNA (Radiological Society of North America) 2013.

The list is as follows:

Web-based DICOM-SR Viewer for CAD data of multiple sclerosis lesions in an imaging informatics-based eFolder
Authors: *Brent Liu, Kevin Ma, Jeff Zhang*

An open source, rich-client web application for visualizing DICOM RT data
Authors: *Brent J. Liu, Ruchi R. Deshpande, David Clunie, John DeMarco, Jorge Documet*

Web-based neurological pain classifier tool utilizing Bayesian decision theory for pain classification in spinal cord injury patients.
Authors: *Sneha K. Verma, Sophia Chun, Brent J. Liu*

An imaging informatics-based system with a novel intelligent workflow engine to support rehabilitation clinical trial research
Authors: *Brent Liu, Ximing Wang, Clarisa Martinez, Carolee Winstein*

We would appreciate your interests in our topics.

IPILab Update: RSNA 2011

August 24th, 2011

IPILab have 10 abstracts accepted to RSNA (Radiological Society of North America) 2011, as well as two abstracts from our collaborators.

The list is as follows:

- Extending Imaging Informatics beyond Radiology: A Web-based Multimedia System to Improve Decision Support through Movement Analysis of Elite Athletes [Educational Exhibit]
- Clinical Experiences and Challenges from the Implementation of a Zero Footprint Mobile DICOM WADO Display Solution for Smartphones and Tablets [Poster]

RSNA 2015 Poster

USC University of Southern California

RSNA 2015

Lesion Registration for longitudinal disease tracking in an imaging informatics-based Multiple Sclerosis eFolder

K C Ma[1], MS; X Zhang[1], PhD; L Amezcua[3], MD; M S Shiroishi[2], MD; A Lerner[2], MD; B J Liu[1], PhD

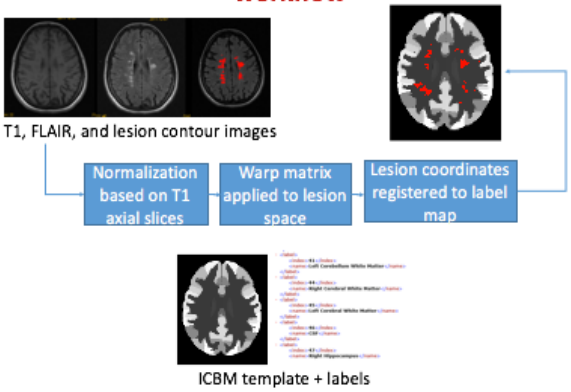
[1]*Image Processing and Informatics Lab, Dept. of Biomedical Engineering, Viterbi School of Engineering, USC*

[2]*Dept. of Radiology, Keck School of Medicine, USC* [3]*Dept. of Neurology, Keck School of Medicine, USC*

Introduction:

- Last year, we displayed big data analysis tools of MS eFolder system
  - integrating patient data with imaging and computational analysis to track patients' progress and lesion volume changes over time
  - Displaying tracking charts and analysis results on web-based user interface
- Major changes this year:
  - Introducing brain normalization and lesion registration techniques
  - Goal: locate and detect individual lesion size changes to identify where the disease is most active and/or where the treatment is most effective
  - Data mining via lesion locations and changes in patients' profiles

Brain normalization and lesion registration workflow



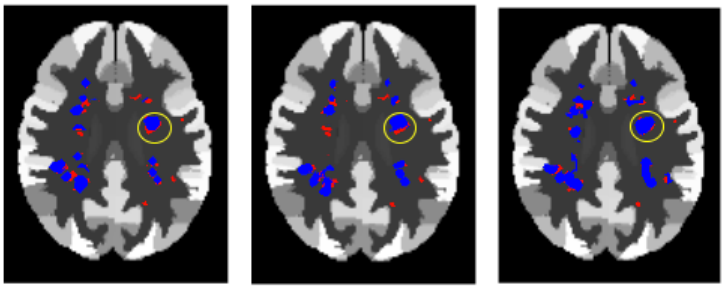
USC University of Southern California

RSNA 2015

Highlighted Features to Demonstrate:

- Lesions that exist in all 4 studies
- Changes in volumes and shapes in those lesions
- Highlight the lesions with the biggest changes in volume
- New lesions from study 1 to study 2, 3, and 4
- Old lesions that disappear from study 1 to study 2, 3, and 4
- Number of lesions and volume of lesions in each of the subcortical structures for data mining

Normalized lesion space: blue is result from current, red is result from 2009



*Yellow circle indicates lesion being tracked

Computerized Medical Imaging and Graphics 46 (2015) 227–236

Contents lists available at ScienceDirect

Computerized Medical Imaging and Graphics

journal homepage: www.elsevier.com/locate/compmedimag

Effective staging of fibrosis by the selected texture features of liver: Which one is better, CT or MR imaging?

Xuejun Zhang[a,b,c,*], Xin Gao[c,*], Brent J. Liu[b], Kevin Ma[b], Wen Yan[d], Long Liling[d], Huang Yuhong[e], Hiroshi Fujita[f]

[a] School of Computer and Electronic Information, Guangxi University, Nanning 530004, Guangxi, PR China
[b] IPILab, Biomedical Engineering Department, University of Southern California, Los Angeles, CA 90033, USA
[c] Department of Medical Imaging, Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou 215163, Jiangshu, PR China
[d] Department of Radiology, First Affiliated Hospital of Guangxi Medical University, Nanning 530021, Guangxi, PR China
[e] No. 1 People's Hospital of Nanning, Nanning 530022, Guangxi, PR China
[f] Department of Intelligent Image Information, Graduate School of Medicine, Gifu University, Gifu 501-1193, Japan

ARTICLE INFO

Article history:
Received 30 October 2014
Received in revised form 22 June 2015
Accepted 2 September 2015

ABSTRACT

Purpose: Texture patterns of hepatic fibrosis are one of the important biomarkers to diagnose and classify chronic liver disease from initial to end stage on computed tomography (CT) or magnetic resonance (MR) images. Computer-aided diagnosis (CAD) of liver cirrhosis using texture features has become popular in recent research advances. To date, however, properly selecting effective texture features and image parameters is still mostly undetermined and not well-defined. In this study, different types of datasets acquired from CT and MR images are investigated to select the optimal parameters and features for the proper classification of fibrosis.

Methods: A total of 149 patients were scanned by multi-detector computed tomography (MDCT) and 218 patients were scanned using 1.5 T and 3 T superconducting MR scanners for an abdominal examination. All cases were verified by needle biopsies as the gold standard of our experiment, ranging from 0 (no fibrosis) to 5 (cirrhosis). For each case, at least four sequenced phase images are acquired by CT or MR scanners: pre-contrast, arterial, portal venous and equilibrium phase. For both imaging modalities, 15 texture features calculated from gray level co-occurrence matrix (GLCM) are extracted within an ROI in liver as one set of input vectors. Each combination of these input subsets is checked by using support vector machine (SVM) with leave-one-case-out method to differentiate fibrosis into two groups: non-cirrhosis or cirrhosis. In addition, 10 ROIs in the liver are manually selected in a disperse manner by experienced radiologist from each sequenced image and each of the 15 features are averaged across the 10 ROIs for each case to reduce the validation time. The number of input items is selected from the various combinations of 15 features, from which the accuracy rate (AR) is calculated by counting the percentage of correct answers on each combination of features aggregated to determine a liver stage score and then compared to the gold standard.

Results: According to the accuracy rate (AR) calculated from each combination, the optimal number of texture features to classify liver fibrosis degree ranges from 4 to 7, no matter which modality was utilized. The overall performance calculated by the average sum of maximum AR value of all 15 features is 66.83% in CT images, while 68.14%, and 71.98% in MR images, respectively; among the 15 texture features, mean gray value and entropy are the most commonly used features in all 3 imaging datasets. The correlation feature has the lowest AR value and was removed as an effective feature in all datasets. AR value tends to increase with the injection of contrast agency, and both CT and MR images reach the highest AR performance during the equilibrium phase.

Conclusions: Comparing the accuracy of classification with two imaging modalities, the MR images have an advantage over CT images with regards to AR performance of the 15 selected texture features, while 3 T MRI is better than 1.5 T MRI to classify liver fibrosis. Finally, the texture analysis is more effective during equilibrium phase than in any of the other phased images.

* Corresponding author at: School of Computer, Electronics and Information, Guangxi University, 100, Daxue Dong Road, Nanning 530004, Guangxi, PR China. Tel.: +86 771 3231260; fax: +86 771 3232214.
E-mail addresses: xjzhang@gxu.edu.cn (X. Zhang), xingaosam@yahoo.com (X. Gao).

http://dx.doi.org/10.1016/j.compmedimag.2015.09.003
0895-6111/

1. Introduction

Patients suffering from chronic liver diseases are at a severe risk from complications such as hepatocellular carcinoma (HCC) and liver failure [1]. In the diagnosis and therapeutic assessment of cirrhosis of the liver and chronic hepatitis, it is necessary to stage the degree of hepatic fibrosis as an important indicator of cirrhosis as well as a critical predictive factor for the occurrence of HCC [2], which is one of the most common malignancies in patients that are affected by these diseases [3]. Fibrosis is caused by excessive deposition of extracellular matrix owing to histological and molecular reshuffling of various components such as collagens, glycoproteins, proteoglycans, and other macromolecules within the extracellular matrix. These features, common to almost all patients with chronic liver disease, lead to the changes in the hepatic morphology, texture pattern, and degree of liver stiffness. Accurate assessment of hepatic fibrosis is crucial for the determination of the appropriate treatment because fibrosis is potentially a reversible process in the early stages.

Liver biopsy, which is used for histological scoring and is still used as a reference test for fibrosis staging, is considered the gold-standard method for the assessment of the degree of fibrosis. However, liver biopsy is an invasive procedure with possible side effects such as pain in 30–40% of the cases, pneumothorax (3‰) or even death (2/10,000) [4,5]. To decrease the need for painful biopsies, non-invasive methods using MR and ultrasound imaging have been proposed to obtain images of the liver. With the development of high-speed imaging devices, highly precise medical imaging services are now widely available. Recently, radiological assessments of hepatic fibrosis by magnetic resonance elastography (MRE) [6–8], gadolinium- or superparamagnetic iron oxide-enhanced MR imaging [9,10], diffusion-weighted MR imaging [11], and real-time ultrasonographic elastography [12] have been reported, and their feasibility, usefulness, and limitations have been shown. Wang et al. [13] have reported on real-time elastography with a new quantitative technology for diffusing histological lesions as a new and promising sonography-based non-invasive method for assessing liver fibrosis in patients with chronic hepatitis B. These new imaging technologies have greatly impacted the traditional diagnostic methods. However, the interpretation of the numerous diseases from different types of medical images is not an easy task, especially for inexperienced residents or general radiologists. In the last decade, different types of computer-aided detection/diagnosis (CAD) systems have been developed to ease the workload of radiologists. There are several promising reports on the CAD of liver fibrosis on MRI, which have been obtained by analyzing morphology changes [14,15] and the texture pattern of fibrosis [16] through images.

Texture patterns of hepatic fibrosis are one of the important features to diagnose the chronic liver disease from initial to end stage on computed tomography (CT) or magnetic resonance (MR) images. However, quantification and classification of hepatic fibrosis with only texture patterns in liver is an extremely difficult task for both the radiologist and computer, resulting in a low accuracy rate of diagnosis. However, subtle differences of tiny structures of fibrosis within a small size of ROI provide the computer with an advantage to the reader as compared with the clinical interpretation using other image features such as shape, volume, surface irregularity, etc. Some work has been developed utilizing intelligent algorithms to stage the fibrosis by feeding the features calculated from 1st and 2nd order texture matrixes into a classifier, and stating their promising results on different kinds of datasets. Gobert et al. [17] reported the use of six features extracted from texture features and two statistical measures as input to an ANN classifier with genetic algorithm, resulting in 73% correct classification rate for the characterization of cirrhotic and non-cirrhotic in MRI images. Li et al. [18] found that integrating the texture features into shape features with ANN classification of cirrhosis may improve the overall performance of cirrhosis detection. However, these studies used samples of a normal and a cirrhotic case and the selected features were based on experience. Other improved classification methods state their effective findings of parameters [19], but did not provide a “real true” selection of features that should be recommended to doctors as their investigation always works on individual sequenced images or modality. In addition, the optimal number and effectiveness of input vectors have not been addressed or identified. These may be due to the limitation of case numbers and large time consumption needed to optimize each combination of features. Although some efficient training methods were proposed to reduce the time in training and testing procedures, their effects are difficult to compare and evaluate as there was no evaluation against the gold standard until now.

China has a high incidence rate of hepatic cancer along the coastal areas, especially in Fushui county of the Guangxi autonomous region, where high liver cancer incidence rate of 56.45/100,000 in average has been reported and for male 91.11/100,000 and female 21.80/100,000, respectively [20]. Therefore, collecting a large amount of fibrotic cases is feasible in our local hospitals. Several hundred patients with different grade distribution of fibrosis were scanned in at least four different sequenced phases using various CT and MRI modalities. In order to find out the “real true” features that impact the classification of texture patterns of hepatic fibrosis utilized in computer-aided diagnosis, we chose the most accurate validation method of leave-one-case-out to check each combination of features. Considering that $2^{15}-1$ different combinations are obtained in a single phased image, the iteration is large and time consuming which would estimate the overall time length to perform the experiment at over 3 years. The accuracy rates outputted from different modalities, sequenced phased images and texture features could not only tell us the “true” of optimal set of fibrotic texture patterns guiding the radiologists to analyze fibrosis efficiently from numerous medical images, but also help to clarify how many and what kind of texture features should be utilized for CAD [21].

This study is focused on the effective selection of texture features for fibrosis classification. We utilized 15 texture features calculated from all of the CT datasets with slice thickness from 0.625 mm to 5 mm, and over four different sequenced CT or MR phased images are investigated. Each combination of features is evaluated using support-vector-machine (SVM) with leave-one-case-out method to select the optimal feature subsets according to their performance. In this paper, Section 2 describes the acquirement of experimental datasets by CT and MR modalities with their imaging protocols. Section 3 describes the methods of the ROI selection, feature extraction and optimization by SVM and leave-one-case-out approach. In Section 4, classification results from the different imaging modalities are demonstrated and discussed. The accuracy rate of each combination of feature subsets is compared to find out the most informative features as well as the optimal number of texture features from the experimental results. The paper concludes in Section 6 with further discussion of the proposed approach and future works.

2. Experimental materials

To evaluate and compare the contribution of texture features derived from CT and MR images towards the diagnosis of liver fibrosis, a total of 367 patients with or without hepatic fibrosis are scanned by 3 different modalities from two different hospitals, including 149 cases from CT and 218 cases from MR.

Table 1
Patients statistic with five fibrous stages according to VHPTA system [22].

Stage of fibrosis	Score	Description of hepatic fibrosis degree	Number of MRI cases (hospital no. 1)	Number of MRI cases (hospital no. 2)	Number of CT cases (hospital no. 1)	Groups
Normal	0	No fibrosis	40	5	36	Non-cirrhosis (Group 0)
Mild fibrosis	1	Fibrous expansion of some portal areas, with or without short fibrous septae	30	4	26	
	2	Fibrous expansion of most areas and bridging fibrosis appeared	23	10	13	
Severe fibrosis	3	Most fibrous septum and lobular structure disorder	18	8	19	Cirrhosis (Group 1)
	4	Early cirrhosis, diffuse fiber hyperplasia	22	16	19	
Cirrhosis	5	Cirrhosis	40	2	36	

2.1. Computed tomography (CT) image

A total of 149 patients were scanned by CT (GE Lightspeed VCT) from June 2009 to March 2012 at the Department of Radiology, First Affiliated Hospital of Guangxi Medical University (Hospital no.1). The data set consists of 36 normal cases, 39 mild fibrosis cases, 38 severe fibrosis cases and 36 typical cirrhosis cases. The imaging protocol is as follows: quad-phase scans are made at 120 kV tube voltage, 250 mA tube current; the image size is 512×512 pixel. The arterial, portal venous and equilibrium phases are acquired at 25, 60, and 120 s after contrast injection, respectively. The injection speed is 3.0 ml/s, concentration is 320 mg/ml, dosage is 85–90 ml. For each case, five different CT datasets with different slice thickness are acquired with interval of 0.625 m, 1.25 mm, 2.5 mm, 3.75 mm and 5 mm, respectively. For a case with one thickness, there are four sequenced phase images: pre-contrast, arterial, portal venous and equilibrium phase. Therefore, a total of 20 sequence/series types were acquired in one CT case. All cases have been verified by needle biopsies as the gold standard of our experiment. Surgical specimens were retrospectively examined by a pathologist who was blinded to patient histories and radiology and surgery reports. The patients' fibrosis stages are evaluated in accordance with the Chinese Viral Hepatitis Prevention and Treatment Plan (VHPTA) [22], ranging from 0 (no fibrosis) to 5 (cirrhosis). In this paper, the entire dataset is divided into two groups: non-cirrhosis (containing normal cases S0 and mild liver fibrosis cases S1, S2); and cirrhosis (containing severe liver fibrosis cases S3, S4 and typical cirrhosis cases S5), as illustrated in Table 1.

2.2. Magnetic resonance (MR) image

From February 2011 to March 2012, a total of 173 patients had an abdominal examination performed using a 3-T superconducting MR scanner (Intera Achieva Quasar Dual; Philips Medical Systems, Netherlands) with a six-channel torso array coil at Hospital no. 1. The imaging protocols used include TR/TE, 2.3/82 ms in plan scanning phase; 7 mm slice thickness and 1 mm inter-slice gap; interpolated imaging matrix, 512×512; slice thickness; 4 mm and 2 mm inter-slice gap. In dynamic contrast scanning phase: 7 mm slice thickness and 1 mm inter-slice gap; contrast injection, Gd-DTPA; the dosage, 0.2 mmol/kg; injection speed, 2–3 ml/s. Since contrast injection after 16 s, 50–60 s and 2–3 min arterial, portal and equilibrium scanning are performed, respectively. Among the datasets, there are 18 patients without liver disease history and hepatic dysfunction, and the fibrosis stage of other patients is confirmed by the result of the liver biopsy procedure.

An additional 45 MR cases were acquired from the No.1 People's Hospital of Nanning (Hospital No. 2) in China using a 1.5 T (Signa Horizon, GE Healthcare) MR scanner and were included in our experiment. The gadolinium-enhanced T1-weighted images (150/1.6; matrix, 512×512; flip angle, 90°; signal acquisition, 1; slices, 18 per 26-second acquisition time) sequences are with section thickness 8 mm with a 2 mm intersection. Table 1 gives the detailed distribution of the 3 datasets. This study was approved by the institutional review board at both hospitals and informed consent was obtained from all patients.

3. Methods

In this section, we give a brief overview of the gray level co-occurrence matrix (GLCM), Support Vector Machine (SVM) and Leave-one-out cross-validation methods, including the guideline of ROIs selection and pre-processing procedures for feature extraction. We then describe a texture features optimizing method for the analysis of the hepatic fibrosis on CT/MR images.

3.1. Selecting ROIs for feature extraction

We have been developing a CAD system named *CirrhoView* for cirrhosis clinical research. Fig. 1 is the graphical user interface (GUI) of our software that enables the radiologists to open DICOM files on a server; manually select ROIs by moving and clicking the mouse button; automatically calculating texture features within the ROI and store to an excel file for further analysis. In this experiment, 10 ROIs in liver were manually picked in a disperse fashion from each sequenced image by L. Long, a diagnostic radiologist with over 20 years' experience of radiological images interpretation. The ROIs are shown as red squares in Fig. 1. Selecting ROIs relies on the guidelines illustrated in Table 2: 1 ROI selected in left medial segment, right anterior lateral segment and right posterior lateral segment, respectively, at the second porta hepatis; 2 ROIs selected in right anterior lateral segment and right posterior lateral segment, respectively, at the first porta hepatis; 1 ROI selected in right anterior lateral segment and right posterior lateral segment, respectively, below the first porta hepatis. Considering the diffuse distribution of liver, large blood vessels and focal liver lesions within the liver are excluded. Three sizes of the ROIs (32×32, 20×20, 16×16 pixel) are generated individually according to the

Table 2
Guideline of selecting ROIs relied on the location of liver.

Number of ROIs	Position of ROIs in hepar	Hepatis partition
2	Right anterior lateral segment	First porta hepatis
2	Right posterior lateral segment	
1	Right anterior lateral segment	Below the first porta hepatis
1	Right posterior lateral segment	
1	left medial segment	Second porta hepatis
1	Right anterior lateral segment	
1	Right posterior lateral segment	
1	Left lateral segment	

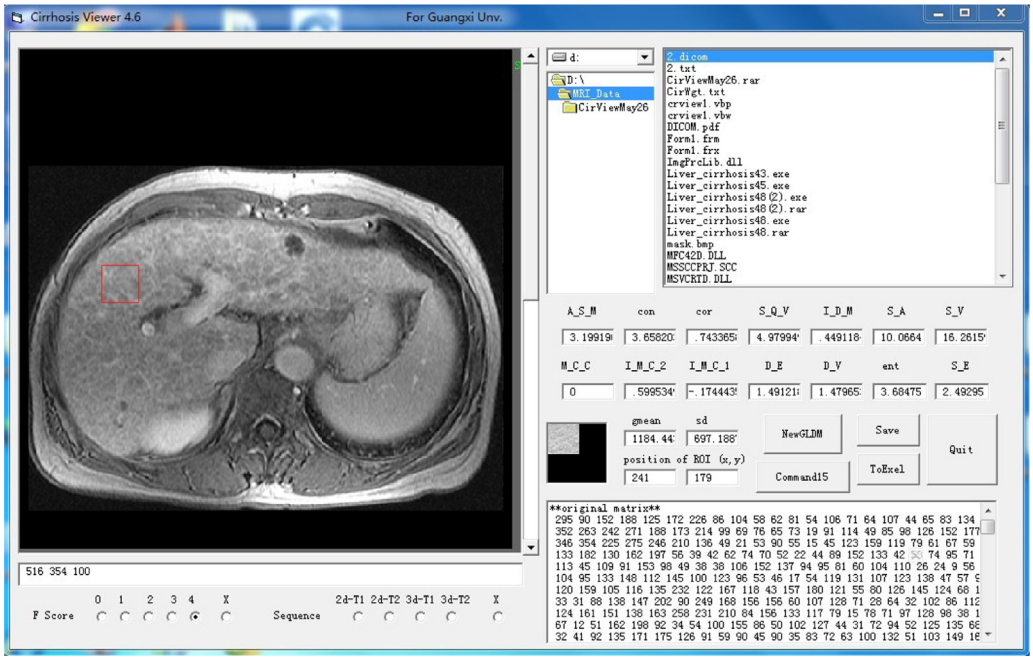


Fig. 1. The graphical user interface of our software CirrhoView, designed for selecting ROIs by radiologists and automatic feature extraction by CAD system. The red square of ROI indicates a texture pattern of liver cirrhosis. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

start point of a square, and all ROIs are pre-processed with or without Sobel filter. Thus we have 10 ROIs with 3 sizes, processed by 2 different filtering procedures, and in 4 phases for calculating features in a single sequenced CT/MR image. Considering there are 5 available slice thicknesses in one CT case, the total number of ROIs reaches to 1200 as illustrated in Fig. 2. Due to the technical difficulty and the focus on this study, only one slice thickness is used in MRI.

3.2. Gray level co-occurrence matrix and texture features

The texture features are obtained by gray level co-occurrence matrix (GLCM) method. As the DICOM images have 12-bit gray levels, which is too large for establishing co-occurrence matrices, the gray levels of an image are reduced to 4-bit, which is accurate enough for this study [17], and the matrix element in row i and column j are denoted:

$$P(i,j,\delta,\theta) = \{[[(x,y),(x+\Delta x,y+\Delta y)]|] \\ f(x,y)=i, \quad f(x+\Delta x,y+\Delta y)=j; \quad x,y=0,1,..,N-1\} \tag{3-1}$$

where $i,j=16$; $\delta=1$; $\theta=0°,45°,90°,135°$.

Assume that P is the matrix of GLCM extracted from ROI, G is the gradation order of GLCM, μ is the mean value of the matrix., μ_x, μ_y, σ_x, σ_y are the means and standard deviations of P_x and P_y, respectively. $P_x(i)$ is the sum of row i of the matrix P, while $P_y(j)$ is the sum of column j of P.

$$P_x(i) = \sum_{j=0}^{G-1} P(i,j)$$

$$P_y(j) = \sum_{i=0}^{G-1} P(i,j)$$

$$\mu_x = \sum_{i=0}^{G-1} i \sum_{j=0}^{G-1} P(i,j) = \sum_{j=0}^{G-1} iP_x(i)$$

$$\mu_y = \sum_{i=0}^{G-1} j \sum_{j=0}^{G-1} P(i,j) = \sum_{j=0}^{G-1} jP_y(j)$$

$$\sigma_x{}^2 = \sum_{i=0}^{G-1} (i-\mu_x)^2 \sum_{i=0}^{G-1} P(i,j) = \sum_{i=0}^{G-1} (P_x(i)-\mu_x(i))^2$$

$$\sigma_x{}^2 = \sum_{j=0}^{G-1} \left(i-\mu_y\right)^2 \sum_{i=0}^{G-1} P(i,j) = \sum_{j=0}^{G-1} \left(P_y(j)-\mu_y(j)\right)^2$$

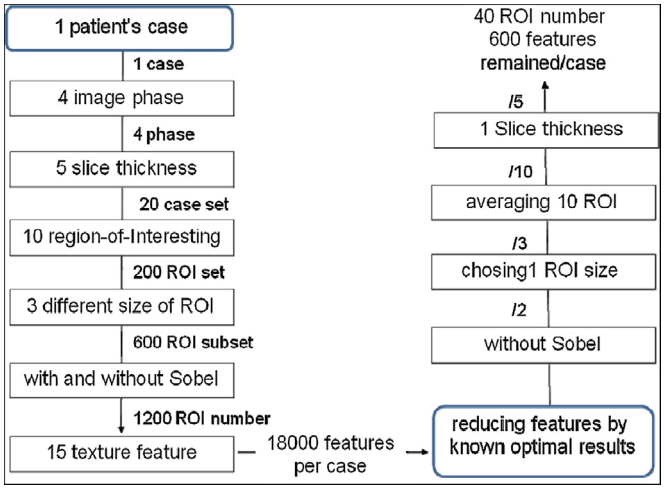


Fig. 2. The evolution of number of features in one CT case demonstrates the input features generated for initial optimization (left) and reduced by some previous investigations (right).

Thirteen texture features are calculated by the method introduced by *Haralick* [23] based on the co-occurrence matrix:

Angular second moment (ASM):

$$\mathrm{ASM} = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\left\{P(i,j)\right\}^2 \tag{3-2}$$

Contrast (CNT):

$$\mathrm{CNT} = \sum_{n=0}^{G-1} n^2 P_{x-y}(n) \tag{3-3}$$

where

$$P_{x-y}(k) = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1} P(i,j) \quad |i-j| = k \quad k = 0, 1, 2, \ldots G-1$$

Correlation (COR):

$$\mathrm{COR} = \sum_{i=1}^{G}\sum_{j=1}^{G} \frac{\left\{i \times j\right\} \times P(i,j) - \left\{\mu_x \times \mu_y\right\}}{\sigma_x \times \sigma_y} \tag{3-4}$$

Entropy (ENT):

$$\mathrm{ENT} = -\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} \left\{P(i,j)\right\} \log\left(P(i,j)\right) \tag{3-5}$$

Sum of squares: Variance (SQV):

$$\mathrm{SQV} = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1} \left\{P(i,j)\right\} \log\left(P(i,j)\right) \tag{3-6}$$

Inverse Difference Moment (IDM):

$$\mathrm{IDM} = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1} \frac{1}{1+(i+j)^2} P(i,j) \tag{3-7}$$

Sum Average ($\mathrm{SUM_{AVE}}$)

$$\mathrm{SUM_{AVE}} = \sum_{i=0}^{2G-2} i P_{x+y}(i) \tag{3-8}$$

where

$$P_{x+y}(k) = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1} P(i,j) \qquad i+j = k, \quad k = 0, 1, 2, \ldots, 2(G-1)$$

Sum Variance ($\mathrm{SUM_{VAR}}$):

$$\mathrm{SUM_{VAR}} = \sum_{i=0}^{2G-2} (i - \mathrm{SUM_{AVE}})^2 \; P_{x+y}(i) \tag{3-9}$$

Sum Entropy ($\mathrm{SUM_{ENT}}$):

$$\mathrm{SUM_{ENT}} = \sum_{i=0}^{2G-2} P_{x+y}(i) \log\left(P_{x+y}(i)\right) \tag{3-10}$$

Difference Variance ($\mathrm{DIF_{AVE}}$):

$$\mathrm{DIF_{AVE}} = -\sum_{I=0}^{G-1} \left\{ i - \sum_{I=0}^{G-1} i P_{x+y}(i) \right\}^2 P_{x-y}(i) \tag{3-11}$$

Difference Entropy ($\mathrm{DIF_{ENT}}$):

$$\mathrm{DIF_{ENT}} = -\sum_{i=0}^{G-1} P_{x-y}(i) \log\left(P_{x-y}(i)\right) \tag{3-12}$$

Information Measures1 (IMC1):

$$\mathrm{IMCI} = \frac{\mathrm{ENT} - \mathrm{HXY1}}{\max\left(\mathrm{ENT}_X, \mathrm{ENT}_Y\right)} \tag{3-13}$$

where

$$\mathrm{ENT}_x = -\sum_{i=0}^{G-1} P_x(i) \log\left\{P_x(i)\right\}$$

$$\mathrm{ENT}_y = -\sum_{j=0}^{G-1} P_y(j) \log\left\{P_y(j)\right\}$$

$$\mathrm{HXY1} = -\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} P(i,j) \log\left\{P_x(i) P_y(j)\right\}$$

$$\mathrm{HXY2} = -\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} P_x(i) P_y(j) \log\left\{P_x(i) P_y(j)\right\}$$

Information Measures2 (IMC2):

$$\mathrm{IMC2} = \sqrt{1 - \exp\left\{-2.0\left(\mathrm{HXY2} - \mathrm{ENT}\right)\right\}} \tag{3-14}$$

In our experiment, two features based on first order statistics of the image: (14). Mean gray value (MGV) and (15). Standard deviation (SD) are also used in this study, thus there are total 15 texture features used in a ROI.

$$\mathrm{MGV} = \frac{1}{N^2}\sum_{x=1}^{N}\sum_{y=1}^{N} f(x,y) \tag{3-15}$$

$$\mathrm{SD} = \sqrt{\frac{1}{N}\sum_{x=1}^{N}\sum_{y=1}^{N}\left\{f(x,y) - \mathrm{MGV}^2\right\}} \tag{3-16}$$

3.3. Support vector machine (SVM)

SVM originated on the basis of statistics by Vapnik et al. [24] is a current general learning method. The discrimination function for linear separable problem is:

$$f(X) = \sum\nolimits_{i=1}^{N} y_i a_i * k(X_i, X) + b* \tag{3-17}$$

where N is the number of support vector, X_i, y_i are the labels of corresponding support vectors, and $\alpha_i{}^*$ and b^* are parameter learned from training samples. The kernel $k(Xi, X)$ is significant for determining the behavior of the classifier. In this study, the Radial Basis Function (RBF) kernel function is used:

$$k(X, Y) = \exp\left\{\frac{\left|X - Y\right|^2}{2\sigma^2}\right\} \tag{3-18}$$

The $|x\text{-}y|$ represents the distance between two vectors, where σ is a constant 1. The input vectors to SVM are 15 texture features in this study while the output is probability for the presence of cirrhosis.

3.3.1. Leave-one-out cross-validation (LOOCV) method

The practical goal of using the SVM model would be to determine which subset of the 15 texture features should be used to produce the best predictive model. For most modelling procedures, if we compare feature subsets using the in-sample error rates, the best performance will occur when all 15 features are used. However

under cross-validation, the model with the best fit will generally include only a subset of the features that are deemed truly informative.

The Leave-one-out (LOO) [25] method, as the name suggests, is to take out a single observation from the original datasets containing M samples as the validation data, and the remaining $M-1$ samples as the training data to build a classification model of SVM. After validating this model, the sample will be moved back into the training data set and another sample is selected. The procedure is repeated such that each observation in the sample is used once as the validation data. After looping a total of M times of training and testing on M samples, all of the cases in one dataset are validated by the SVM model. In this study, the number of samples M is selected as 50, 50, 18, corresponding to CT set, 3.0 T MR and 1.5 T MR sets.

Leave-one-out cross-validation is typically used in the analysis of very small datasets. Although usually the LOOCV method is expensive to train since such training must be carried out repeatedly, it is still the most robust and accurate way to evaluate a predictive model. In some studies, a few modified methods such as the least squares and kernel regression, cross-validation can be sped up significantly by pre-computing certain values that are needed repeatedly in the training, or by using fast updating rules. However, careful attention is needed to preserve the "total blinding" of the validation set from the training procedure, otherwise bias may result. In order to develop the modelling procedures accurately, we choose the traditional LOOCV method in our experiment.

Among the testing results, the ratio between the number of correctly classified cases as cirrhosis and total number of cirrhosis cases is defined as true positive (TP) while the ratio between the number of correctly classified cases as non-cirrhosis and total non-cirrhosis cases is defined as true negative (TN). The average of TP and TN is defined as the accuracy rate (AR).

3.4. Calculation of texture features with different conditions

Texture features will present different values if calculated from varying pre-processing methods, image quality or ROI sizes. Therefore, we should optimize each of these conditions before analysing and optimizing feature subsets.

First, the accuracy rate detected w/o a Sobel Filter is tested in a sequenced dataset to determine the best input pattern of SVM, and half of the datasets will be excluded with this result [18]. Second, for CT image sets, accuracy rate calculated from 5 different slice thicknesses are obtained by testing the samples from portal venous phased images. Only one slice thicknesses with highest rank will be kept in further study and other data is excluded [19]. This step is skipped in MR image sets with single slice thickness. Third, one optimal size is chosen among the three different ROI sizes [26].

Since the numbers of ROIs selected are too large, 15×10 texture features in one sequenced case are reduced by averaging the features value from 10 selected ROIs, respectively, and turned into a new dataset with only 15 texture features. After all of the above processing steps, the total number of ROIs in a case is reduced from 1200 to 40 as shown in Fig. 2. This step may significantly decrease the number of input samples to ensure SVM efficient training.

3.5. Optimizing the features

Besides the above optimizing problem, there remains the challenge to optimize the 15 texture features, e.g. how many and which ones are the best features to represent fibrosis? In this paper, the data samples are divided into two categories in the SVM classification model: the mild liver fibrosis cases S0, S1, S2 as the negative samples group 0, while S3, S4, S5 severe liver fibrosis cases and typical cirrhosis cases as the positive samples group 1. Among sequenced datasets, each possible combination of 15 features is performed by the SVM model with LOOCV method to compare their accuracy. The method is executed as the following steps in detail:

- Step 1: Non-cirrhosis and cirrhosis groups are read into Group 0 and Group 1, respectively, containing M samples in each group. 15 features are extracted by GLCM and averaged over 10 chosen ROIs. ($2M \times 15$ features as input candidates).
- Step 2: The number of input items n is selected from the combinations of 15 features. ($2^{15}-1$ different combinations obtained, where $n \in [1,15]$).
- Step 3: According to the selected n items in a subset from Step2, $2M \times n$ input features are picked up from Group 0 and Group 1, respectively. A single sample with n features is chosen as the validation data, and the remaining $2M-1$ samples are used as the training data.
- Step 4: SVM classifier is modelled with $2M-1$ training samples and validates the remained test sample k. If k belongs to Group 0 and is correctly classified, then TN = TN + 1; likewise, if Group 1 is correctly classified, then TP = TP + 1.
- Step 5: If there are undetermined samples within $2M$ then run Step 3; if not, calculate the accuracy rate (AR): AR0 = TN/M, AR1 = TP/M, and AR = (AR0 + AR1)/2.
- Step 6: If the ergodic process does not finish, then run Step 2, otherwise, run Step 7.
- Step 7: Complete. $2^{15}-1 = 32{,}767$ numbers of AR are produced for further processing.

3.6. Evaluation of the performance of texture features

Each sequenced datasets has 32,767 numbers of AR values that can be divided into 15 groups $[AR^1, AR^2, \ldots, AR^i, \ldots, AR^{15}]$ according to the numbers of selected items in a subset. Each AR^i also has numbers of subsets by the combination of i items out of 15 features, among which a maximum value of AR^i is defined as Max AR^i. Since the CT and MR datasets contain 4 phased images, the maximum value of AR^i in a certain jth sequenced datasets is defined as $Max\ AR_j{}^i$. In addition, an average sum of maximum value in a whole m sequenced datasets acquired by one modality k can be calculated as

$$AR_k = \left[\frac{1}{m}\sum_{j=1}^{m} \mathrm{MaxAR}_j^1, \frac{1}{m}\sum_{j=1}^{m} \mathrm{MaxAR}_j^2, \ldots, \frac{1}{m}\sum_{j=1}^{m} \mathrm{MaxAR}_j^i, \ldots, \frac{1}{m}\sum_{j=1}^{m} \mathrm{MaxAR}_j^n \right] \quad (3\text{-}19)$$

where k = 1:CT, 2: MRI 3T, 3: MRI 1.5T.

4. Results and discussion

The grade accuracy rate varies with different imaging conditions and combinations of different number of texture features. In our previous study [18], the performance of classification of fibrosis on MRI w/o Sobel filter was investigated. The result showed a best variation accuracy of 0.76 without Sobel pre-processing vs. 0.64 with filter. Comparing with the high score of using raw data directly, Sobel filter is insignificant in pre-processing the MR or CT images. This may be due to the fact that the Sobel filter not only enhances the fibrosis pattern but also the background noises. We also concluded that different sizes of ROI may impact the classification, and 20×20 is the optimal size among the three ROI sizes [26]. Therefore, this experiment was performed by using 20×20 ROI raw data. For the CT datasets, the optimal size of CT image slice thickness is 1.25 mm, which demonstrates the fibrosis information

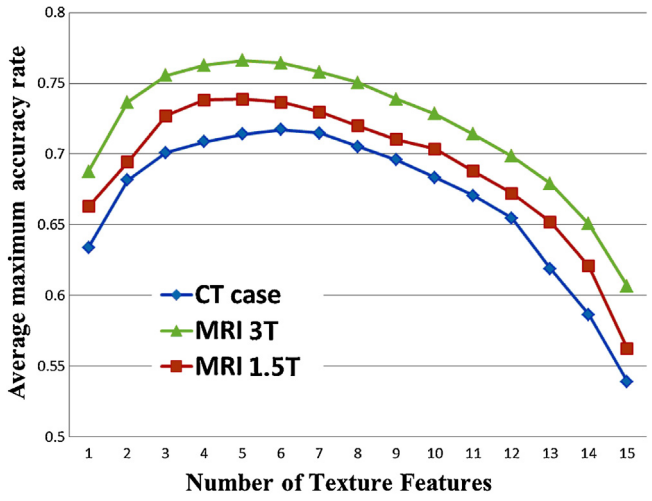


Fig. 3. Accuracy Curve by selection of number of features indicates the different performance on 3 different modalities. Where ▲ MR images from Hospital no. 1; ■ MR images from Hospital no. 2; and ◆ CT images from Hospital no. 1.

in this interval is the most adequate on CT image. Likewise, the 20×20 ROI raw data with 1.25 mm interval is the default setting for this experiment on CT datasets.

4.1. Selecting optimal number of features

In order to determine a subset of features that are deemed truly informative, LOOCV is performed on each of the sequenced imaging datasets. In this modelling procedure, every combination of 15 features is generated to provide the best AR from 1 to 15 feature subsets in Table 3. As shown in Fig. 3, the optimal number of texture features to classify liver fibrosis degree is from 4 to 7, no matter what modalities are used. The best testing average sum of maximum value of AR are 71.72% and 76.61%, obtained by 6 features from CT and MRI (3.0T) images at hospital no. 1, respectively, and 73.22% in case of 5 features from MRI (1.5 T) at hospital no. 2. The overall performance calculated by the average sum of the maximum value of all 15 features is 66.83% in CT images, with 68.14% (1.5 T), and 71.98% (3 T) in MR images, respectively. It is obvious that the AR of MR images is better than that of CT images by comparing the three curves in Fig. 3. The results also imply that 3 T MR has better image quality for classifying fibrosis than that of 1.5 T, as both average sum of maximum value of AR's in MR are higher than CT. We should notice that the fibrotic texture is better observed on MR than on CT images in clinical practice and the results reflect this trend. In addition, the lowest average sum of maximum value of AR indicated in Table 3 occurs when using all of the 15 features together to build the predictive SVM model, which implies the importance of selecting informative features.

4.2. Distribution of informative features

From the results of the average sum of maximum values of AR, we have concluded that the number of feature subset ranging from 5 to 7 features has better performance than other numbers. Another interesting point to investigate is what these 5 to 7 features actually are. A threshold value of AR value is set to 0.7 so as to pick the subset features having excellent performance. If an AR value is >0.7, then the number of each of the items in this AR is plus 1. Based on this statistics, a discriminative power histogram of every feature contributed to the classification is generated in Tables 4a and 4b and shown in Fig. 4 after normalizing the scale into a range of [0–1].

Table 3
AR values of MRI and CT by using different number of features as input vectors of SVM.

Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
MRI3T	0.687	0.737	0.755	0.763	0.766	0.764	0.758	0.750	0.739	0.728	0.714	0.699	0.679	0.651	0.607
MRI1.5T	0.663	0.694	0.727	0.738	0.739	0.737	0.730	0.720	0.710	0.704	0.688	0.672	0.652	0.621	0.562
CT	0.634	0.682	0.701	0.709	0.714	0.717	0.715	0.705	0.696	0.683	0.670	0.655	0.619	0.587	0.539

Table 4a
Value of feature importance in MRI and CT modalities (1).

Feature	ASM	CNT	COR	SQV	IDM	SUM_{AVE}	SUM_{VAR}	SUM_{ENT}
MRI3T	0.644	0.034	0.000	0.070	0.647	0.736	0.006	0.484
MRI 1.5T	0.213	0.568	0.000	0.491	0.476	0.315	0.502	0.698
CT	0.164	0.375	0.000	0.347	0.474	0.147	0.253	0.243

Table 4b
Value of feature importance in MRI and CT modalities (2).

Feature	ENT	DIF_{AVE}	DIF_{ENT}	IMC1	IMC2	MGV	SD
MRI 3T	0.581	0.026	0.061	0.452	0.007	1.000	0.336
MRI 1.5T	0.596	0.225	0.253	0.227	0.282	0.614	1.000
CT	0.386	0.399	0.330	0.284	0.163	1.000	0.475

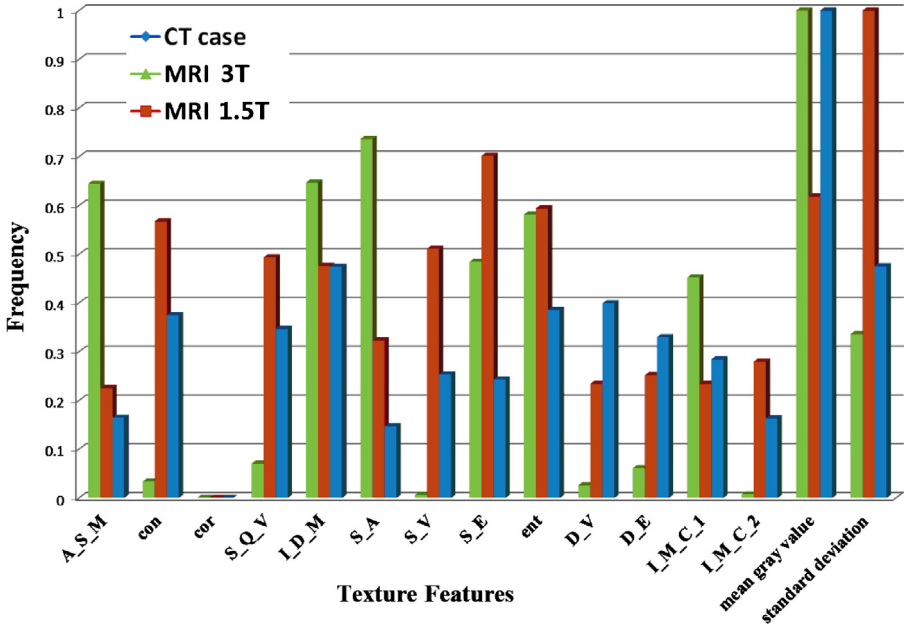


Fig. 4. From the discriminative power histogram for all of the 15 characteristics extracted from 3 different datasets, the ability of each feature in the differentiation of fibrous texture is investigated by counting the frequency of high scores appeared in AR values.

Fig. 5 indicates the top 7 texture features extracted in CT images and MR images in the discrimination power histogram. The red circle contains the top 7 features extracted in CT images: mean gray value (MGV), standard deviation (SD), inverse difference moment (IDM), difference variance (DV), entropy (ENT), contrast (CON), and sum of squares variance (SQV). The blue circle contains the top 7 features extracted in MR images from hospital no. 2: SD, sum entropy (SE), MGV, ENT, CON, sum average (SV), SQV. On the other hand, the green circle contains the top 7 features extracted in MR images from hospital no. 2: MGV, sum average (SA), IDM, angular second moment (ASM), ENT, SE, and information measures of correlation1 (IMC1). It should be noted that among the 15 texture features, mean gray value and entropy are the most commonly included as high performing in all 3 imaging datasets. With the development of fibrosis, more and more fibrous pattern appear in liver region. This phenomenon significantly changes the intensity and the complexity of an image/ROI, which corresponds to the features of MGV and ENT, respectively. Other features such as standard deviation, contrast, sum of squares variance, inverse difference moment, and SE are selected by 2 datasets. Correlation (COR) has the lowest AR value and is removed the feature set in all image datasets. Selecting features based on this ranking will be more efficient in future studies.

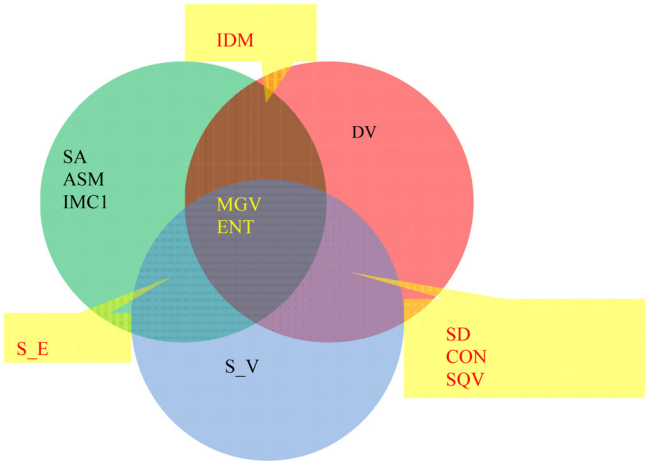


Fig. 5. The distribution of features used in 3 datasets with top ranking 7 from Fig. 3.

4.3. Performance of different phased images

AR values tend to increase with the injection of contrast agency, and both CT and MR images reach highest performance in the equilibrium phase as shown in Fig. 6. MR can demonstrate fibrotic

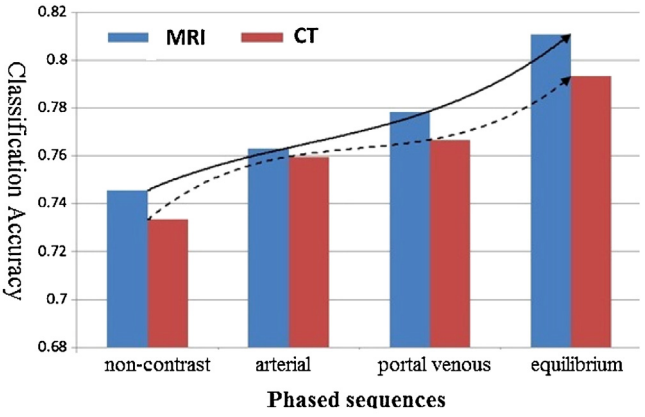


Fig. 6. The AR value in 4 phased images show that the texture analysis is effective in equilibrium phase than in other phased images.

texture efficiently and the equilibrium phase image is recommended as a main tool for interpretation of cirrhosis.

Exhaustive validation processing by LOO with SVM method is very time consuming because the number of combination of 15 texture features is very large. In our experiment, a 40 group 0 + 40 group 1 dataset in one sequenced image dataset utilizing the loop procedure of training and testing, for example, takes an average of 6 to 9 h on a workstation (4.3G CPU with 4G RAM, Win7 64bit version). However, once the modeling is completed, we can use this predictive model to diagnose a new dataset within a second. Although some studies announced their effective methods for training samples with improved algorithms that may greatly reduce the validation processing time, the robustness and accuracy still has problems. Furthermore, it is hard to determine a gold standard to verify their results. Therefore, in this study we chose the most critical validation method of LOOCV to ensure the objectivity of our results.

It will be a great challenge to include a same procedure of human's observation experiment to check if its conclusion is identical with CAD's. One of our former studies [27] demonstrated some preliminary results on the comparison of the human's observation with computer-aided diagnosis. Although radiologists' performance was significantly lower than that of the computer algorithm with only the texture pattern of fibrosis, the trend of accuracy changing is the same on both sides.

There are some limitations to our study, and these should be improved upon in our future work. First, the datasets are only divided into two groups. This is due to the unbalanced distribution of each of the fibrous stages and the purpose of decreasing the complexity of the validation. In the next step, our studies will extend to the quantitative assessment of fibrosis into 5 stages with the increasing datasets. Second, this study does not combine images from multiple phases due to the computational complexity. Top features from each phase could be combined to get up to 60 (15×4) features per case, and the performance of such a future work will tell us the variable importance from all phases. Third, this study was conducted as preliminary research, and considerable time and effort was expended to manually define the ROIs and record the values for validation. We have now started to develop an automated algorithm to select ROI within the extracted liver region, as well as integrate the GLCM, LOOCV and SVM method into one program. Finally, the overall AR's classified by texture analysis are not high (less than 0.8) enough to be used as an individual model to stage fibrosis in clinical practice. This may due the fact that the fibrous pattern is not obviously different from different stages, and even the experienced radiologist cannot interpret fibrosis well only with texture on images. It is possible to differentiate individual fibrosis stages by using texture features with combination with other effective features such as shape, volume, elasticity, and so on [28].

5. Conclusion

This study demonstrated that SVM and leave-one-case-out method are effective in feature selection from a large number of datasets. The iteration test on all of the subsets indicates that 4 to 7 features are the optimal number of features to classify liver fibrosis into two groups. Comparing the accuracy of classification on two modalities, we have shown that MR images have an advantage over CT images, while 3 T MR is better than 1.5 T MR to classify liver fibrosis, and mean gray value and entropy are the most useful features. The texture analysis is more effective in equilibrium phase than in other phased images and thus is recommended as a main tool for interpretation of cirrhosis.

Acknowledgments

This work was supported in part by two research support from the National Natural Science Foundation of China (Nos. 61262027 & 45627390), and in part by a financial support from China Scholarship Council (201308455018), and by JSPS Grant-in-Aid for Scientific Research on Innovative Areas (grant number 26108005). The authors are grateful to Yufan Zeng, Hiroshi Kato and members at Fujita lab in Gifu University, for their early efforts and discussions on this research.

References

[1] El-Serag HB, Mason AC. Rising incidence of hepatocellular carcinoma in the United States. N Engl J Med 1999;340(March):745–50.
[2] Yatsuhashi H, Yano M. Natural history of chronic hepatitis C. J Gastroenterol Hepatol 2000;15(May (Suppl. s2)):111–6.
[3] Wynn TA. Cellular and molecular mechanisms of fibrosis. J Pathol 2008;214(Jan (2)):199–210.
[4] Saadeh S, Cammell G, Carey WD, Younossi Z, Barnes D, Easley K. The role of liver biopsy in chronic hepatitis C. Hepatology 2001;33(1):196–200.
[5] Poynard T, Ratziu V, Bedossa P. Appropriateness of liver biopsy. Can J Gastroenterol 2000;14(6):543–8.
[6] Muthupillai R, Lomas DJ, Rossman PJ, Greenleaf JF, Manduca A, Ehman RL. Magnetic resonance elastography by direct visualization of propagating acoustic strain waves. Science 1995;269(Sep (5232)):1854–7.
[7] Rouvière O, Yin M, Dresner MA, et al. MR elastography of the liver: preliminary results. Radiology 2006;240(Aug (2)):440–8.
[8] Huwart L, Sempoux C, Salameh N, Jamart J, Annet L, Sinkus R, et al. Liver fibrosis: noninvasive assessment with MR elastography versus aspartate aminotransferase-to-platelet ratio index. Radiology 2007;245(Nov):458–66.
[9] Lucidarme O, Baleston F, Cadi M, Bellin MF, Charlotte F, Ratziu V, et al. Non-invasive detection of liver fibrosis: Is superparamagnetic iron oxide particle-enhanced MR imaging a contributive technique? Eur Radiol 2002;13(Sep (3)):467–74.
[10] Aguirre DA, Behling CA, Alpert E, Hassanein TI, Sirlin CB. Liver fibrosis: noninvasive diagnosis with double contrast material-enhanced MR imaging. Radiology 2006;239(May):425–37.
[11] Taouli B, Tolia AJ, Losada M, Babb JS, Chan ES, Bannan MA, et al. Diffusion-weighted MRI for quantification of liver fibrosis: preliminary experience. AJR 2007;189(Oct (4)):799–806.
[12] Friedrich-Rust M, Ong MF, Herrmann E, Dries V, Samaras P, Zeuzem S, et al. Real-time elastography for noninvasive assessment of liver fibrosis in chronic viral hepatitis. AJR 2007;188(3):758–64.
[13] Wang J, Guo L, Shi XY, Pan WQ, Bai YF, Ai H. Real-time elastography with a novel quantitative technology for assessment of liver fibrosis in chronic hepatitis B. Eur J Radiol 2012;81(1):31–6.
[14] Zhang XJ, Li WG, Fujita H, Kanematsu M, Hara T, Zhou XR, et al. Automatic segmentation of hepatic tissue and 3D volume analysis of cirrhosis in multi-detector row CT scans and MR imaging. IEICE Trans Inf Syst 2004;E87-D(8):2138–47.
[15] Goshima S, Kanematsu M, Kobayashi T, Furukawa T, Zhang X, Fujita H, et al. Staging hepatic fibrosis: computer-aided analysis of hepatic contours on gadolinium ethoxybenzyl diethylenetriaminepentaacetic acid-enhanced hepatocyte-phase magnetic resonance imaging. Hepatology 2012;55(1):328–9.
[16] Li WD, Zeng YF, Zhang XJ, Huang Y, Long LL, Fujita H. Staging the hepatic fibrosis on CT images: optimizing the slice thickness and texture features. In: IEEE 2011 international symposium on bioelectronics and bioinformatics. 2011. p. 267–70.

[17] Lee GN, Zhang X, Kanematsu M, et al. Classification of cirrhotic liver on MR images using texture analysis. Int J Comput Assist Radiol Surg 2006;1(Suppl. 1):379–81.
[18] Li W, Zhang X, Kanematsu M, Hara T, Zhou X, Fujita H, et al. Development of an automated method for differentiation of cirrhotic liver in abdominal MR images. Med Img Inf Sci (MII) 2004;21(2):137–43.
[19] Wu C, Chen Y. Texture features for classification of ultrasonic liver images. IEEE Trans Med Imag 1992;11(2):141–52.
[20] Wei Z, Li Y, He M. Analysis of incidence rate of hepatic cancer in Fushui county from 2004 to 2008. Chin J Oncol Prev Treat 2011;3(2):137–8 (in Chinese).
[21] Zhang X, Zeng Y, Fujita H, Wen L, Long L, Huang Y. Effective staging of fibrosis by the selected texture features of liver: which one is better, CT or MR imaging? Radiological Society of North America scientific assembly and annual meeting program, Radiological Society of North America (*RSNA*) SSG08-03. 2013 (abstract).
[22] Viral Hepatitis Prevention and Treatment Plan, Chinese Society of Hepatology. Chin J Infect Dis 2001;19(1):56–62.
[23] Haralick R. Statistical and structural approaches to texture. Proc IEEE 1979;67(5):786–804.
[24] Vapnik VN. The nature of statistical learning theory. Springer Verlag: Heidelberg, DE; 1995.
[25] Schapire R. Leave one out error stability, and generalization of voting combinations of classifiers. Mach Learn 2004;55:71–97.
[26] Kato Hiroki, Zhang Xuejun, Kondo Hiroshi, et al. Computer-aided diagnosis of chronic liver disease using 3 Tesla MRI artificial intelligence: optimization of the region of interest in texture analysis. In: Proceeding of 38th Japan magnetic resonance medicine conference. 2010 (in Japanese).
[27] Kato Hiroki, Kanematsu Masayuki, Zhang Xuejun, et al. Computer-aided diagnosis of hepatic fibrosis: preliminary evaluation of MRI texture analysis using the finite difference method and an artificial neural network. Am Roentgen Ray Soc (AJR) 2007;189:117–22.
[28] Zhang Xuejun, Goshima Satoshi, Zhou Xiangrong, Hara Takeshi, Kanematsu Masayuki, Fujita Hiroshi, et al. Quantitative staging the hepatic fibrosis: computer-aided diagnosis by shape, texture, volume, and elasticity analysis. In: Proc. of Radiological Society of North America Scientific Assembly and annual meeting program Radiological Society of North America (RSNA), LL-INE1211. 2011.

Xuejun Zhang (1969–, Naning, Guangxi, China) received the B.S. degree in Physics from Guangxi University, PR China, in 1991; the M.S. degree in Electronics and Information Systems Engineering from Gifu University, Japan, in 2001; and the Ph.D. degree in Electronics and Information Systems Engineering from Gifu University, Japan, in 2004. From 2004 to 2007, he continued his research in Computer-Aided Diagnosis as a postdoctoral researcher at Department of Intelligent Image Information, Division of Regeneration and Advanced Medical Science, Graduate School of Medicine, Gifu University. He was invited as a visiting professor of Center for Medical Diagnostic Systems and Visualization, University of Bremen, Bremen, Germany at 2006. Since 2007, he was a Professor of School of Computer and Electronic Information, Guangxi University, Nanning, Guangxi, China. Now he is a visiting scholar in University of Southern California. His research interests include computer-aided diagnosis system, image segmentation, pattern recognition, visualization in medicine. He has published over 70 papers in Journals, Proceedings, Book chapters and Scientific Magazines.

Hiroshi Fujita received the B.S. and M.S. degrees in Electrical Engineering from Gifu University, Japan, in 1976 and 1978, respectively, and Ph.D. degree from Nagoya University in 1983. He was a Research Associate at University of Chicago, USA, from 1983 to 1986. He is currently a Chairman and a Professor in the Department of Intelligent Image Information, Graduate School of Medicine, Gifu University, Japan. His research interests include computer-aided diagnosis system, image analysis and processing, and image evaluation in medicine. He has published over 500 papers in Journals, Proceedings, Book chapters and Scientific Magazines. He is currently President of the Japan Society of Medical Imaging and Information Sciences.

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

Development of a novel imaging informatics-based system with an intelligent workflow engine (IWEIS) to support imaging-based clinical trials

Ximing Wang [a,*], Brent J Liu [a], Clarisa Martinez [b], Xuejun Zhang [c], Carolee J Winstein [b]

[a] Image Processing and Informatics Lab, Department of Biomedical Engineering, Viterbi School of Engineering, University of Southern California, Los Angeles, CA 90089, United States
[b] Division of Biokinesiology & Physical Therapy, University of Southern California, Los Angeles, CA 90089, United States
[c] School of Computer, Electronics and Information, Guangxi University, Nanning, Guangxi 530004, PR China

ARTICLE INFO

Article history:
Received 18 November 2014
Accepted 22 March 2015

Keywords:
Imaging-based clinical trials
Workflow engine
System framework
Imaging informatics system

ABSTRACT

Imaging based clinical trials can benefit from a solution to efficiently collect, analyze, and distribute multimedia data at various stages within the workflow. Currently, the data management needs of these trials are typically addressed with custom-built systems. However, software development of the custom-built systems for versatile workflows can be resource-consuming. To address these challenges, we present a system with a workflow engine for imaging based clinical trials. The system enables a project coordinator to build a data collection and management system specifically related to study protocol workflow without programming. Web Access to DICOM Objects (WADO) module with novel features is integrated to further facilitate imaging related study. The system was initially evaluated by an imaging based rehabilitation clinical trial. The evaluation shows that the cost of the development of system can be much reduced compared to the custom-built system. By providing a solution to customize a system and automate the workflow, the system will save on development time and reduce errors especially for imaging clinical trials.

1. Background

Clinical trials have rapidly evolved during the past decade. According to U.S. National Institutes of Health, the total registered clinical studies rose from 10,241 in 2003 to 158,829 in 2013 [1]. Specifically, imaging techniques have been used widely in clinical trials to track and evaluate biomarkers [2]. Massive amounts of imaging data are produced during the clinical trial and the workflow has become more and more complex. Moreover, since these data are collected from different departments, collaboration between healthcare staff plays a more essential role and substantially impacts the overall efficiency of healthcare [3]. Unfortunately, the cost of collecting and managing the large volumes of data and communication processes has negatively impacted the productivity.

Electronic data collection (EDC) has been proven to reduce the monitoring and data management cost in comparison to paper data collection (PDC) [4]. For years, hospital information systems (HIS) and picture archiving and communication systems (PACS) are prevalent in hospital environment for data management. In contrast to the hospital environment, academic research data are usually stored in disparate environments and processed with distinct and highly customized, single silo workflows. Conventional PACS designed for the hospital environment is limited in the flexibility for various research-related workflows [5]. Furthermore, clinical trial workflows are usually agile and need to be adapted to disparate research methods that require the data management to be customized. Typically, the need for clinical trial workflow is addressed by custom-developed data management systems. However, custom-built software development can be time-consuming and resource-intensive, and most of small-scale clinical trials are limited for the cost of a custom-built informatics system. Data capture database development platforms, such as REDCap [6], are widely used for small-scale trials as a framework. Such software is mainly textual form based and not integrated with imaging data and multimedia data. In such systems, the workflow view is not presented, activities are not automated and relations between forms cannot be defined. As a result, a solution to provide efficient workflow development and refinement along with data management and decision support to serve the small-scale imaging-based clinical trials is urgently needed.

Workflow automation technology has been demonstrated to improve the efficiency of business workflow for over ten years [7]. In the business and industry field, workflow automation products enable the

* Corresponding author. Tel.: +1 213 821 8395.
E-mail address: ximingwa@usc.edu (X. Wang).

http://dx.doi.org/10.1016/j.compbiomed.2015.03.024
0010-4825/

company to create a workflow with a variety of components, mostly electronic forms, and utilize the workflow to manage the data flow in their production line. The advantage of the workflow automation software is that it allows user to create a data management system based on specific workflow without software development. Additionally, the workflow engine is able to process a sequence of steps automatically based on the predefined rules. For example, a workflow requires the data management system to parse the data, store the data to the database and email the reports to related staff once the data is received. With the workflow engine, this set of actions can be easily defined. The avoidance of the software development reduces the cost and effort in the business model and minimizes the potential errors in human action [8].

The workflow engine concept has already been introduced in healthcare, e.g. Huser et al. presented a system for clinical decision support in 2011 [9]. However, there is no similar system designed for imaging based clinical trials. Compared with business model, imaging based clinical trials require more compatibility with multimedia data [10], more intuitive workflows, and particular analysis of imaging biomarkers. To address the challenges in the imaging-based clinical trial, we present a novel prototype imaging-informatics based prototype system named IWEIS (imaging workflow engine informatics system) to support imaging-based clinical trials.

As a case study of the imaging informatics system, the IWEIS prototype was implemented and applied to an imaging-based stroke rehabilitation phase I clinical trial named Dose Optimization for Stroke Evaluation (DOSE). This paper will discuss IWEIS based on this DOSE clinical trial. The DOSE clinical trial aims to understand the optimal dose of a principle-based rehabilitation intervention for stroke [11]. The trial aims to recruit 40 subjects in 4 years. During the trial, complex multi-media data including laboratory-based measurements such as Wolf Motor Function test (WMFT) and behavioral questionnaires, such as Motor Activity Log (MAL) will be collected. Imaging studies, such as structural MRI, Transcranial Magnetic Stimulation (TMS), and Diffusion Tensor Imaging (DTI) will also be conducted and collected. As an important component, the imaging data will be used to track the changes in stroke lesion size and location through the physical therapy treatment. Potential correlation of lesion size and rehabilitation progress will also be investigated.

By utilizing IWEIS, the deployment of the DOSE informatics system will be different from traditional methods. Fig. 1 demonstrates the traditional scenario and the proposed new scenario. In the traditional scenario, after the design of the study, clinical designers need to set up multiple meetings with software developers to explain the requirements. This step is critical and could be time-consuming as the software developers may not have adequate clinical trial knowledge. The software development step, which is carried out by software developers, normally lasts several weeks to several months. The final step, quality assurance of the software, requires the collaboration between clinical researchers and software developers. Sometimes the system can have a lot of deficiencies due to developer's lack of the knowledge of the clinical trial. In the IWEIS scenario, clinical researchers are able to deploy a system without software engineers. That means clinical researchers controls all the development processes and the system is deployed by people with thorough understanding of the clinical trial. Therefore, the IWEIS scenario improves the efficiency compared to traditional scenario. In Fig. 1, red stages show negative impact on the cost of time and resources with traditional scenario, while green stages shows where IWEIS can reduce of the time and cost of the system deployment.

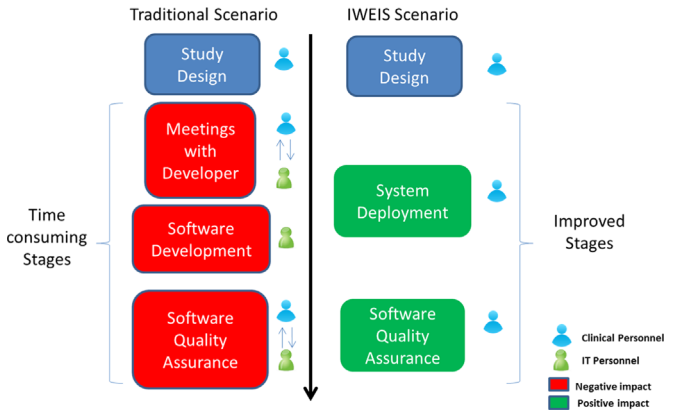


Fig. 1. Traditional scenario and IWEIS scenario in the development of clinical trial informatics system. Stages with negative impacts on the cost of system development are marked as red, and stages with positive impacts are marked as green. This figure shows where the challenges are in the traditional scenario and where the IWEIS scenario will improve the efficiency of the system development. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2. Methods

IWEIS was developed in Hypertext Preprocessor (PHP) 5.3.10 along with MySQL 5.5.32 and Apache server 2.2 under the HTML5 standard and deployed within a Linux Environment. The IWEIS software also adopts a variety of open-source software such as JQuery [12], a prevalent JavaScript library, Diagramo [13], a web-based flowchart software with pure HTML 5, and Twitter bootstrap [14], an intuitive and powerful front-end style framework for web development etc. To meet privacy requirements, the system requires a SSL certificate and authentication with username and password.

The system is developed based on a two-tier architecture (See Fig. 2). The base layer is comprised of a MySQL database, a file storage system, a DICOM (Digital Imaging and Communications in Medicine) viewer and the Apache web based server. The base layer provides necessary services for system operations and stores all the data to be collected. The second tier consists of two main components: the designer module and the application module. Each module incorporates a web-based graphical user interface (GUI) with a unique Uniform Resource Locator (URL). The designer module is the back-end view of the system, which allows a project designer to deploy and customize the informatics system. Once an instance of the system is deployed, the application module shows the front-end view of the created imaging informatics system, allowing client users to collect and manage the data by the workflow of the clinical trial.

2.1. System components

2.1.1. The designer module

The designer module is the back end of the system that can create and deploy the information system instance. Typically, the

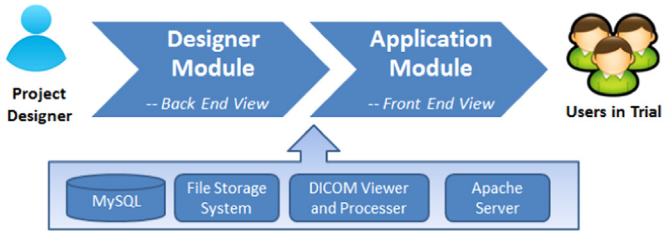


Fig. 2. Schematic presentation of the IWEIS system architecture. The system architecture is a two-tier layer comprised of a base layer with four main components. In the second tier, the designer module is utilized for deployment and the application module is utilized as the imaging informatics based system for front end users in clinical trial applications.

Please cite this article as: X. Wang, et al., Development of a novel imaging informatics-based system with an intelligent workflow engine (IWEIS) to support imaging-based..., Comput. Biol. Med. (2015), http://dx.doi.org/10.1016/j.compbiomed.2015.03.024

deployment of the proposed system for an imaging based clinical trial is comprised of 5 steps.

1. A principle investigator (PI) designs the trial with specific workflows and requirements. In the DOSE trial, the PI defines the data to be collected in each stage and each staff's tasks in each stage.
2. Develop the workflow with the IWEIS workflow designer. In DOSE trial, the manager is able to build the workflow with the workflow painter in IWEIS.
3. Define tasks in each stage of the workflow. In the DOSE trial, the manager is able to create a variety of tasks that need to be done in each stage. The tasks include collecting form data, upload imaging data etc.
4. Assign the rules in the workflow engine. The rules are used to automate some processes and assist in the decision making. In the DOSE trial, rules are used to assist screening patients. The criteria for screening are entered as rules, and the system shows if a patient is satisfied with these rules. The rules for restrict tasks access can also be defined. Firstly, the arrows in the workflow define the basic rules. Each stage can only be accessed when all the previous stages are completed. Secondly, rules can be set to each task. User can define a set of prerequisite tasks that must be completed or approved before access of the target task.
5. The front-end system is generated automatically and ready for use for the clinical trial.

Once a clinical trial workflow has been designed (step 1), the project manager is able to develop the workflow-based imaging informatics system using the designer module. The designer module is composed of three main components, which are related to three main steps (step 2–4).

2.1.1.1. Component 1: workflow designer. The designer module has an integrated workflow designer utilizing a thin-client HTML5 based workflow diagram painter (see Fig. 3) for users to design the workflow (step 2). The painter has a menu bar and three panels. Users are able to pick a shape from the left panel to draw on the canvas. Functions including showing grids, group/ungroup, text and customizing styles are also available within the tool. The canvas in Fig. 3 illustrates the current workflow for the DOSE trial as the first initial application example. In the workflow, each shape stands for a stage of the workflow trial. By choosing a specific shape and color, each shape can be used to indicate similar types of stages within the workflow. In each stage, the user is capable of combining one or several workflow tasks within the stage by selecting the shape and clicking the button "Edit contents" (Arrow B). By clicking "Engine Rules" (Arrow A), the user is also able to configure the rules for each shape in the workflow.

2.1.1.2. Component 2: workflow stage editor and tasks builder. The next step is to define tasks in each workflow stage (step 3). The stage editor enables user to build one or more tasks in each stage. As long as a task is created in a stage, the task will be shown in the same stage of the created instance system for data collection and data presentation. A toolbox with five task builders is provided in the stage editor.

- General tasks builder
  - o *Digital forms builder*: The textual forms are the most widely used method for data collection within the clinical trial. IWEIS provides a digital form builder tool. To further facilitate the production of forms, a template library is also incorporated. The user is able to create a template, load a template from the library, and customize the form from the loaded template.
  - o *File uploalder builder*: Typically, a number of files relevant to the clinical trial need to be stored for review and other uses. For example, within the DOSE clinical trial, the patients' signed HIPAA (The Health Insurance Portability and Accountability Act) authorization and patient's motion sensor data during physical therapy must be stored and associated with each patient and specific workflow stages. The builder allows user to create a file uploading and downloading tool within any of the workflow stages.
- Multimedia tasks builder
  - o *Medical images task builder*: This task builder enables the user to create an uploader and viewer to upload DICOM format files and view the DICOM studies through a zero-footprint Web Access to DICOM Objects (WADO) viewer. To respect patient privacy and abide by HIPAA's Privacy Rule, an anonymizer is integrated with the uploader and DICOM headers of all uploaded images will be de-identified automatically. User can

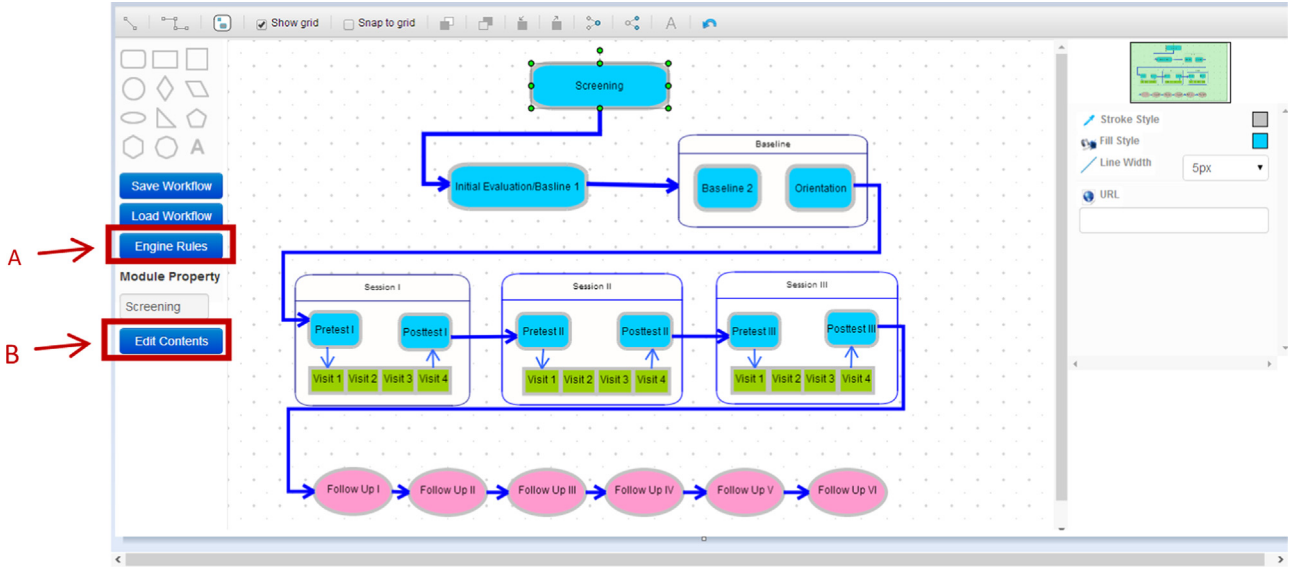


Fig. 3. Workflow Designer Module in the back end. The user is able to pick a shape in the left panel, drag and drop it into the canvas. The canvas illustrates a sample workflow design example of the DOSE rehabilitation clinical trial. Arrow A links to the engine rules page. Arrow B allows the user to edit contents in each stage of the workflow.

Please cite this article as: X. Wang, et al., Development of a novel imaging informatics-based system with an intelligent workflow engine (IWEIS) to support imaging-based..., Comput. Biol. Med. (2015), http://dx.doi.org/10.1016/j.compbiomed.2015.03.024

also link the images task with a textual form for collecting results of the images analysis. Additionally, user could upload template image as a mask to assist image interpretation. More details are further discussed in the imaging section within the Application module.

- o *Video and audio plugin builder*: Video and audios are commonly used file format for clinical trials. This builder allows user to create a video uploader and viewer, which enables the user to upload a video and then review it with a video player.

In summary, in step 3, the user defines the data that needs to be collected in each stage and establishes the data collection tools.

2.1.1.3. Component 3: rules manager. Rules are essential for workflow engine systems. Within the clinical trial environment, the rules are less complicated than in the business environment. However, well-defined rules are needed to automate certain processes, provide decision-support, reduce human error, and minimize resource overhead. Based on the common requirements within clinical trial environments, the IWEIS prototype supports two categories of rules. The first category is workflow stage rules, which is built-in within the workflow engine. In the workflow diagram painter, the user is able to link stages within the clinical trial workflow by an arrow. These arrows define the sequence of the study, and a stage can only be accessed if the previous stages are completed. In other words, the user is not allowed to skip prerequisite stages during the clinical trial. The second category is task rules, which are applied to specific tasks in a workflow stage. There are three types of task rules: (1) Completion restriction rules; (2) Approval restriction rules; and (3) Complex requirements which will be described further below.

1. *Completion restrictions*: The task access is restricted unless specific prerequisite tasks are completed. IWEIS allows user to select the completion-required specific tasks for each task.
2. *Approval restrictions*: The task access is restricted unless specific tasks are approved by quality assurance (QA) users.
3. *Complex restrictions*: The workflow stage access is restricted by a set of rules, including mathematical comparisons of a value within tasks or complex states. For example, a task can be restricted by two rules. Suppose the first rule is limited by the "Age" value in the "Profile" form. The participant's age must be less than a value defined by the user. The second rule is limited by the "stroke length" value in the "Phone Screen" form. The subject's stroke length must be greater than a certain time defined by the user. Only if the subject's data meet both rules, this task can be accessed.

In addition to the application of the rules discussed before, the rules can also be used as a decision support tool. For example, within the DOSE trial, a patient can only be enrolled if he/she satisfied all requirements within the screening workflow stage, e.g. age must be greater than 21, stroke severity is within a range etc. The complex restrictions are used as restrictions for the access of the enrollment form of the subject which would be the next step in the clinical workflow. Instead of manually looking through all the conditions and verifying the enrollment by the study coordinator, the system automatically shows unsatisfied requirements and prevents the enrollment of the subject resulting in fewer potential human errors.

2.1.2. The application module

After the deployment of the system, the application module creates an actual instance of the system that was designed in the designer module. The newly created system will be used to create new patients and enter data in the real clinical trial. The homepage of the system is a dashboard for users. The dashboard is composed of two sections, my tasks and patient worklist. The user with permissions is also able to create a new patient for the study from dashboard. The "my tasks" section shows all tasks that are ready for the user to process. The workflow engine checks all the rules of the tasks and pushes the available task to the user in charge. Pick a task will lead to the data entry page of that stage of the workflow for the patient. The patient worklist section shows all enrolled patient and link to patient profile page. The patient profile page is designed to show the study workflow with status tracker since a workflow showing the overview of the study process for each patient will benefit data collection and monitoring in clinical trial.

To utilize the system, the first step is to create a new subject with a subject ID. Creating a new subject or choosing one of the subjects from the worklist leads to the subject profile page with all the data collection links activated. The left part of Fig. 4 illustrates the subject profile page with the clinical trial workflow example. In each stage of the workflow, a progress indicator shows the number of tasks in the workflow stage that have been completed. Clicking the workflow stage results in a popup widow with all the tasks related to that particular stage. In the right part of Fig. 4, the first task shows a text-based digital form with a rule restriction. The rule is not met since a prerequisite task has not been completed. The second task shows a link for uploading videos and a list of the videos already uploaded. The uploaded video can be downloaded or streamed for viewing. The third task shows the link to the DICOM WADO viewer. An upload button allows the user to upload DICOM images, anonymize the DICOM header, and extract the header information into the database for efficient queries. The worklist within the viewer shows all the images uploaded. These images can be viewed through a zero-footprint DICOM WADO web viewer or downloaded to the local client for further image processing analysis. Task 4 illustrates the file uploader task.

2.2. System features

2.2.1. Images analysis module

Imaging data have been used widely as the biomarkers in current medicine. More and more clinical trials use imaging biomarkers as evidences to detect subtle changes. For example, the lesion size, rate of growth and location detected in brain images can be used to predict the severity of the stroke. As a result, a platform that can combine the traditional textual data and the imaging data will be beneficial to researchers in clinical trials. The IWEIS system has integrated a vendor neutral imaging data analysis tool, the Images Analysis Module, which could upload, view, and facilitate data analysis online.

The Images Analysis Module includes both a DICOM uploader and a WADO (Web Access to DICOM Objects) viewer. The DICOM uploader allows loading of image studies either from a CD or a local client machine. In order to comply with HIPAA, all images are de-identified and transmitted under Secure Sockets Layer (SSL) to the system. Once uploaded, all data will be parsed and the header information will be extracted to a database for efficient query.

To facilitate the analysis of the imaging data, IWEIS utilizes a vendor neutral WADO viewer developed by Imaging Processing and Informatics Laboratory (IPILab) at the University of Southern California for online image viewing [15]. The viewer is a stand-alone PHP-MySQL-Apache based software package. A PHP DICOM parser library NanoDICOM [16] is utilized to process the DICOM files. Images are stored in the file system and accessible by WADO identifiers through Apache. An interface was developed and the viewer is integrated with the IWEIS system. In the image viewer (See Fig. 5), the user can navigate through all the image series within each study and change the layout of the viewer. The user is also able to scroll, pan, zoom, change window/level and show

Please cite this article as: X. Wang, et al., Development of a novel imaging informatics-based system with an intelligent workflow engine (IWEIS) to support imaging-based..., Comput. Biol. Med. (2015), http://dx.doi.org/10.1016/j.compbiomed.2015.03.024

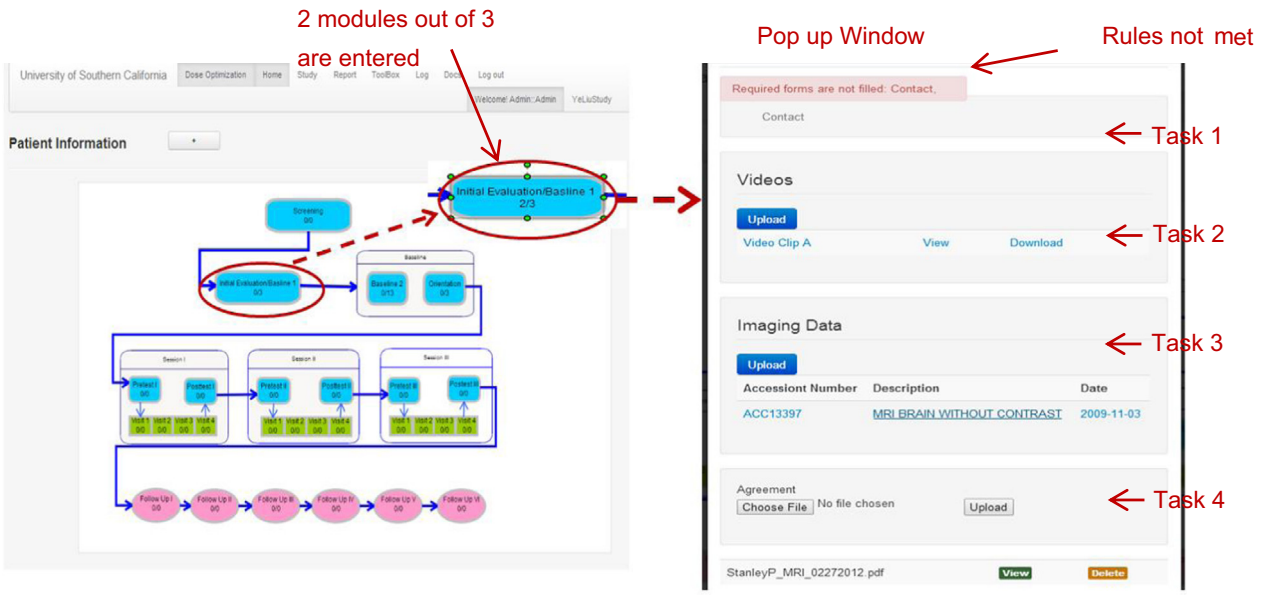


Fig. 4. Screenshot of Patient profile page with integrated workflow in the front end. The left panel illustrates the patient's page. In each stage of the workflow, the number shows the completion status of the modules. Clicking the stage leads to a popup window with all the modules in that stage (right panel). If the rules for a stage are not satisfied, the workflow stage will be inaccessible.

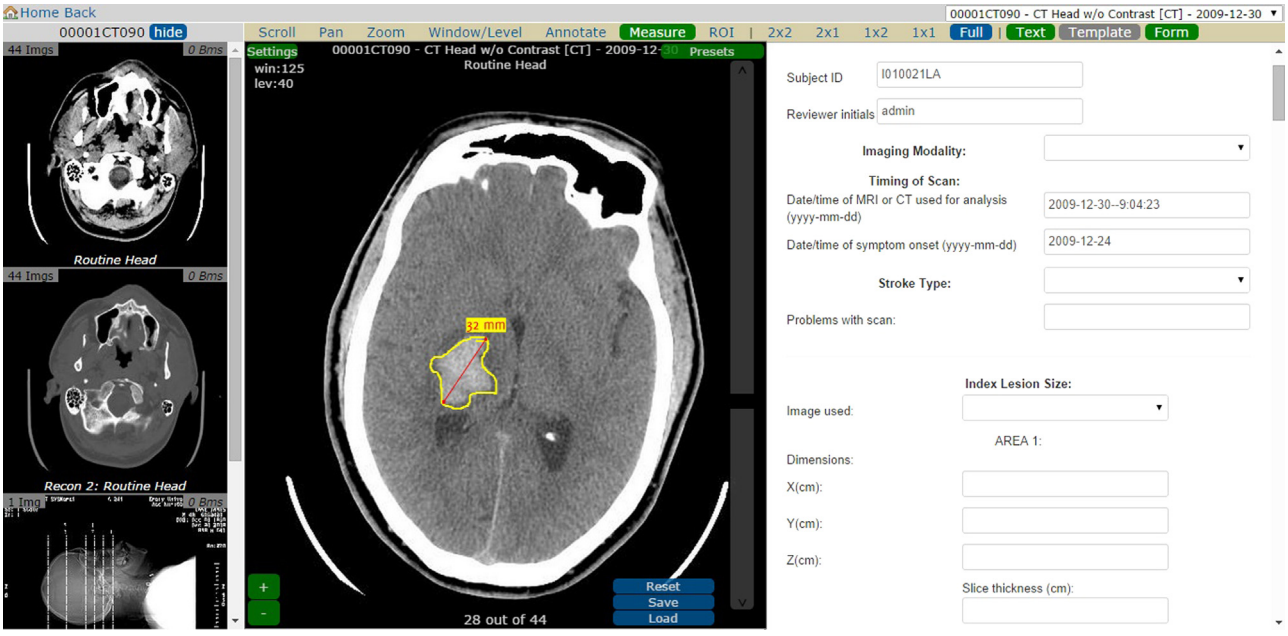


Fig. 5. Web-based zero footprint DICOM WADO viewer. The user is able to combine a self-designed data collection form with the images. Therefore user is able to read the image, fill out the reading result with a customized form and send it to database. The viewer also provides annotation, measurement, and ROI marker tool to assist the identification of lesions or other pathologies related to the clinical trial.

textual information extracted from DICOM header. Different from traditional image viewers, this viewer is a zero-footprint solution and no client side plugin is needed. This is extremely advantageous due to the common restrictions on computers used within the clinical environment and where most of the clinical trials personnel can benefit from the portability of the image viewer.

Distinguished from other image viewing software, IWEIS facilitates the analysis of the images by novel features. Similar to other software, measuring, ROI (Region of Interest) marking and annotation tools are available with the image viewer. Fig. 5 illustrates an example screenshot where the diameter of the lesion is being measured and a ROI markup annotation of the lesion has been added. The measurement and markup results are stored within the database for future retrieval. Despite these traditional tools, Images Analysis Module is able to link a data collection form with the image viewer. In the practical application, radiologists usually need to record the image reading results into a textual form and send it to the database. IWEIS allows the user to create a textual data entry form in the back end and link it with the images task. Once an image task is linked with the form, the user is able to show the image and textual form side by side (see Fig. 5). Therefore, the user is able to collect the image reading results conveniently with a customized form. Based on the data collected by the forms, user is able to compare longitudinal reading results of the same patients and investigate the rate of growth of the lesion.

Please cite this article as: X. Wang, et al., Development of a novel imaging informatics-based system with an intelligent workflow engine (IWEIS) to support imaging-based..., Comput. Biol. Med. (2015), http://dx.doi.org/10.1016/j.compbiomed.2015.03.024

In order to assist the identification of the location of lesion, the Images Analysis Module is equipped with a template masking tool. To use this tool, the user needs to upload a PNG format semi-transparent template image, and the tool allows user to overlay the template on top of the image. Fig. 6 illustrate an example of the tool. The user adapts published vascular territory templates [17,18] and uploads these template images. At the right part of the window, all templates uploaded are shown. The user can pick any of the templates and overlay it on top of the image. User can also zoom in, zoom out and rotate the template mask to match the image better. By matching the template with the images, the identification of the lesion location within the organ will be easier and more accurate.

2.2.2. *System modification and expansion*

As discussed before, one of the challenges is that sometimes the researchers need to redesign or improve the informatics system even after some patients have been enrolled. For instance, after the treatment of several subjects, clinical researchers may find out that some forms need to be expanded and some stages of the workflow need to be modified. In the traditional scenario, clinical researchers need to contact the software developer and explain the requirements. The software developer needs to back up the data and reprogram the system. This process may take several days to several weeks, and the cost will be high. Sometimes, the modification will be expensive and buggy if the system is not designed flexibly at the beginning. To address this challenge, IWEIS system has several mechanisms in various situations.

- Tasks level.
  All the task tables have a “stage ID” data field and stores which stage it belongs to. When a task needs to be moved to another stage, the “stage ID” is updated and hereby the task is moved at the application end. New tasks can also be created in any stage. Deletion of a task will result in hiding the tasks in the workflow and the application site. However, the data that have already been collected will not be lost and can be retrieved at any time.
- Workflow level.
  The workflow level modification means that new stages will be inserted into the workflow or some stages will be moved to other positions in the workflow. To implement this feature, a “stage sequence” table is utilized to store the previous and next stage for each stage. When the workflow is modified, the database will be backed up before the modification, and the “stage sequence” table is updated. The tasks will be associated with the original stages no matter where they are moved. Therefore the tasks in each stage will not be affected and the collected data will be retained. To remove a stage, the user needs to move or delete all the tasks in the stage, and delete the stage, which will removed it from the “stage sequence” table.
- Rules level
  Users can set new rules or update existing rules. Once the rules have been updated, the system will check all the rules with the collected data and warn all the violations under the new rule system. The new rule system will also be used in future data collection to prevent errors.

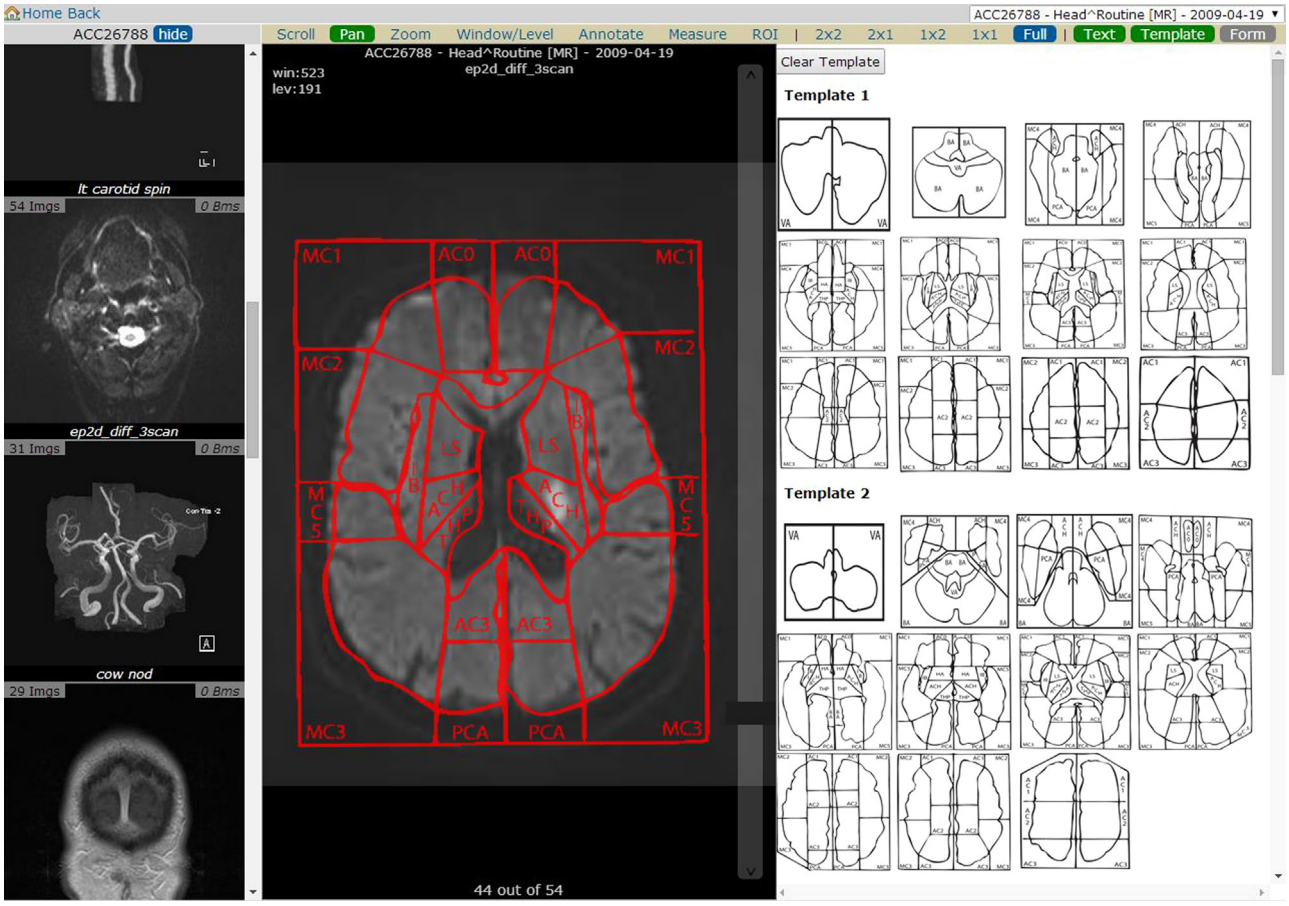


Fig. 6. Template masking tool. On the right part of the window, a library of sample templates (Templates adapted from Damasio et al.[17], courtesy of Dr. Matthew Edwardson) uploaded by user can be chosen. The uploaded templates can be shown on top of the image and can be rotated, zoomed in and zoomed out. Overlaying the template on top of images facilitate the identification of lesion location.

Please cite this article as: X. Wang, et al., Development of a novel imaging informatics-based system with an intelligent workflow engine (IWEIS) to support imaging-based..., Comput. Biol. Med. (2015), http://dx.doi.org/10.1016/j.compbiomed.2015.03.024

- Data fields level
  The data fields in each task can also be updated. User is able to insert new data field, update data field name and delete a data field. Deletion the data field result in hiding the data field from the tasks, while the data is still preserved and can be retrieved. After updating the data fields, the system will also automatically run the rule system to search for any violation of the predefined rules. Warnings will be shown for violations found.

2.2.3. Data query and presentation

IWEIS is more than a data collection tool or digital form builder. To further simplify the process for data analysis, IWEIS provides multiple solutions for data mining and presentation. To avoid potential data exposure, only approved users with data retrieval rights are able to access the data mining tools. The data can be queried in three ways. A query by patient ID leads to all the relevant data categorized by workflow stages in a patient record manner similar to an Electronic Medical Record (EMR). A query by the task and workflow stage leads to a cross-project sheet with all patients' data in the task group. A comprehensive query allows user to specify multiple search criteria. In the comprehensive query mode, the user can pick values from any of the tasks or data collected from the digital forms and set search conditions for them. For example, in DOSE clinical trial, the users may need to search the patients with stroke history for more than 3 years that have an MRI scanned at the initial evaluation stage. The query would be the “stroke length” value in the “history and demographics” form to be greater than 3 years, and the images in the initial evaluation stage that have been uploaded. The results return a list of patients that satisfy the conditions.

The data retrieved can be displayed online or downloaded to the local client machine. Text-based digital form data will be converted to an excel sheet for download and further analysis. Video and audio files are able to be streamed online or downloaded for post-processing. The imaging data can be viewed through the DICOM WADO viewer discussed above and can be compressed as a zip file for download.

2.2.4. Quality assurance

Clinical trials are usually under the monitoring of sponsors for quality control. A clinical trial has to be performed complying with standards and the data has to be collected with highest quality. To facilitate the quality assurance, the IWEIS adopts several mechanisms. Firstly, to be compliant with Title 21 CFR Part 11, all the activities are logged automatically. The log can be queried by username or downloaded directly from the graphical user interface. Secondly, data collected has to be approved at least one time (Please see Fig. 7). During the data collection process, the data need to be reviewed by the data collector when they entered the data. After that, project managers are able to define additional reviews from other users for each task group. This step is optional and can be defined based on the importance of data and the scale of trials. The system also provides an option of randomly selecting part of data for review, which is designed for large scale clinical trials, since it is unfeasible to review all the data in some cases. Thirdly, the system allows creating a user group (QA group) account to review all the data. This mechanism aims to provide an account for third party reviewers.

3. Evaluation

The IWEIS was evaluated with the DOSE clinical trial. The DOSE trial aims to collect 40 subjects in a four year period and currently 30 subjects have been enrolled. The subjects are randomized into four dosage groups to receive treatment with physical therapy. A baseline evaluation, three sessions of treatments, and six follow-ups are conducted for each patient. The data involved includes digital questionnaires, MRI and DTI images, videos, and motion sensor data. These data need to be collected and associated with the various stages in the workflow. Currently, a custom-built system is developed and being used for DOSE trial [19]. To evaluate the IWEIS system, we created an instance in IWEIS and compared it with no-system situation and custom-built system situation.

As shown in Table 1, the development time and the financial cost with IWEIS are much less than a custom-built system. To develop a custom-built system, the clinical trial leader needs to look for a developer, define the needs, communicate with the developer and review the delivered systems. Typically, the communication can be challenging for both clinical researchers and software developers. Miscommunication will result in a delay in delivery and more expenditure of man hours in re-development. In contrast, IWEIS allows researchers to deploy the system directly and eliminate the communication with developers. Thus, the development/refinement cycles timelines are reduced significantly. In DOSE trial, the custom-built system takes about 5meetings and 6 months with a part-time developer for the development. During the development of the custom-built system, regular meetings were also held for progress report. In contrast, the deployment of IWEIS system takes a few days without any meeting.

Another common situation is that clinical trials are easy to change in workflow and forms. Therefore, the flexibility of a system is another essential feature. In the experience of the DOSE clinical trial, there were quite a number of modifications of forms and requirements after the development started. For example, in the system development with DOSE trial, some forms were revised after the custom-system was developed. This modification demands communication with the developer, a re-programming of the software, a revising of the database, and a final review from the clinical trial researchers. Thus a custom-built system is able to be modified, but requires nontrivial efforts. Hence, the cost of a modifying a custom-built system is high and may not be appropriate choice for small-scale clinical trials. In contrast, the IWEIS system allows user to modify the form without re-programming. The data is backed up automatically, and the modification cost only several hours. IWEIS rules system's warning of violations also facilitate the validation of the system modification. In DOSE trial, the modification of the custom-built system takes a day to 2 weeks depending on the complexity, while the IWEIS system takes hours for basic modification. However, the custom-built system is able to realize advanced features but IWEIS system is not. For example, in the DOSE

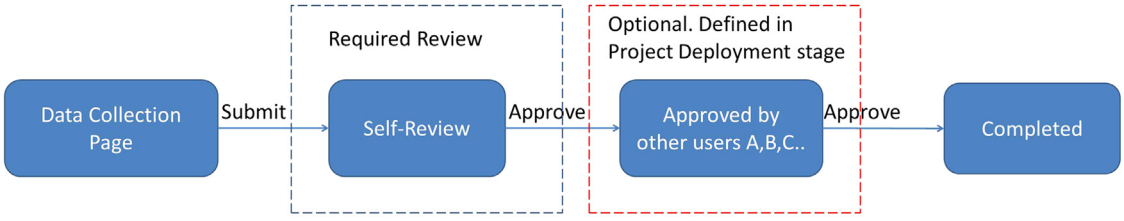


Fig. 7. Data collection review flowchart. For each data collection, the data entry person is required to review and approve the data to complete the data entry. Moreover, the study manager is also able to define if the data needs to be approved by other users, such as supervisors or a QA group to complete the data entry.

Please cite this article as: X. Wang, et al., Development of a novel imaging informatics-based system with an intelligent workflow engine (IWEIS) to support imaging-based..., Comput. Biol. Med. (2015), http://dx.doi.org/10.1016/j.compbiomed.2015.03.024

Table 1
Comparison of IWEIS system with no-system situation and custom-built system situation in DOSE clinical trial.

Tasks	No electronic system	IWEIS system	Custom-built system
Pre-development meeting	NA	0 times	~5 times
System development	NA	3 days	6 months with a part time developer
System modification	NA	1–8 h	1 day–2 weeks with a part time developer
Advanced features	NA	No	Yes
Financial cost	NA	~$0	~$5000
Data collection	Paper based, 1–2 min per question	1–2 min per question.	1–2 min per question
Imaging data	Stored in CD	Uploaded to system in minutes	Uploaded to system in minutes
Textual data query	Manually look through paper forms, 1–20 min	< 1 s by database query	< 1 s by database query
Imaging data query	Look through CDs manually. 1–20 min	< 1 s by database query	< 1 s by database query
Collaborations between staff	Need to deliver paperwork physically. Up to several days	Instantly. Users can login to see the tasks need to process	Instantly. Users can login to see tasks need to process
Accidental mistakes	Non-avoidable due to human errors	Accidental mistakes will be avoided if rules are set	Avoidable

trial, the custom-system is able to provide an interactive data collection tool for treatment and is able to clips videos automatically, while IWEIS system is limited in these functions.

From the functional perspective, both IWEIS and the custom-built system significantly reduce the time in data retrieval and collaboration compared with no-system situation. Without an electronic system, the data retrieval cost up to several hours for manually looking through paperwork or imaging CDs. With an electronic system, a simple data query is able to retrieve textual data and multimedia data instantly. The collaboration cost is also significantly reduced. Without an electronic system, physically delivery of paperwork and CDs may delay the progress for days. With an electronic system, the physical delivery is not necessary and system pushes tasks to user in charge promptly.

Based on the evaluation, the prototype IWEIS may have less functionality than a custom-built system today, but it is capable to meet regular needs with much less resources, which is beneficial for small scale clinical trials. In addition, once these generic tools are developed, they can be easily integrated into IWEIS and thus provide the same additional functionality as the custom-built system.

4. Discussion

Most of the imaging informatics systems are designed for hospital environment and not suitable for clinical trials. In the field of clinical trials, data capture software such as XNAT [20] and REDCap are the main software used to address the data management needs. However, current solutions are inadequate for imaging based clinical trials. Therefore IWEIS system aims to provide a solution specifically for imaging based clinical trials.

In imaging based trials, imaging biomarkers are used as evidences and researchers normally need to acquire information like lesion size, location, rate of growth. The reading results need to be collected in forms for longitudinal studies. Due to such characteristics of imaging based clinical trials, IWEIS provides a web viewer with measuring tool, side-by-side forms and template masking tool. Once the images reading results are collected, the user could query all the data for each patient and investigate the changes through different stages of the trial.

A comparison between IWEIS prototype and existing similar software REDCap and XNAT is illustrated in Table 2. All three software are free of programming for the deployment. However, firstly, REDCap only supports textual data, and XNAT only supports imaging data. IWEIS supports textual data, imaging data and multimedia data. Secondly, In REDCap and XNAT, the workflow concept is not illustrated. The data in REDCap and XNAT managed on a "visit" manner. The data is organized by visit number and there are no intuitive views

Table 2
Comparison of IWEIS with existing similar software.

	IWEIS	REDCap	XNAT
Non-programming deployment	Yes	Yes	Yes
Textual data	Yes	Yes	No
Imaging data upload and display	Yes	No	Yes
Multimedia data	Yes	No	No
Data security and access	Yes	Yes	Yes
Management by workflow	Yes	No	No
Management by visits	Yes	Yes	Yes
Rules engine	Yes	Yes	No
Automate data collection by workflow	Yes	No	No
System modification/expansion	Yes	Yes	Yes

of the how the clinical trial look like. IWEIS presents the workflow and provides a workflow based manner for data collection, presentation, and data query. In IWEIS, the data are linked by the workflow. Instead of providing a set of non-related forms, IWEIS presents the workflow and status indicator in each subject profile page. Visit based data management is also supported in IWEIS. Thirdly, in IWEIS, rules can be setup to reduce errors and provide decision support. REDCap integrates a basic method for data validation while IWEIS integrates a rule system which allows setting up more complex rules. Fourthly, although all three systems allow flexibly modification of the system, IWEIS' rule system helps validation and reducing errors during the system modification and expansion. Therefore, IWEIS is able to further facilitate the data capture and management for small scale imaging based clinical trials.

In practice, to support a collaborative clinical trial in a large hospital, the administrator needs to plan thoroughly before the actual deployment of the system. The protocol, workflow, related personnel and each person's responsibility and access to tasks need to be defined carefully. The deployed system can be utilized even in the pilot study for testing. In some cases the patients are screened from the HIS/EHR database, hence the prior background data needs to be retrieved. The system also needs to be hosted on a secure server behind firewall. The training for users is necessary as the users are also involved in other studies or normal clinical environment, which may be confusing and thus, reduce efficiency of the overall clinical workflow.

In addition to improving the efficiency, physicians can also benefit from the system. First of all, the physicians will benefit from the system by saving time on administrative work and focus on the diagnosis and treatment. Moreover, they also gain a platform for future data mining. The advantage of the IWEIS system is that it collects all the data related to the clinical trial from a variety of sources and store the information in a centralized database. This is a significant step in implementing a generalized platform for data mining and decision support. In the future, new patients can be matched to

Please cite this article as: X. Wang, et al., Development of a novel imaging informatics-based system with an intelligent workflow engine (IWEIS) to support imaging-based..., Comput. Biol. Med. (2015), http://dx.doi.org/10.1016/j.compbiomed.2015.03.024

the system through data mining to produce decision support for the new patients' treatment plan. Data analysis toolbox modules are also planned to be integrated with the system which allows researcher to perform statistical analysis online.

5. Limitations and future work

Clinical trials usually enroll patients by advertisement or by screening from HIS (Hospital Information System) or EHR (electronic Health Records system) database. In the second case, some data like prior background data, patient registry data can be reused. Hence an interface between the IWEIS system and HIS/EHR can benefit clinical trials by auto-population of the background data. In the subsequent version, the system will be integrated with HIS/EHR system and be able to receive HL7 messages for patient-specific health data. Once a patient is enrolled, the system will auto-populate the database based on HIS/EHR and reduce the unnecessary efforts for collection of redundant data. The system will also be equipped with a DICOM receiver and be able to communicate with PACS directly. In addition, the current system is developed as a web-based zero footprint application. Therefore, we can integrate with current web-based HIS/EHR through a patient portal.

Compared with custom systems, IWEIS system are limited in functionality and advanced features. Custom systems usually are better designed in software architecture and can meet the specific needs from an individual clinical trial. Moreover, in the generalized system, when the workflow and rules become complex, the possibility of the occurrence of software bugs increases during the integration of various stages and workflow modification. Therefore, development of a generalized system is more complex than a custom system and needs to be tested with more clinical trials with a variety of situations. However, as discussed in the introduction, custom system are more time consuming to develop while the deployment of IWEIS system is able to reduce the cost for development and keep general functionalities. In the long term, with more advanced features developed in subsequent versions of the software, the IWEIS will finally be capable to support complex clinical trials like custom systems.

The workflow expansion and modification feature provides more flexibility than custom systems. However, it may also bring data inconsistency issue and possible conflicts in rules. For example, the participants' datasets collected before the modification of the workflow may be different from new participants which may come after the modification of the system. Another possible example is that new rules require the participant to collect new data while the existing participant has already passed the collection time for new data. In these cases, the researcher needs to mark it as “not applicable” or “missing data” to continue. This is an inherent issue when the workflow is modified in middle of the clinical trial. To address this data inconsistency issue, the future work of the system will be the development of a new feature that can back up various versions of workflow and link the workflow versions with data versions. This feature will facilitate researchers to analyze the data with an intuitive view of the corresponding workflow and help resolve the potential data inconsistency issue.

6. Conclusion

Workflow engines have been applied widely across the business field for years. Complex clinical trials can also benefit from workflow engines to address the demands for data collection, analysis, and distribution. To address these special requirements of imaging-based clinical trials, this paper discussed a novel prototype system with a workflow engine. The system was developed and evaluated for an imaging-based clinical rehabilitation trial named as Dose Optimization of Stroke Evaluation (DOSE) trial and the features of the system have been discussed. By providing flexibility in building and tailoring the workflow in various stages of clinical trials, the system will ultimately save time and reduce errors for clinical trials.

Conflict of interest statement

None declared

Acknowledgment

This work was supported by NIH/NICHD R01HD065438

Reference

[1] ClinicalTrials.gov. ⟨http://clinicaltrials.gov/ct2/resources/trends⟩.
[2] B.J. Erickson, J.C. Buckner, Imaging in clinical trials, Cancer Inform. 4 (2007) 13–18.
[3] J. Parker, E. Coiera, Improving clinical communication: A view from psychology, J. Am. Med. Inform. Assoc. 7 (5) (2000) 453–461.
[4] Ivan Pavlović, Tomaž Kern, Damijan Miklavčič, Comparison of paper-based and electronic data collection process in clinical trials: costs simulation study, Contemp. Clin. Trials 30 (4) (2009) 300–316. http://dx.doi.org/10.1016/j.cct.2009.03.008.
[5] H.K. Huang, PACS and Imaging Informatics: Principles and Applications, John Wiley & Sons, Hoboken, New Jersey, 2010.
[6] P.A. Harris, R. Taylor,R. Thielke, J. Payne, N. Gonzalez, J.G. Conde, Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. J. Biomed. Inform. 2009;42(2):377-381. 10.1016/j.jbi.2008.08.010.
[7] K.L. Rayan, Ko, A computer scientist's introductory guide to business process management (BPM), Crossroads 15 (4) (2009) 11–18. http://dx.doi.org/10.1145/1558897.1558901.
[8] Ximing Wang, Clarisa Martinez, Jing Wang, Ye Liu, Brent Liu, Development of a user customizable imaging informatics-based intelligent workflow engine system to enhance rehabilitation clinical trials, in: Proceedings of the SPIE 9039, Medical Imaging 2014: PACS and Imaging Informatics: Next Generation and Innovations, 90390G, March 19, 2014, doi: 10.1117/12.2044411.
[9] Vojtech Huser, Luke V Rasmussen, Ryan Oberg, Justin B Starren, Implementation of workflow engine technology to deliver basic clinical decision support functionality, BMC Med. Res. Methodol. 11 (2011) 43.
[10] Ximing Wang, Sneha Verma, Yi Qin, Josh Sterling, Alyssa Zhou, Jeffrey Zhang, Clarisa Martinez, Narissa Casebeer, Hyunwook Koh, Carolee Winstein, Brent Liu, Imaging informatics-based multimedia ePR system for data management and decision support in rehabilitation research, in: Proceedings of the SPIE 8674, Medical Imaging 2013: Advanced PACS-based Imaging Informatics and Therapeutic Applications, 86740P, March 29, 2013, doi: 10.1117/12.2008585.
[11] DOSE clinical trial website. ⟨http://pt.usc.edu/ongoing_studies/stroke/⟩.
[12] JQuery. ⟨http://jquery.com/⟩.
[13] Diagramo. ⟨http://diagramo.com/⟩.
[14] Twitter Bootstrap. ⟨http://getbootstrap.com/⟩.
[15] Ximing Wang, Jorge Documet, Kathleen A. Garrison, Carolee J. Winstein, Brent Liu, A multimedia comprehensive informatics system with decision support tools for a multi-site collaboration research of stroke rehabilitation, in: Proceedings of the SPIE 8319, Medical Imaging 2012: Advanced PACS-based Imaging Informatics and Therapeutic Applications, 83190U, February 23, 2012, http://dx.doi.org/10.1117/12.910971.
[16] NanoDICOM. PHP DICOM Toolkit. ⟨http://www.nanodicom.org/⟩.
[17] Hanna Damasio, Antonio Damasio, Lesion Analysis in Neuropsychology, Oxford University Press, Cambridge, UK, 1989.
[18] Laurent Tatu, Thierry Moulin, Julien Bogousslavsky, Henri Duvernoy, Arterial territories of the human brain cerebral hemispheres, Neurology 50 (6) (1998) 1699–1708.
[19] Brent J. Liu, Carolee Winstein, Ximing Wang, Matt Konersman, Clarisa Martinez, Nicolas Schweighofer, An imaging informatics-based ePR (electronic patient record) system for providing decision support in evaluating dose optimization in stroke rehabilitation, in: Proceedings of the SPIE 8319, Medical Imaging 2012: Advanced PACS-based Imaging Informatics and Therapeutic Applications, 83190V, February 23, 2012, doi: 10.1117/12.912952.
[20] Daniel S. Marcus, Timothy R. Olsen, Mohana Ramaratnam, Randy L. Buckner, The Extensible Neuroimaging Archive Toolkit: an informatics platform for managing, exploring, and sharing neuroimaging data, Neuroinformatics 5 (1) (2007) 11–34.

Computerized Medical Imaging and Graphics 46 (2015) 257–268

Contents lists available at ScienceDirect

Computerized Medical Imaging and Graphics

journal homepage: www.elsevier.com/locate/compmedimag

Design and development of an ethnically-diverse imaging informatics-based eFolder system for multiple sclerosis patients

Kevin C. Ma[a,*], James R. Fernandez[b], Lilyana Amezcua[c], Alex Lerner[b], Mark S. Shiroishi[b], Brent J. Liu[a]

[a] IPILab, Dept. of Biomedical Engineering, Viterbi School of Engineering, University of Southern California, Los Angeles, CA 90089-1111, United States
[b] Dept. of Radiology, Keck School of Medicine, University of Southern California, Los Angeles, CA 90089-1111, United States
[c] Dept. of Neurology, Keck School of Medicine, University of Southern California, Los Angeles, CA 90089-1111, United States

ARTICLE INFO

Article history:
Received 29 December 2014
Received in revised form 6 September 2015
Accepted 29 September 2015

SUMMARY

Purpose: MRI has been used to identify multiple sclerosis (MS) lesions in brain and spinal cord visually. Integrating patient information into an electronic patient record system has become key for modern patient care in medicine in recent years. Clinically, it is also necessary to track patients' progress in longitudinal studies, in order to provide comprehensive understanding of disease progression and response to treatment. As the amount of required data increases, there exists a need for an efficient systematic solution to store and analyze MS patient data, disease profiles, and disease tracking for both clinical and research purposes.
Method: An imaging informatics based system, called MS eFolder, has been developed as an integrated patient record system for data storage and analysis of MS patients. The eFolder system, with a DICOM-based database, includes a module for lesion contouring by radiologists, a MS lesion quantification tool to quantify MS lesion volume in 3D, brain parenchyma fraction analysis, and provide quantitative analysis and tracking of volume changes in longitudinal studies. Patient data, including MR images, have been collected retrospectively at University of Southern California Medical Center (USC) and Los Angeles County Hospital (LAC). The MS eFolder utilizes web-based components, such as browser-based graphical user interface (GUI) and web-based database. The eFolder database stores patient clinical data (demographics, MS disease history, family history, etc.), MR imaging-related data found in DICOM headers, and lesion quantification results. Lesion quantification results are derived from radiologists' contours on brain MRI studies and quantified into 3-dimensional volumes and locations. Quantified results of white matter lesions are integrated into a structured report based on DICOM-SR protocol and templates. The user interface displays patient clinical information, original MR images, and viewing structured reports of quantified results. The GUI also includes a data mining tool to handle unique search queries for MS. System workflow and dataflow steps has been designed based on the IHE post-processing workflow profile, including workflow process tracking, MS lesion contouring and quantification of MR images at a post-processing workstation, and storage of quantitative results as DICOM-SR in DICOM-based storage system. The web-based GUI is designed to display zero-footprint DICOM web-accessible data objects (WADO) and the SR objects.
Summary: The MS eFolder system has been designed and developed as an integrated data storage and mining solution in both clinical and research environments, while providing unique features, such as

Abbreviations: CNS, central nervous system; CSF, cerebral-spinal fluid; DICOM, digital imaging and communication in medicine; EDSS, expanded disability status scale; ePR, electronic patient record; FLAIR, fluid attenuated inverse recovery; GUI, graphical user interface; IHE, integrating the healthcare enterprise; HTTP, hypertext transfer protocol; LAC, Los Angeles county hospital; MRI, magnetic resonance imaging; MS, multiple sclerosis; PACS, picture archiving and communication system; PHP, PHP: hypertext preprocessor; SC, secondary capture; SR, structured reporting; UID, unique identification; USC, university of Southern California; WADO, web-accessible DICOM object; WAMP, windows/Apache/MySQL/PHP; XML, extensible markup language.

* Corresponding author. Tel.: +1 213 821 8395.
E-mail address: kevincma@usc.edu (K.C. Ma).

http://dx.doi.org/10.1016/j.compmedimag.2015.09.007

quantitative lesion analysis and disease tracking over a longitudinal study. A comprehensive image and clinical data integrated database provided by MS eFolder provides a platform for treatment assessment, outcomes analysis and decision-support. The proposed system serves as a platform for future quantitative analysis derived automatically from CAD algorithms that can also be integrated within the system for individual disease tracking and future MS-related research. Ultimately the eFolder provides a decision-support infrastructure that can eventually be used as add-on value to the overall electronic medical record.

1. Introduction

1.1. Multiple Sclerosis and MRI

Multiple Sclerosis (MS) is a demyelinating disease in which the patient's central nervous system (CNS) degenerates and causes inflammation and brain and spinal atrophy [1,2]. The scarred tissues leftover from the autoimmune attacks are lesions, or plaques, in the white matter [3]. As the disease progresses, the effect on a patient's life can be devastating. Disability can continue to progress until an effective treatment regimen is developed for an individual patient to slow disease progression. Direct medical costs and associated assistive and rehabilitative expenditures can quickly reach into tens of thousands of dollars per year.

There is currently no cure for MS because of the irreversible axonal injuries and the complex nature as to what triggers the autoimmune attacks [3,4]. MS manifests itself differently amongst different ethnicities [5,6], such as differing prevalent symptoms, disabilities, MS lesion locations, and response to treatments. The purpose of an MRI scan is to visually locate lesions in the central nervous system. MS lesions appear hypo- or iso-intense in a T1-weighted scan and hyper-intense in a T2-weighted and FLAIR (fluid attenuated inverse recovery) scans. Location and morphology of lesions are also used to identify lesions caused by multiple sclerosis [7,8]. Fig. 1 shows the three aforementioned axial slices of a patient's brain with multiple sclerosis.

MS disease progression can be monitored by subsequent MR scans for changes in lesion characteristics in longitudinal studies. In general, following initial diagnosis, disease activity is assessed by the number of relapses per year (relapse rate), accumulation of disability as measured by the expanded disability status scale (EDSS) [9] and changes in MRI lesion characteristics. The progression of the disease is variable, and requires routine follow-up imaging studies to document disease exacerbation, improvement, or stability of the characteristic MS lesions.

1.2. Current challenges with MS disease management and research

There are a few challenges that exist in the current status for the diagnosis and treatment of multiple sclerosis in both clinical and research settings. MS is a complicated disease that manifests itself differently for different individuals as well as different ethnicities. Therefore, MS disease management and treatment become complex and highly individualized. A more detailed study of specific clinical markers at the time of presentation and during the disease course is required to ascertain and define the distinct features of MS.

Currently, MS lesion quantification requires a manual approach to lesion measurement on MRI [9]. It is time-consuming to quantify lesion load (total lesion volume in a study), and volumetric calculations with current techniques only yielding, at most, a rough estimate. The task of tracking MS lesion changes becomes even more difficult if several MR longitudinal studies of the same patient (follow-up studies) require quantitative comparison, which is currently not utilized in clinical diagnosis. In addition, it has been shown that MS diagnosis, lesion detection and lesion load calculation suffer from inter- and intra-observer variability [10].

1.3. MS eFolder design concept

In order to build a complete system to aid in multiple sclerosis diagnosis, treatment, and research, we have designed and developed an ethnically-diverse imaging informatics-based system, called MS eFolder, designed specifically for MS clinical care and research. The MS eFolder provides a database solution for storing patients' clinical information, an effective way for accessing patients' MR images and associated data, and also provides a lesion quantification tool. By integrating these components compliant to the DICOM standard in the clinical environment, the eFolder system is able to provide a data repository for treatment planning,

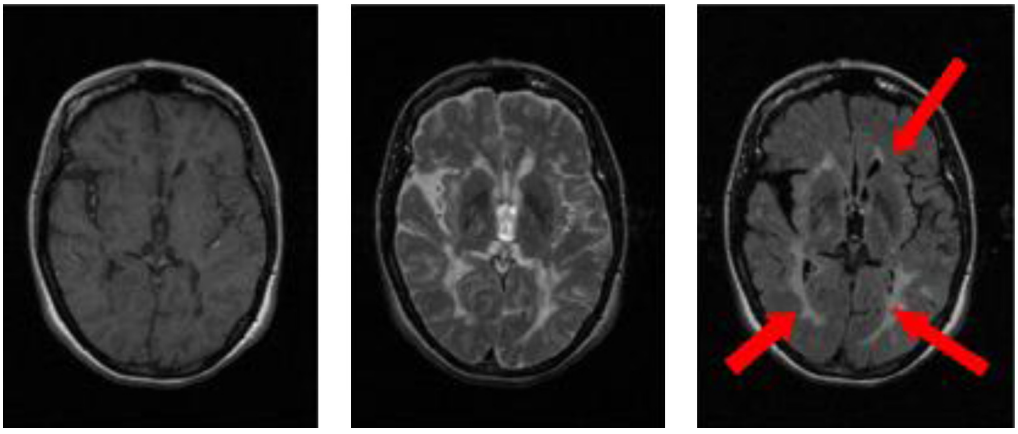


Fig. 1. Three axial brain images of an MS patient. The left-most image is T1-weighted, the middle image is T2-weighted, and the right-most image is FLAIR. White regions in FLAIR image (as pointed by arrows) indicates MS lesions in white matter.

correlate MR lesion quantification data with clinical manifestations more effectively, and provide a data mining tool for research purposes. In the future this system can serve as the decision-support component in an overall EMR for treatment of MS patients. The concept of eFolder is derived from electronic patient records (ePR). ePR is a digital, comprehensive patient database that stores patients' demographic information, medical history, and any information that may be needed and included in ePR's designs. For example, an ePR developed for surgery [11] may contain a subset of patients that have a particular back pain related disease, for example, patients undergoing spinal discectomy.

The MS eFolder is developed to aid in diagnosis, disease tracking, treatment, and research of MS. After examining the needs of clinicians and for research purposes, the system has three main software components:

1. A DICOM-based patient record database: The database needs to store the patients' various information, such as their demographic information, medical history, MS history, and any survey results that are gathered by researchers and clinicians. The system also needs to link to medical images (in this case, patients' brain and spinal MRI) to aid in diagnosis and disease tracking. MR images and DICOM-SR objects are stored in the eFolder's image repository component, which will be explained in detail in the next section.
2. Image processing module: In order to quantify 3-dimensional white matter lesion volumes and brain matter volumes, an image processing module is included in the MS eFolder system. The module allows users to manually identify and contour lesions in the brain MRI. Another algorithm is able to segment gray matter, white matter, and cerebro-spinal fluid (CSF). A voxel-based algorithm is able to cluster those contours and segmentations and quantify volumetric data. Its advantages include a standardized quantification process, allowing MS lesion tracking in longitudinal studies, and offering a quantifiable results for clinical trials and research. The module is integrated with the database, thus the results are stored and patient data can be queried via lesion characteristics. While the module currently utilizes manual contours to create MS lesion segmentation for volumetric calculations, we can replace the manual contouring component with an automatic or semi-automatic computer-aided detection (CAD) algorithm. The eFolder system's modular design allow the components to be replaced with minimal disruption of system services.
3. A web-based graphical user interface (GUI): A GUI is needed for patient data viewing, input, management, and data querying. It is web-based such that the system can be accessed via Hypertext Transfer Protocol (HTTP) and Internet connections. The GUI needs to be comprehensive such that all patient data, including diagnostic MR images and quantification results, are all available for viewing. The interface has the potential to be a powerful tool in data mining for researchers and data gathering for both clinical and research purposes.

The last step of MS eFolder system design is to integrate all of the components of eFolder together and fit into a real-life clinical workflow. IHE (Integrating the Healthcare Enterprise) provides a design of how to include a post-processing step inside a typical clinical workflow [12]. The MS eFolder system workflow is modeled after this IHE workflow profile to ensure successful DICOM-based integration when the system is available for clinical use.

2. Material and methods

Fig. 2 shows the components diagram of the MS eFolder system. The eFolder system consists of three components: eFolder web-based services, the image processing and quantification module, and a graphical user interface.

The system diagram also includes DICOM services (shown in Fig. 3 as orange block on the left) for receiving and storing MR images and DICOM SR objects, and also directing image workflow within the eFolder system. Patient data and images are stored in the eFolder archive server. The DICOM receiver acts as the gateway between image source and image archive. The control device directs image dataflow within the eFolder server, and the data archive acts as image repository in the system. The following sections detail design and development of these components.

2.1. EPR design

The eFolder database, written in MySQL, stores text data including patient history, MR image locations, and lesion quantification

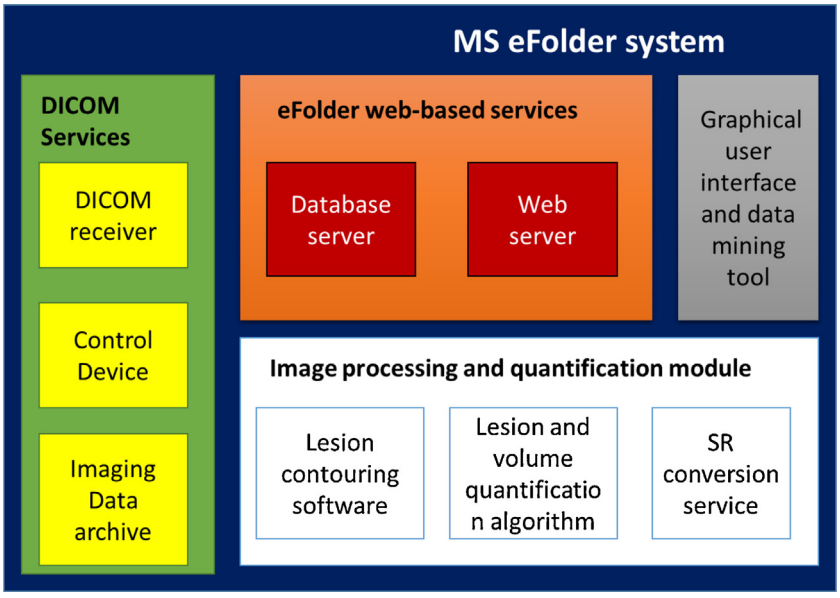


Fig. 2. Modular diagram of Multiple Sclerosis (MS) eFolder system. SR: structured reporting; DICOM: Digital Imaging and Communications in Medicine.

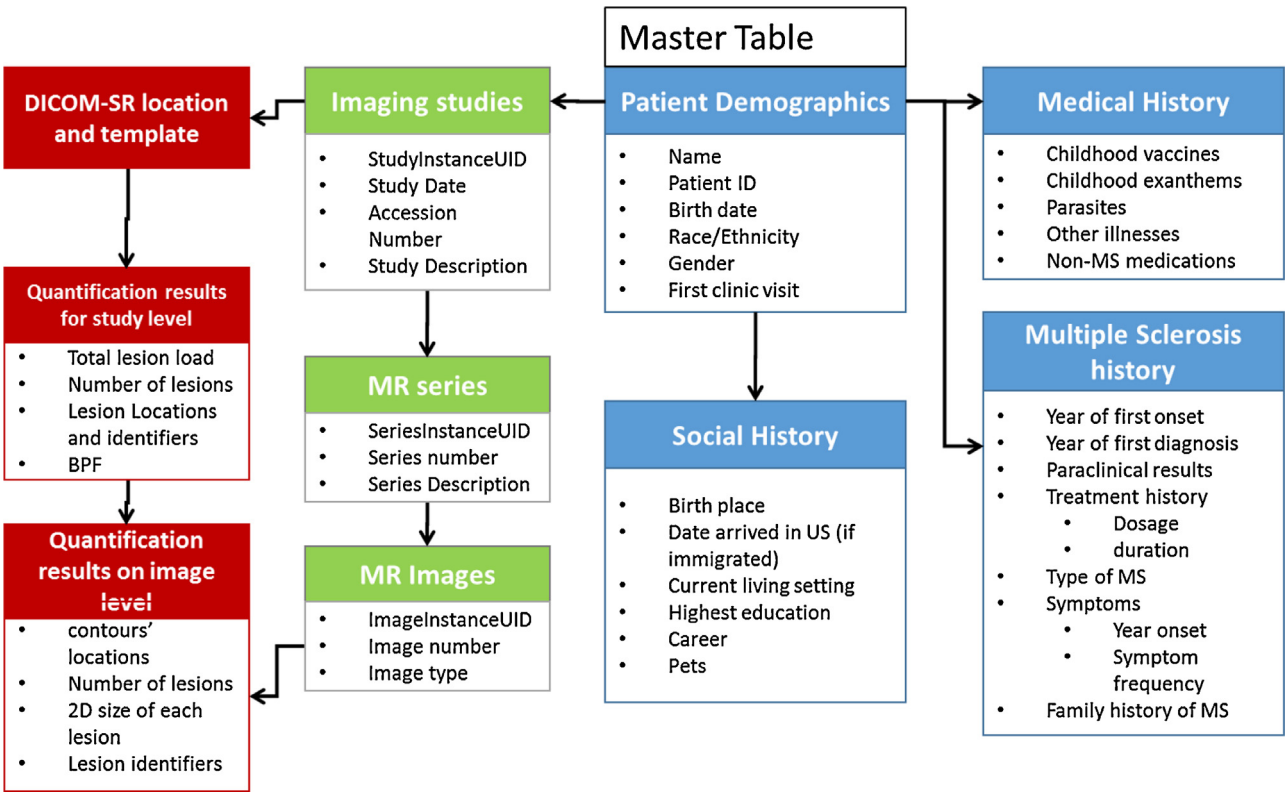


Fig. 3. General database schema for the MS eFolder. Arrows indicate how each table is related to each other in the schema.

results. The database structure is built such that one single patient has a unique data entry regarding demographics and social data, has a list of all MR studies regarding to MS, and a list of all quantification results (in SR format) available for that patient. Fig. 3 shows the general database structure.

The information stored in demographic data includes name, gender, ethnicity, birthplace, and social history. The medical history database module stores records of any childhood illnesses, vaccines, and other past medical histories. The MS history database includes the year of first diagnosis/symptom, any family members with MS, MS type, MS symptoms and frequency, treatment history, and so forth. Data collection is conducted by patient interviews reading of medical reports, and physician inputs. Columned items are collected from surveys designed by neurologists and research project leaders.

The imaging database stores patients' MR studies. The database structure is designed following the DICOM standard, from the patient level to the study level, series, and finally individual images. Data in the imaging database is automatically populated via PHP-based DICOM parsing scripts. Table values, such as study instance UID (unique identification) and series instance UID are parsed directly from the header files of the uploaded DICOM images. DICOM patient identifiers are matched with patient records in the demographics database during the uploading procedure via web-based GUI.

The lesion quantification algorithm produces results from both the study level and at image level. At the study level there are total lesion load, number of lesions, three-dimensional lesion centroid coordinates, lesion sizes in three-dimensions, and the overall report of findings and three-dimensional view of the lesion contours. At the image level, the program calculates lesion locations on that image, size of each lesion on the image, and (x,y) coordinates of lesion contours. Lesion volume data calculation and population of database is covered in the next sections.

2.2. Lesion detection and quantification

The eFolder system is designed to include a post-processing module that detects MS lesions and calculates lesion volumes in the white matter region of central nervous system. The purpose of the module is to objectively quantify lesion volumes and able to track lesion volume changes in longitudinal studies. The quantified changes enables users to track disease progress and monitor treatment responses over time. The module receives raw DICOM data from the modality or any other image data source and automatically detects abnormal voxels as MS lesions. Then, the lesion quantification algorithm is designed to output lesion volumes, lesion locations, and total lesion load from lesion contours on the brain MRI.

While the lesion detection component is being concurrently explored and developed, we have designed the system based on the assumption that the detection and quantification component has been completed and able to generate reliable results. Therefore, the system is modular, and the detection component can be replaced by any new or existing MS lesion detection algorithms. In recent years, there have been several interests and attempts at MS lesion detection with various degrees of success [13–15]. Any existing or future MS lesion detection and segmentation can be implemented in the MS eFolder system because of the modular design.

With this in mind, we currently use lesion contours manually completed by radiologists and evaluated by neuroradiologist fellows in order to achieve better accuracy and detection and contouring, as well as completing overall system development. The MRI studies are uploaded to Synapse3D® software at Healthcare Consultation Center II (HCCII) at the University of Southern California. With the software, users are able to manually draw lesion contours directly on the DICOM images. The contours are exported via DICOM files. The output of lesion identification is then evaluated by the quantification algorithm.

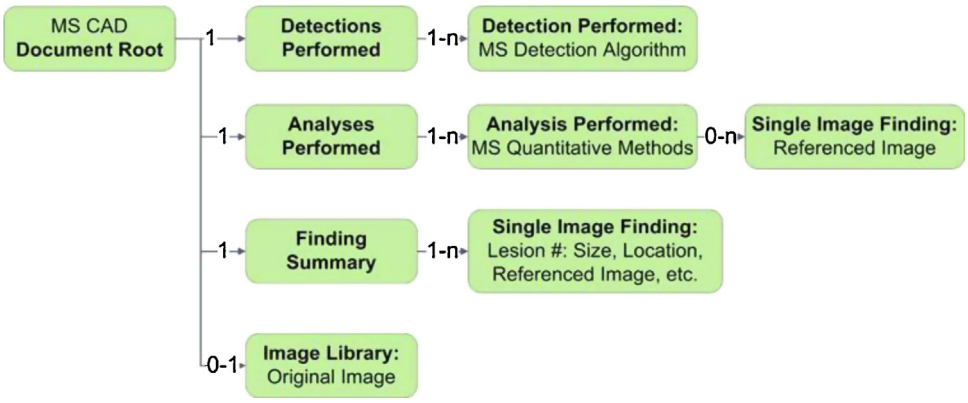


Fig. 4. DICOM-SR template used for the image processing module for the eFolder, as published by Le, A. (2009). [17] The figure shows the tree structure that can be stored in DICOM-compliant file structure.

Lesion voxels in the binary segmentation are clustered in 3-D with 26-connectivity principle. The lesion load (total lesion volume in the brain) and the number of lesions can be calculated using these clusters. Lesion volume is obtained by multiplying number of voxels in a lesion and the voxel size, which is extracted from DICOM headers of images. Lesions are separated into three subgroups: small (<1 cm^3), medium (between 1 and 5 cm^3), and large (>5 cm^3). Lesion load is obtained by summing up all lesion volumes, and lesion locations are identified by the coordinates of their centroids.

2.3. Quantification results output and reporting

The initial output of the lesion quantification program is a MATLAB file containing total lesion load, lesion coordinates, volume of each lesion, number of lesions, 2D MR slices containing lesion contours and lesion reconstructed in 3-D space. The algorithm produces output data that needs to be standardized and integrated with other system components for data queries and look-ups. DICOM structured reporting (DICOM-SR) provides templates to include quantification results in a standardized format [16]. The structured reporting would allow searching, storage, and comparison with other similar data better than traditional paper report format.

For integrating the image processing module in the overall system, we are using the methodology outlined in a CAD–PACS integration toolkit. A CAD–PACS integration toolkit is used to store CAD results in a Picture Archiving and Communications System (PACS) environment [14]. While the eFolder system currently does not use CAD programs and CAD results to identify MS lesions, the output of lesion quantification algorithm is considered as post-processed results similar to CAD outputs and can be stored via a DICOM-SR. An integration of quantification results with DICOM-SR is designed and implemented to directly store quantification results in a tabulated format in the eFolder, as well as to produce DICOM structured reports that may be displayed from a PACS workstation and stored in a conventional PACS. The utilization of DICOM-SR and CAD–PACS toolkit for MS eFolder is for a more streamlined workflow in a standardized way that can be implemented in existing PACS environments. Fig. 4 is a template designed for storing lesion quantification results.

After designing the custom template for lesion quantification results, Fig. 5 shows the workflow steps of converting MATLAB-based quantification results into DICOM-SR objects.

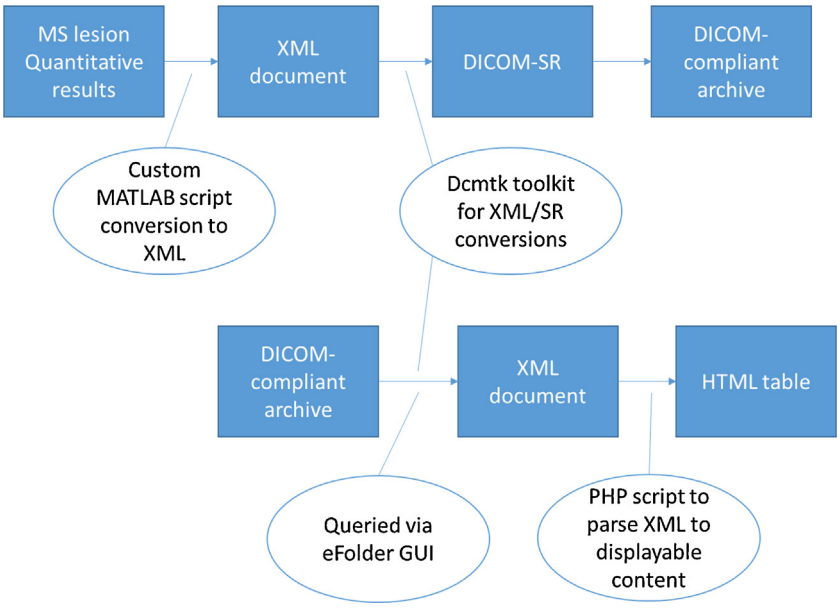


Fig. 5. Workflow diagram for DICOM-SR conversion and display. Top row: converting MS quantification results to DICOM-SR for storage. Bottom row: querying and displaying information within DICOM-SR objects.

Fig. 6. XML document to store MS lesion quantitative analysis results based on DICOM SR template.

The first step is to convert quantitative results to an XML (Extensible Markup Language) document, which then is converted to DICOM-SR via the dcmtk open-source toolkit [18]. In order to generate the correct XML document, sample DICOM SR files obtained and converted to XML. The XML template then is modified and customized to store the quantification values. A MATLAB script is used to convert quantitative output from a study to the XML file according to the customized template. Fig. 6 shows a partial screenshot of the resultant XML document.

2.4. Visualization

The following are the design criteria for the eFolder user interface:

- The GUI needs to be web-based to allow remote access using thin-client architecture. Computations and visualizations are completed on the server side for a light-weight and fast GUI.
- The GUI needs to be comprehensive. It needs to display patient clinical data, imaging data, and quantification results on the same interface. It allows physicians and radiologists to access all of the information related to the data query.
- The system needs to be dynamic and allow display of 3D images and manipulations of images presented. An organized viewing interface allows for a more clarified presentation
- The GUI needs to allow flexible and intelligent data mining. With a large number of patients' information stored in the eFolder system, any clinician and researcher should be able to look up MS patients on a variety of different search criteria, ranging from patient demographic data to lesion analytical results.

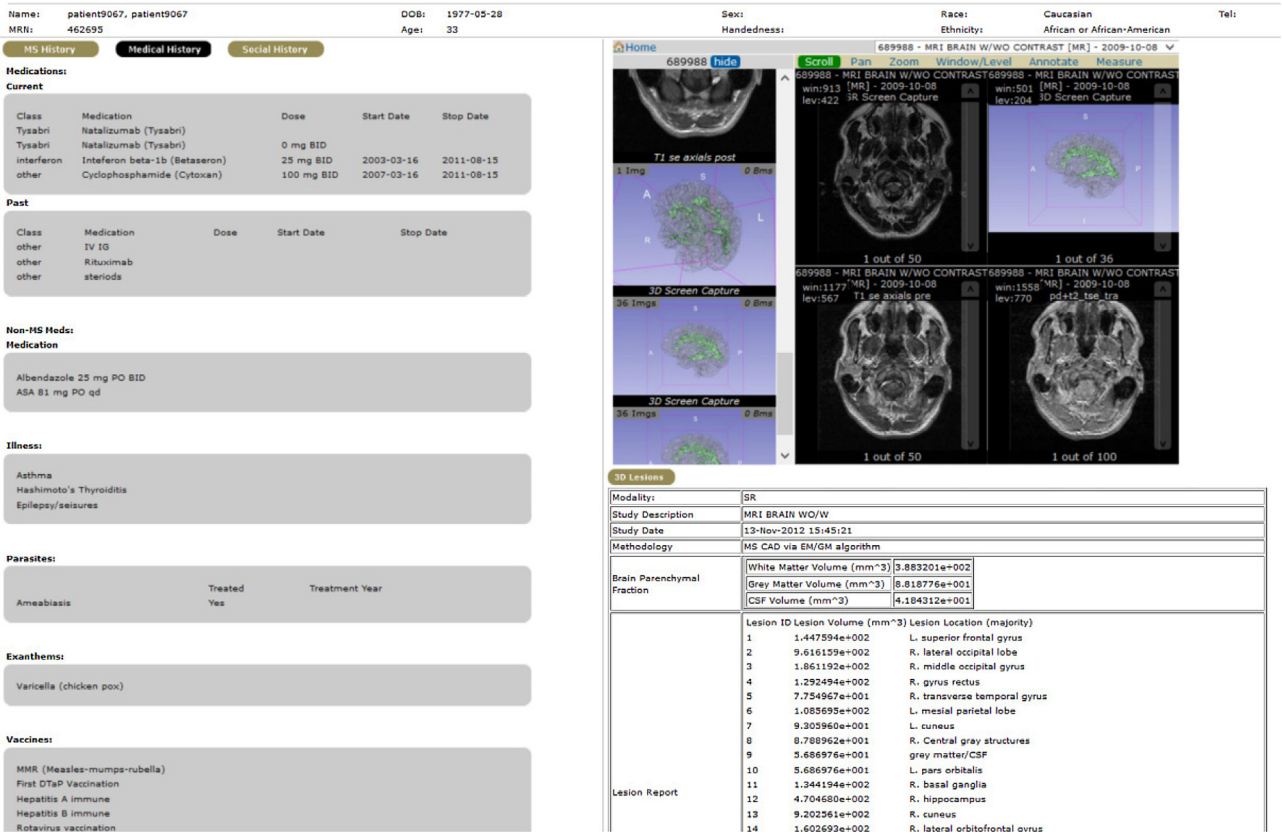


Fig. 7. Screenshot of comprehensive MS eFolder web-based GUI. Left panel: patient's clinical and demographic data. Right top panel: WADO image viewer embedded in eFolder GUI. Right bottom panel: SR document content in tabulated format.

Base on PHP scripting language, the dynamic GUI guides the user to look up a specific patient's disease history with images, and it allows querying for patients with various different criteria. Viewing of DICOM images allows zoom, pan, window/level, and scrolling on the webpage. Fig. 7 shows a screenshot of the comprehensive web-based GUI for MS eFolder.

The developed GUI fits all of the design criteria described above. The GUI is web-based and hosted on a secure HTTP server. Users are able to log in using his or her credentials to view patient data. The main page of the GUI displays all of the essential patient information, such as name, ID, date of birth, gender, and ethnicity. The GUI is then divided up into different tabs, both within the main GUI as well as separated browser tabs. The main window module of patient information (on the left panel of Fig. 7) contains "MS History", "Medical History", and "Social History". User may click on each tab to switch between viewing the three data sets. The top right module contains the WADO-based image viewer embedded within the main window. The user may view the imaging studies separately in a full-size DICOM image viewer tab. The right bottom panel displays the DICOM-SR result from the patient's study in a tabulated format. The main window's purpose is to aggregate all of the information and display it in a clear and organized way.

The GUI also allows patient data input and patient data look-up via web-based forms. The data lookup feature is able to allow user to access all of the patients within the query criteria in the form. The criteria include patient clinical data as well as imaging and lesion quantification data. This feature is designed to allow data mining of eFolder patient database, allowing user to access a large data source to obtain data. The resultant data analysis may create observable trends that were previously difficult to obtain. For example, user may look up patients that have had certain drug treatment for over 5 years, while simultaneously looking up if the patients has had imaging scans during that time. The changes in lesion volumes and locations may shed a light on treatment's effects on the imaging level over time.

In addition, the system GUI is designed for user to view quantification results visually via the WADO viewer. DICOM secondary capture, or DICOM-SC, are DICOM image objects generated by post-processing clients and software packages. In the eFolder project, DICOM SC are image captures of lesion contour results. DICOM-SC, along with SR, aids users to visualize MS lesion contours in both 2D and 3D space. The secondary captures of contour results are converted to the DICOM format for storage and display from a DICOM image viewer, for streamlining the DICOM-compliant workflow in the MS eFolder system. For 2D image captures, a MATLAB script is used to overlay 2D lesion contours on top of the FLAIR axial images to create a new series under the DICOM study. DICOM-SCs are stored in the archive and can be viewed alongside the original DICOM images in the DICOM web-based viewer. Fig. 8 shows how the DICOM images and SCs are viewed in a WADO (web accessible DICOM objects) viewer.

2.5. Integration

The MS eFolder design is modeled after the IHE workflow profile to show its use in a clinical environment. To accomplish the tasks, a simulated clinical environment with MS eFolder is set up in the Image Processing and Informatics laboratory (Fig. 9).

The workflow for MS eFolder integration is defined in five steps:

1. MR images are sent from modality simulator to the eFolder server for archiving

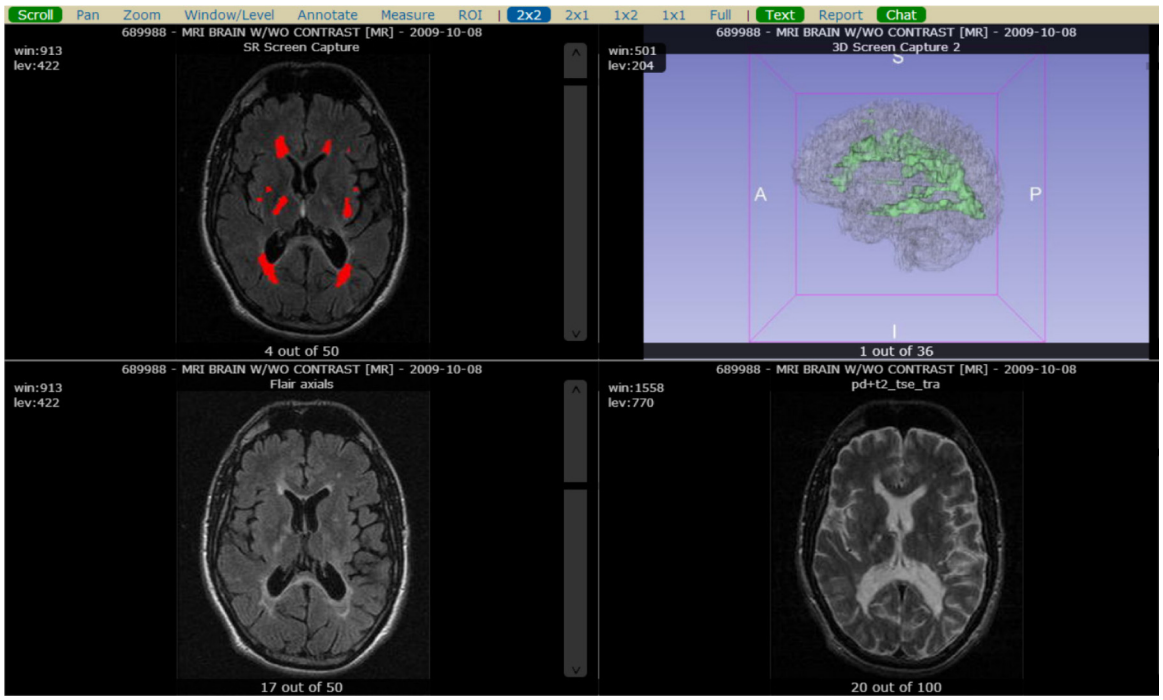
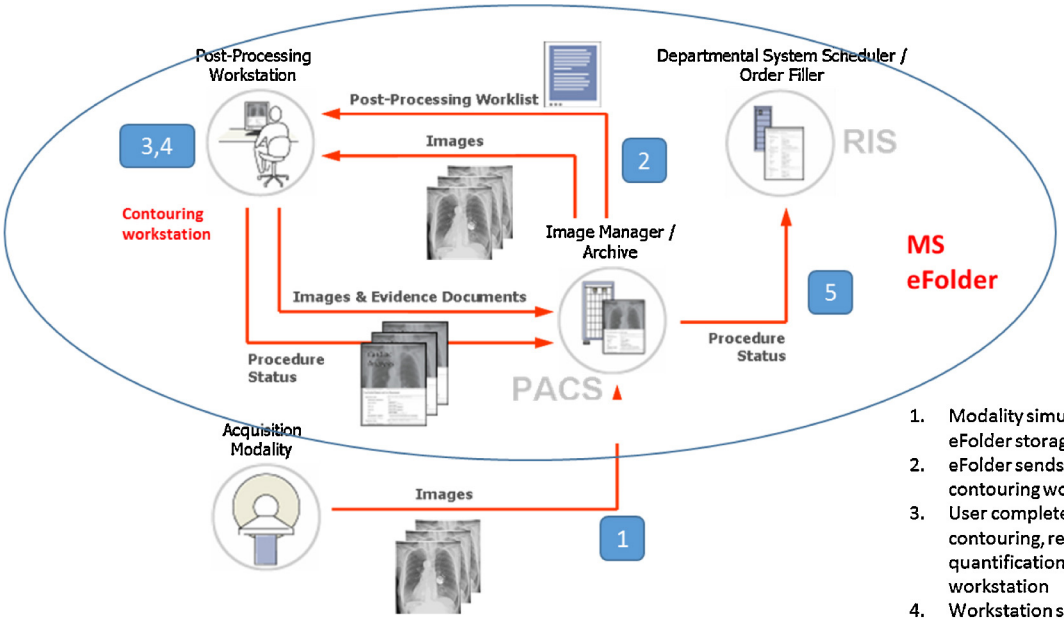


Fig. 8. Screenshot of WADO viewer displaying DICOM images and DICOM-SC for the MS eFolder. Upper left panel displays the 2D contour overlay, upper right panel displays the 3D lesion rendering from different angles by the Slicer 3D software.

2. The eFolder server sends a copy of the images to the post-processing workstation for postprocessing analysis
3. User at the contouring workstation is able to access the images via contouring software
4. Lesion contour files are analyzed by the volume quantification module, which converts quantification results into DICOM_SR and sent back to the eFolder server for storage
5. At the completion of each of the previous steps, a status tracking tool inside eFolder displays alerts of the study progress to the user

Fig. 10 displays the physical components diagram of the MS eFolder system setup in order to simulate the workflow steps outlined in Fig. 11.

The hardware environment is set up in a laboratory environment. There are two Windows-based desktop towers, one performs as the eFolder archiving and web server and the other as the post-processing workstation. Another laptop is connected to the network to simulate the user accessing the eFolder system. The eFolder network is LAN-based, connected via Ethernet cables and managed by a router. The router, a D-Link EBR 2310



1. Modality simulator sends images to MS eFolder storage
2. eFolder sends copy of image to contouring workstation
3. User completes lesion identification and contouring, results are sent to the quantification module on the same workstation
4. Workstation sends quantification results (in DICOM-SR) back to storage
5. eFolder GUI accesses image, CAD data, and monitors workflow progress

Fig. 9. MS eFolder workflow diagram with IHE postprocessing profile. The steps 1 through 5 indicates the order of workflow of the demonstration.

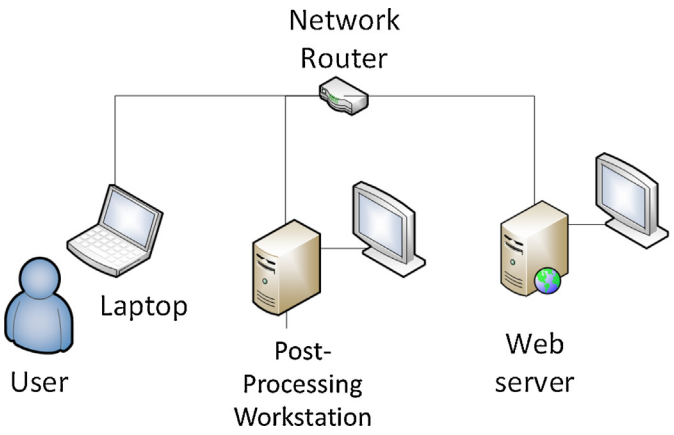


Fig. 10. MS eFolder components diagram. The entire system includes 3 hardware components (Windows-based personal computers) and a closed network for demonstration purposes.

series wired router, assigns unique IP addresses to the hardware components in the network. The eFolder web/archiving server is a Dell® Dimension 9150 series desktop tower, running Windows XP Professional. A WAMP (Windows Apache/MySQL/PHP) server program is installed on the desktop to run the eFolder's web-based interface, PHP scripts such as database connection and DICOM parsing, and the MySQL-based eFolder database. The post-processing workstation is a Dell® Dimension 9200 series desktop tower, running Windows 7 Professional. MATLAB version 2011b is installed on the post-processing workstation to run the volume quantification module. The post-processing workstation also includes web-based Synapse3D interface for lesion contouring. The laptop is used to access the system, and there are no system requirements as long as the laptop has a web browser to view the eFolder GUI. This allows the eFolder system to be accessed on the network by portable devices such as tablet computers or smart mobile phones. Fig. 11 shows a status tracking module of the MS eFolder interface that tracks a study's progress in the eFolder study acquisition workflow.

3. Results and discussion

3.1. Data collection

The MS eFolder system requires patient clinical data and MR imaging data for designing and building the infrastructure in this phase. In order to build and validate the system, sample patient data are acquired from the Department of Neurology and Department of Radiology at University of Southern California (USC). There are two main goals of data collection: the first goal is to acquire sample data to build the system infrastructure, and the second goal is to acquire an adequate number of data to complete system validation and evaluation.

Research regarding differences in Hispanic American MS patients is the primary motivation for data collection among the different ethnic groups. Compared to the detailed MS studies regarding African American and Asian ethnic groups, there has not been these kind of studies performed related to Hispanic MS patients. A research project on Hispanic-American MS patients has been ongoing at USC to observe any significant differences in disease manifestation or correlations between Hispanic and Caucasian MS patients. The MS eFolder is thus currently designed to fit the data model of Hispanic American-related research data.

For each MS patient, two types of data are collected: clinical data and imaging data. Patients' clinical data are collected via physicians' questionnaires about their social and medical backgrounds. Patient candidates' criteria include having been diagnosed of multiple sclerosis, having had appointments at the Neurology department at USC, and having had MR brain images (T1-w, T2-w, and FLAIR axial slices) taken at various imaging centers and available in digital format at USC. Patient data collection has begun at USC since 2008. Patients visiting the MS clinic were asked to fill out a questionnaire form regarding their basic information, social history, ethnic background, medical history, and childhood medical history. A sample form is shown in Fig. 12.

All brain MR studies are collected at University of Southern California Academic Medical Center and Los Angeles County Hospital. A total of 72 patients are collected: 36 Hispanic and 36 Caucasian patients. The patients of two groups are matched by gender, age (within 5 years), disease duration (within 5 years), and disease type (all are relapse-remitting). For each patient, the images, in DICOM format and anonymized, are acquired from Siemens® Symphony Maestro 1.5T and GE® Signa HDt 1.5T MR scanners and include at least T1, T2, and FLAIR axial images as well as sagittal/coronal views, diffusion images, etc.

3.2. Lesion quantification and data analysis results

Fig. 15 shows the manual contour results from Caucasian and Hispanic patients, comparing total lesion volume and disease duration.

Fig. 13 is also an example of data mining tool of MS eFolder, shown in Fig. 2 as part of the GUI. The system is able to relate disease duration, obtained from patients' medical history, and lesion volume, which is obtained from MS eFolder's image processing

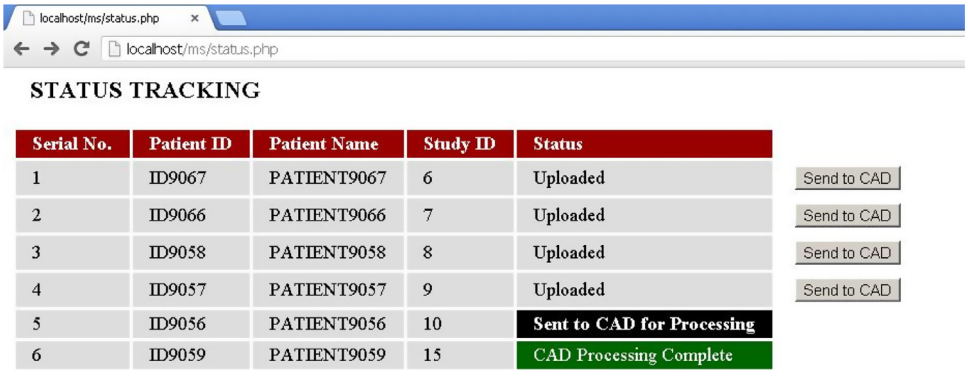


Fig. 11. Status tracking page for MS eFolder post-processing workflow after a study (Study ID 10) has been sent to the post-processing (CAD) workstation, and after another study (Study ID 15) has been processed and sent back to the server.

MULTIPLE SCLEROSIS QUESTIONNAIRE

Study ID number: _____ Today's Date:___/____/_____

Directions: Read and answer the short-answer questions in the space provided. You may also answer with a check mark and/or circle all the answers that may apply. We have tried to make the questions quite specific. If, however, some confusion should arise, feel free to ask any research team member. We sincerely appreciate your participation.

I.- PARTICIPANT INFORMATION:

Name: _____

Date of Birth: _____

Gender: ☐Male ☐Female

Please circle: Right handed or Left handed or Ambidextrous

II.- SOCIAL HISTORY:

a) Birth Place: _____

b) Residence for age under 15 yrs: _____

c) Time in the United States: _____

d) Family Generation: _____

e) Type of Residence: ☐Rural or ☐Urban

f) Highest level of education: ☐Elementary ☐Middle School ☐High School ☐College ☐Post Graduate

g) Occupation:

Present: _____ Past: _____

h) Have you been exposed to pesticides? ☐Yes or ☐No

i) Do you smoke? ☐Yes or ☐No If Yes, for how many years? _____

Fig. 12. First out of 6 pages of Multiple Sclerosis questionnaire to collect patient data.

module. The comprehensive data collection, storage, and analysis highlights the eFolder's capability of integrating patient data into a complete patient profile.

A longitudinal study tracking viewer module has been developed to help users track a patient's disease progress visually. Fig. 14 shows an example of data tracking of 4 longitudinal studies, and Fig. 15 shows comparison of lesion contours from two studies of the same patient. The two figures have both been included in the web-based GUI as separate viewing modules.

The GUI is able to display the lesion contour results in its quantitative analysis module, which can be viewed in a separate browser. The quantitative results, longitudinal analysis results, and lesion comparison results between the initial patient groups are all presented in the analysis module

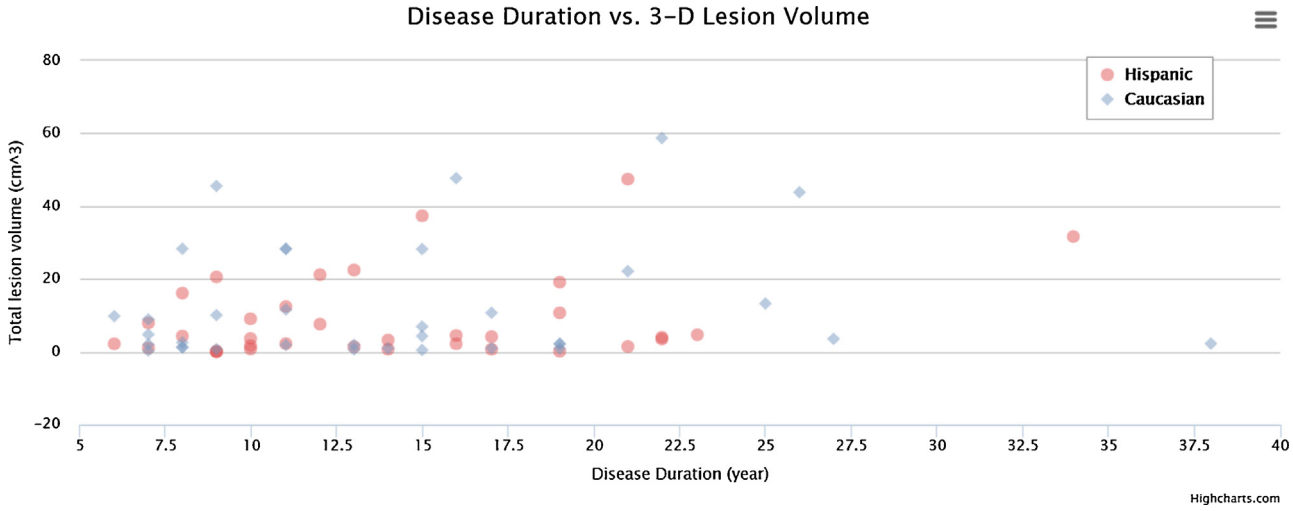


Fig. 13. Scatter plot of total lesion volume between Caucasian and Hispanic patients.

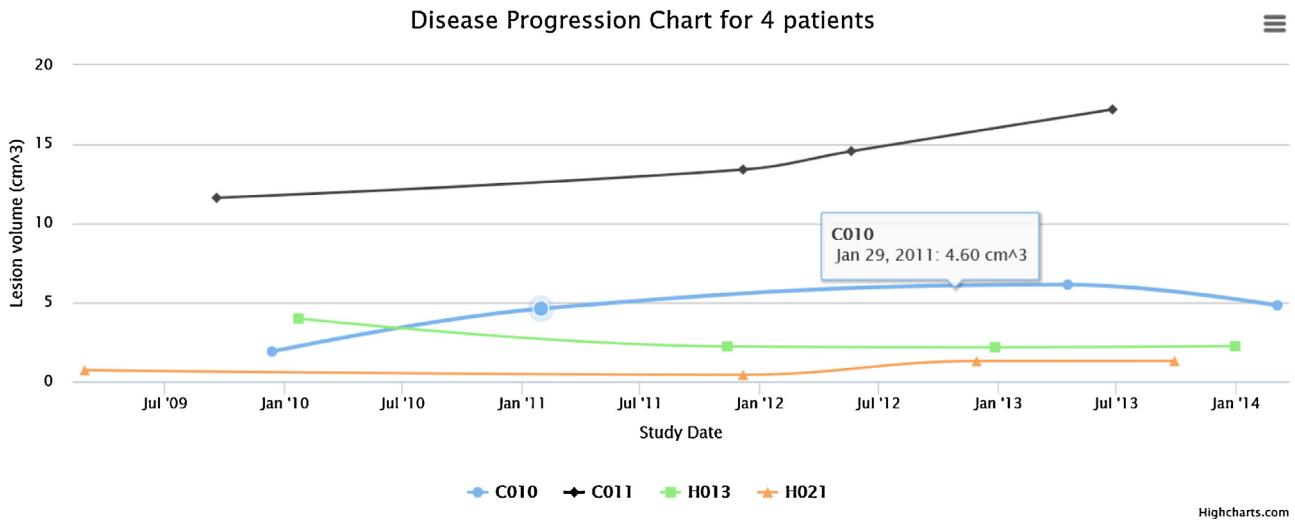


Fig. 14. Lesion volume of 4 patients, each with 4 studies from 2009 to 2014. Each dot represents a CAD result showing the overall lesion volume for that study.

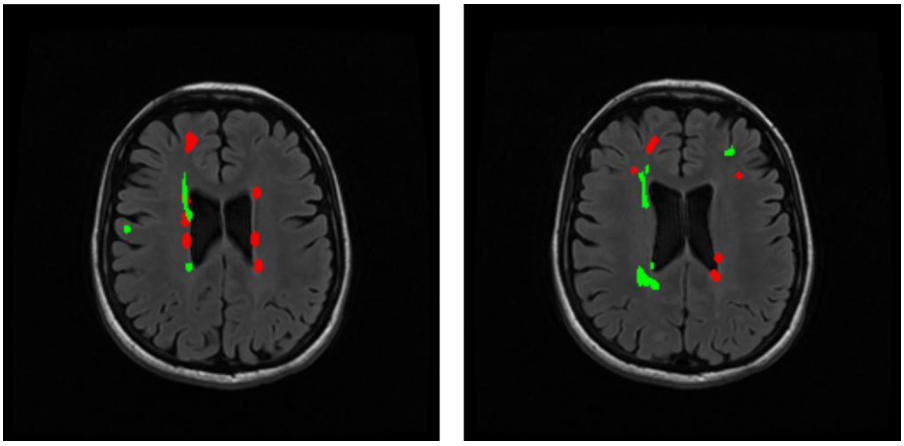


Fig. 15. Lesion contours from two separate studies superimposed on the same image. Red/darker contour is from a study in 2011, and green/lighter contour is from 2014. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.3. Preliminary system development results

Preliminary tests of the MS eFolder system with integrated IHE workflow were successful in the laboratory environment. All 72 patients' information have been entered into the MS eFolder database, and the MR studies have been uploaded into the system for data transfer and data upload based on the DICOM protocol. Radiologists have completed manual contouring on all 72 patients plus any additional studies being collected. The output of manual contouring steps have successfully been quantified by the quantification module, and DICOM-SR objects of quantification results have been generated and stored. The comprehensive eFolder GUI and database has been successfully completed and is available for viewing via a web browser. The DICOM-SR conversion workflow has been completed and successfully tested within the system. Overall, the comprehensive GUI design and layout to display lesion quantification data in DICOM-SR has been completed.

The eFolder system will be implemented in a clinical environment as part of the future work. The system will be evaluated based on system performance, data uploading and viewing time, as well as patient data look-ups and data mining modules. The system will be able to display a patient's profile, with treatment history and disease changes over time, which includes quantified lesion changes. Users will also be able to look up lesion data and clinical data for a large group of patients in order to observe and discover any disease trends. The eFolder system, combined with very large amount of data generated in the clinical environment, is capable of performing complex big data analysis.

3.4. Future work

A major goal for our current development is to develop a reliable and accurate lesion detection algorithm in order to decrease manual contouring man-hours.

Our current progress on developing an automatic lesion voxel classification algorithm is based on Statistical Parametric Mapping (SPM) brain image analysis toolkit for MATLAB [19]. At first, gray matter and white matter of the brain are segmented via voxel intensity estimation and probabilistic maps provided from SPM, which utilizes voxel-based morphometry [20], or VBM. VBM of brain MRI involves spatially normalizing all of the images in a study into a standardized stereotactic space. The algorithm then extracts the gray and white matter voxels from normalized space, performs smoothing, and then performs a Bayesian statistical analysis to further calculate gray and white matter voxel probabilities [21]. The output of this SPM-based preprocessing is a probabilistic map of gray matter and white matter segmentation. After gray matter, white matter, and CSF have been segmented, we can calculate brain parenchymal fraction, as well as starting on segmenting MS lesions in white matter.

Once the automatic detection component is completed, the module will be integrated into the overall eFolder system by replacing the manual contouring component. The eFolder system has been designed in such a way that DICOM-compliant modules can be inserted into the system without significant changes in the overall system workflow. The automatic lesion detection will enhance the overall eFolder system, making it more efficient and consistent in quantifying lesion volume changes and achieving overall system goals. Since the eFolder system is designed with IHE integration protocols, any new or existing lesion segmentation programs can be integrated into the eFolder system with DICOM-compliant outputs and protocols.

There are a few other areas of the eFolder system that need further work and is currently under development. First, 3-dimensional DICOM-SC objects is being developed to be displayed in WADO viewer. Currently, a 3-D contour object and is created via the Slicer 3D software. We would like to able to automatically create 3-D views directly from WADO viewer.

Second, the completed eFolder system with IHE post-processing workflow needs to be evaluated in a real clinical setting. Performance evaluations and user feedbacks will be used to improve system performance and GUI design. The current GUI can be more interactive to include dynamic referencing on individual lesions in WADO viewer and DICOM-SR viewer. Additional features will be discussed and planned according to user feedback.

4. Conclusions

In this paper, we present the design and development a comprehensive imaging-informatics based eFolder system for multiple sclerosis. It combines the concept of electronic patient record along with disease-centric database design, MR images, and an automated lesion quantification tool for an easier, more efficient system for longitudinal disease tracking, decision support, data mining, and data repository for both MS clinical and research environments. The eFolder system is capable of aiding the comparison studies between Hispanic-American and Caucasian MS patients via integrating MRI and clinical findings. The system is DICOM-complaint, while being completely web-based to allow remote access and telemedicine. MS lesions are identified and contoured by expert neuroradiologists, and a lesion quantification system has been developed in MATLAB. Quantification results have been converted into DICOM-SR for DICOM-compliant long-term storage. System workflow, based on IHE, has been developed and tested in a laboratory environment. Imaging data and patient disease data have been collected, while a graphical user interface has been developed to bring ease of access and user-friendliness to the MS eFolder. While complete integration is ongoing and several improvements are being refined, the eFolder concept aims to improve on MS diagnosis, tracking, and research. The eFolder will bring a novel and comprehensive approach to observe longitudinal MS lesion changes in MR, and deciding treatment plans that best suit the MS patient's profile.

Acknowledgements

Kevin C. Ma and James R. Fernandez were supported by NIH NIBIB T32 EB00438 Training Grant.Mark S. Shiroishi was partially supported by SC CTSI (NIH/NCRR/NCATS) Grant # KL2TR000131.

References

[1] Benedict R, Bobholz J. Multipl sclerosis. Semin Neurol 2007;27(1):078–85.
[2] Ringold S, Lynn C, Glass R. Multiple Sclerosis. JAMA 2005;293(4):514.
[3] Hafler D. Multiple sclerosis. J Clin Invest 2004;113(6):788–94.
[4] Hafler D. The distinction blurs between an autoimmune versus microbial hypothesis in multiple sclerosis. J Clin Invest 1999;104:527–9.
[5] Cree B, Khan O, Bourdette D, et al. Clinical characteristics of African Americans vs Caucasian Americans with multiple sclerosis. Neurology 2004;63:2039–45.
[6] Yamasaki K, Kira J, Kawano Y, et al. Western versus Asian types of multiple sclerosis: immunogenetically and clinically distinct disorders. Ann Neurol 2004;40(4):569–74.
[7] Tas MW, Barkhol F, van Walderveen MA, et al. The effect of gadolinium on the sensitivity and specificity of MR in the initial diagnosis of multiple sclerosis. Am J Neuroradiol 1995;16:259–64.
[8] Lovblad KO, Anzalone N, Dorfler A, et al. MR imaging in multiple sclerosis: review and recommendations for current practice. Am J Neuroradiol 2010;31:981–9.
[9] (a) Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 1983;33:1444–52;
(b) Miller DH, et al. MRI outcomes in placebo-controlled trial of natalizumab in relapsing MS. Neurology 2007;68:1390–401.
[10] Filippi M, Horfield MA, Rovaris M, et al. Intraobserver and interobserver variability in schemes for estimating volume of brain lesions on MR images in multiple sclerosis. Am J Neuroradiol 1998;19:239–44.
[11] Documet J, Le A, Liu B, et al. A multimedia electronic patient record system for image-assisted minimally invasive spinal surgery JCARS 2010;5(3):195–209.
[12] (Dec. 23rd, 2014) IHE Post-processing Workflow ⟨http://wiki.ihe.net/index.php?title=Post-Processing_Workflow⟩.
[13] Ghribi O, Njeh I, Ben Hamida A, Zouch W, Mhiri C. Brief review of multiple sclerosis lesions segmentation methods on conventional magnetic resonance imaging. In: Advanced technologies for signal and image processing (ATSIP). 2014. p. 249–53, http://dx.doi.org/10.1109/ATSIP.2014.6834616.
[14] Admasu F, Al-Zubi S, Toennies K, Bodammer N, Hinrichs H. Segmentation of multiple sclerosis lesions from MR brain images using the principles of fuzzy-connectedness and artificial neuron networks. In: Image processing. 2003., http://dx.doi.org/10.1109/ICIP.2003.1246873.
[15] Roy PK, Bhuiyan A, Ramamohanarao K. Automated segmentation of multiple sclerosis lesion in intensity enhanced flair MRI using texture features and support vector machine. In: Image processing (ICIP). 2013. p. 4277–81, http://dx.doi.org/10.1109/ICIP.2013.6738881.
[16] Hussein R, Engelmann U, et al. DICOM structured reporting. Radiographics 2004;24(May):891–6.
[17] Zhou Z, Liu B, Le A. CAD-PACS Integration Tool Kit Based on DICOM Secondary Capture, Structured Report and IHE Workflow Profiles. In: Computerized Medical Imaging and Graphics (June). 2007. p. 346–52, http://dx.doi.org/10.1016/j.compmedimag.2007.02.015.
[18] (Dec. 23rd, 2014) DCMTK - DICOM Toolkit ⟨http://dicom.offis.de/dcmtk.php.en⟩.
[19] Friston KJ. Statistical parametric mapping the analysis of functional brain images. London: Academic; 2007 (Print).
[20] Ashburner J, Friston KJ. Unified segmentation. NeuroImage 2005;26.3:839–51 (Web).
[21] Ashburner J, Friston KJ. Voxel-based morphometry—the methods. NeuroImage 2000;11.6:805–21 (Web).

Computer Aided Drafting, Design and Manufacturing
Volume 25, Number 1, March 2015, Page 28

CADDM

Vessel Extraction on Ocular Fundus Images by Using Gabor Filter Bank

SU Ming-jian[1], *ZHANG Xue-jun*[1,2], *WANG Xi-ming*[2], *Brent J Liu*[2], *GAO Xin*[3], *ZHANG Zuo-jun* [4], *ZHOU Bin*[1]

1. School of Computer, Electronics and Information, Guangxi University, Nanning 530004, China;
2. IPILab, Biomedical Engineering Dept, University of Southern California, Los Angeles, CA 90033;
3. Department of Medical Imaging, Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou 215163, China;
4. Cancer Hospital of Guangxi Medical University, Nanning, China.

Abstract: Fundus diagnosis is an important part of the whole body examination that may provide rich clinical information to doctors for diagnostic reference. Manual fundus vessel extraction is helpful to quantitative measurement of diseases but obviously it is a tough work for physicians. This paper presents an automatic method by using Gabor filter bank to extract the artery and vein separately in the ocular fundus images. After preprocessing steps that include gray-scale transform, gray value inversion and contrast enhancement, the Gabor filter bank is applied to the extraction of the artery and vein in the ocular fundus images. Finally these two different width types of vessels are selected by post-processing methods such as labeling, corrosion, binarization, etc. Evaluation results show an accurate rate of 90% in vein and 82% in artery from 20 cases, that indicates the effectiveness of our proposed segmentation method.

Key words: Gabor filter bank; vessel extraction; ocular fundus; segmentation

1 Introduction

Ocular fundus has the greatest concentration of arterial, venous and capillaries in human body, which can be only observed with the naked eye. It can reflect dynamic blood circulation and body health, such as hypertension, diabetes[1], tumours of the brain[2], Uremic[3], etc. Therefore, the funduscopic examination is one of the important indexes for physical examination. Retinal blood vessels are small with dense distribution, and manual retinal vessel extraction could extract ocular fundus artery and vein, cross, capillary hemangioma, vascular proliferation and bleeding spots. However, it is a very time-consuming task. According to recent studies, some automatic methods have been undertaken to facilitate this procedure. The Gabor filter was used to enhance blood vessels by exploiting their shape feature[4], and the pattern feature of human face[5] or fingerprint[6] could be also segmented by the Gabor filter bank methods. According to their results, we suppose that a fundus images with different vessel directions and thickness should have a similar output patterns after the procedure of Gabor filter bank. Chang et al.[7] use a 12 direction line operator to apply to the filtered fundus images by the Canny edge detector, and gradient orientation based feature detection method is also used for extracting retinal blood vessels[8]. However, simply using edge detectors couldn't identify the pattern direction or thickness. Lin et al.[9] introduced a vessel segmentation method that is free from background interference. Other methods such as machine learning or filtering by various classifiers are also proposed in [10], [11] and [12]. Although their results are very promising, none of these methods could extract fundus artery and vein separately, and with higher time complexity than our method. This paper presents an automatic method for vessel segmentation with different line widths that includes: enhancement of the fundus images to yield a best contrast of the image; separation of the fundus artery and vein base on the 2D Gabor filter bank extraction; finally the whole performance is evaluated by comparing our results with gold standard drawn by

Project Item: Supported by National Natural Science Foundation of China (Nos. 61262027, 45627390), and Research Foundation Project of the Guangxi Ministry of Education (No. 200810MS048).

Corresponding author: ZHANG Xue-jun, Male, Ph.D., Professor, **E-mail:** xjzhang@gxu.edu.cn.

doctors.

2 Methodology

2.1 Gabor Filter Bank

(1) 1-Dimension Gabor filter

1-D Gabor filter can be defined as the product of a Gaussian kernel times a complex sine function.

E.g. $g(t)=k\times w(at)*s(t)$

Here,

$w(t)=e^{-\pi t^2}$,

$s(t)=e^{j(2\pi f_0 t)}$,

$w(at)s(t)=(\sin(2\pi f_0 t+\theta), j\cos(2\pi f_0 t+\theta))$.

The k, θ, f_0 are filter parameters. We could think of the complex Gabor filter is made up of a real and a complex function. The real part is:

$$g_r(t)=w(t)*\sin(2\pi f_0 t+\theta)$$

And the imaginary part is:

$$g_i(t)=w(t)\cos(2\pi f_0 t+\theta)$$

(2) 2-D Gabor filter

Study on biology discovered that the distribution of 2D Gabor filter can well describe the acceptable information domain single cells in primary visual cortex part of the vertebrate brain, and it has the similar characteristics both in the spatial and the frequency domain that is consistent with the human visual system. Gabor transform has been proven to be the best description of the signal spatial domain and frequency domain in the 2D uncertainty case. Therefore, theory and experimental basis based on the image segmentation with Gabor filter is adequate.

The 2D Gabor wavelet transform is a powerful tool for multiscale image representation and analysis. 2D Gabor filter function can be represented by:

$$\varphi_j(x)=\frac{\|\vec{k}_j\|^2}{\sigma_1\sigma_2}\exp(-\frac{\|\vec{k}_j\|^2\|\vec{x}\|^2}{2\sigma_1^{\ 2}}-\frac{\|\vec{k}_j\|^2\|\vec{y}\|^2}{2\sigma_2^{\ 2}})$$

$$\cdot[\exp(i\vec{k}_j\vec{z})-\exp(-\frac{\sigma_1\sigma_2}{2})]$$

$$\vec{k}_j=\begin{pmatrix}k_{js}\\k_{jy}\end{pmatrix}=\begin{pmatrix}k_v\cos\varphi_u\\k_v\sin\varphi_u\end{pmatrix}$$

Here $\vec{z}$ is the coordinate of a given position of image $\vec{x}$, $\vec{y}$ that represents for the horizontal coordinate and the vertical coordinate in the image metrics (one gray pixel for one $\vec{z}$ value). $\vec{k}_j$ is the center frequency of the filter, φ_u is the filter's direction selectivity. $\|\vec{k}_j\|^2/\sigma_1\sigma_2$ is used for compensation energy spectrum decay dependent on the frequency. $\exp(-\frac{\|\vec{k}_j\|^2\|\vec{x}\|^2}{2\sigma_1^{\ 2}}-\frac{\|\vec{k}_j\|^2\|\vec{y}\|^2}{2\sigma_1^{\ 2}})$ is used for constraint the 2D Gauss envelope. σ_1, σ_2 determine the scale and shape of Gauss envelope. When $\sigma_1=\sigma_2$, Gauss envelope is round, if $\sigma_1\neq\sigma_2$, Gauss envelope is oval. $\exp(i\vec{k}_j\vec{z})$ is the complex 2D function, where the real part is 2D cosine function $\cos(\vec{k}_j z)$, the imaginary part is 2D sine function $\sin(\vec{k}_j z)$. In order to obtain a zero DC response, we minus $\exp(-\sigma_1\sigma_2/2)$ in the real part. This makes the 2D Gabor wavelets transform not affected by the absolute value of gray level of the image, and the image is not sensitive to the change of illumination.

2D Gabor wavelet is composed of 2D Gabor filter function through a set of filter generated from scaling and rotation, while the selection of its parameters are usually considered in frequency space. In order to sample the entire frequency of an image, the sampling method has a plurality of center frequency and direction space. Parameter σ determines the bandwidth of the filter, the relationship between the two parameters is

$$\sigma=\sqrt{2\ln 2}\left(\frac{2^{\phi}+1}{2^{\phi}-1}\right)$$

Here, ϕ is the half-magnitude bandwidth, when ϕ=1 octave, $\sigma\approx\pi$. Here is some illustration of our Gabor filters:

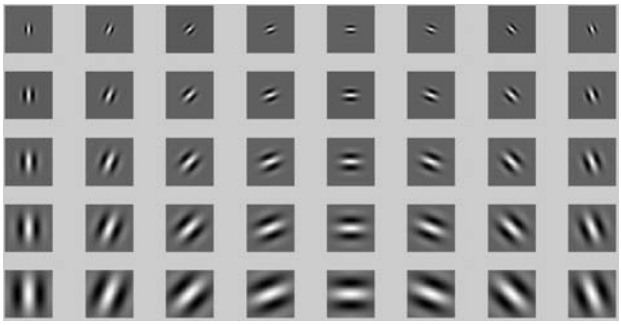


Fig.1. Real part of 5 scales 8 directions of GABOR filter.

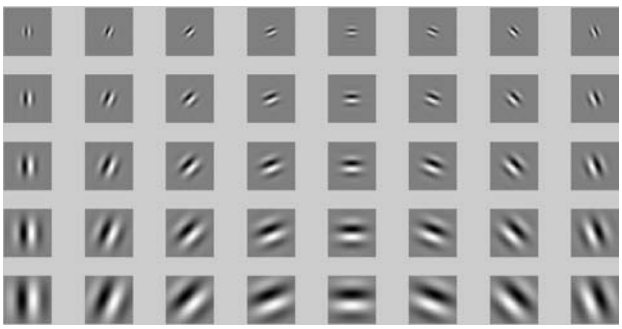


Fig.2. Imaginary part of 5 scales 8 directions of GABOR filter.

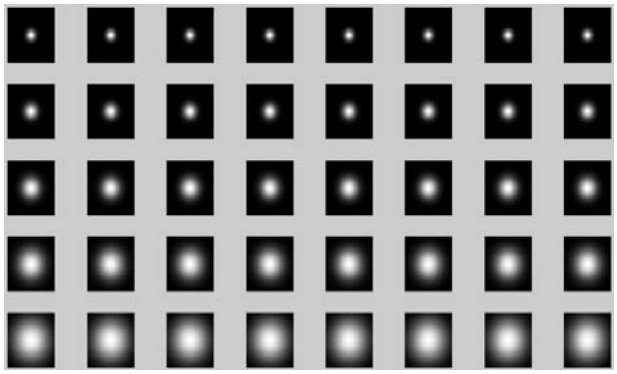


Fig.3. 5 scales 8 directions GABOR filter.

These images illustrate single Gaussian envelope is made up of Gabor filter, each have different scales and directions. Fig.1 represents the formula

$$g_r(t) = w(t) * \sin(2\pi f_0 t + \theta),$$

Fig.2 represents the formula

$$g_i(t) = w(t)\cos(2\pi f_0 t + \theta)$$

and Fig.3 represents the formula

$$g(t) = \sqrt{g_r^2(t) + g_i^2(t)}.$$

According to the Gaussian envelope term $\exp(-\frac{\left\|\overrightarrow{k_j}\right\|^2\left\|\vec{x}\right\|^2}{2\sigma_1^{\ 2}} - \frac{\left\|\overrightarrow{k_j}\right\|^2\left\|\vec{y}\right\|^2}{2\sigma_1^{\ 2}})$, if $\sigma_1 \neq \sigma_2$, Gauss envelope is oval, if $\sigma_1=\sigma_2$, Gauss envelope degenerate to round. The description in Matlab language is: G(t)= exp(-(Kv*Kv*((x+5)*(x+5))/(sigma1))-(Kv*Kv*((y-8)*(y-8))/(sigma2))), degenerate to round is: G(t)'= exp(-(Kv*Kv*((x+5)*(x+5)+(y-8)*(y-8))/(2*sigma1*sigma2))) .

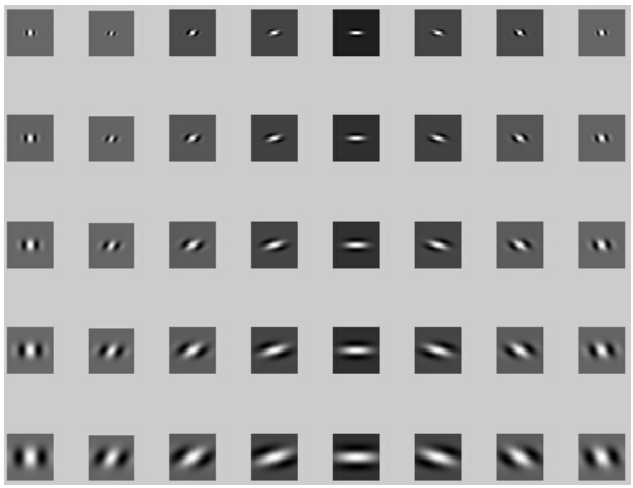


Fig.4. Real part of 5 scales 8 directions of oval GABOR filter.

Here *Kv* could adjust the scale of Gaussian envelope. The smaller *Kv* we choose, the bigger Gaussian envelope could generate. The *x*+5, *y*-8 represent the shift of Gaussian envelope in the range of filtering flat.

(3) Gabor filter bank

A 2-D Gabor filter bank can be made up of several 2-D Gabor filters contain in one filter bank. It could be describe in Matlab language as follow:

```
bank(t) = w(1) + w(2) + ..... w(n)
w(1) = sqrt(GaborReal(:,:,j).^2 + GaborImg(:,:,j).^2);
Where
   GaborReal(:,:,j) = Kv*Kv*tmp1*tmp2/sigma.^2;
   GaborImg(:,:,j) = Kv*Kv*tmp1*tmp3/sigma.^2;
tmp1 = exp(-(Kv*Kv*(j*j+i*i)/(2*sigma.^2)));
tmp2 = cos(Kv*cos(θ)*i+Kv*sin(θ)*j) - exp(-sigma.^2/2);
tmp3 = sin(Kv*cos(θ)*i+Kv*sin(θ)*j);
...
And so is
w(n)=sqrt(GaborReal_n(:,:,j).^2 + GaborImg_n(:,:,j).^2).
```

Here, we use different sigma, *Kv*, θ to form each Gabor kernel. A Gabor filter bank could be formed after the convolution is applied to each $w(n)$ by sinusiod. bank(t) is Gabor filter bank, (x,y) is the center of Gaussian kernel in Cartesian coordinates plane. We can shift a Gaussian kernel position by changing (x,y) value. Changing the term gsize could setup the scale of Gaussian kernel. Here is an example of a dual-kernel Gabor filter bank which is implemented by Matlab:

$bank(t) = w(1) + w(2)$;

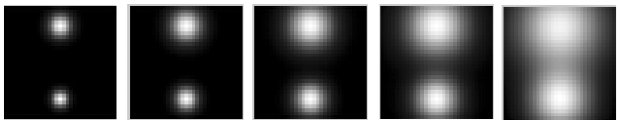


Fig.5. A matlab instance of dual-core Gaussian envelop.

2.2 Sample Pattern Simulation

At the early stages of the experiment, we use the Gabor filter bank apply to test patterns. There are test pattern Fig.6, Fig.7 and Fig.8. The test patterns are as follows.

Fig.6 shows 6 directions thin lines, Fig.7 shows 6 directions thicker lines, and Fig.8 shows a joint picture is made up of Fig.6 and Fig.7. First of all, we use single Gabor filter convolve Fig.6 with the parameter *sigma*=5. The result pattern is shown in Fig.9, which

we could find out the most accurate signal response appear in row 2. We also use the parameter *sigma*=π which generates another result show in Fig.10.

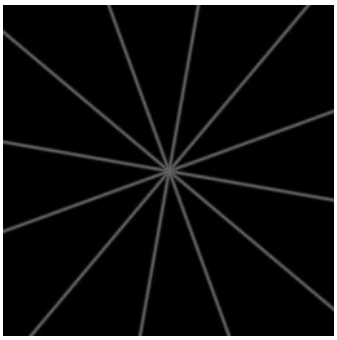


Fig.6. Test image with 6 directions of thin line.

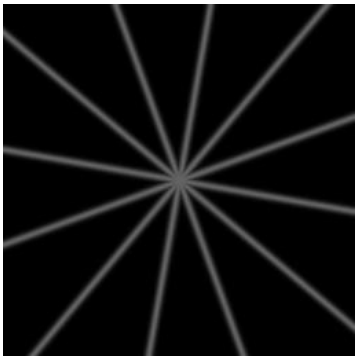


Fig.7. Test image with 6 directions of thicker line.

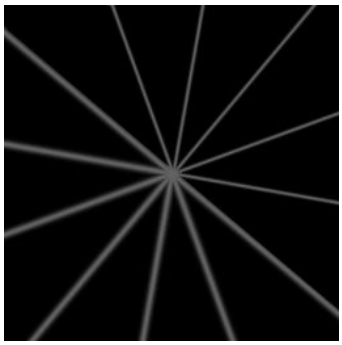


Fig.8. A combination image of Fig.6 and Fig.7.

We can find out a regularly phenomenon by comparing with Fig.9 and Fig.10. If we pick parameter sigma appropriate, we should get to the specific direction of the line in the result image. If we choose sigma smaller, the result set will show multi-direction of the lines. In *sigma*=5 case, the lines of directions of 3π/8, 4π/8, 7π/8, π appears properly, corresponding to the angle of complex sinusoid in Gabor kernels.

According to the method proposed in 2.1(3), a two kernels Gabor filter bank with parameter *sigma*1=3π, *sigma*2=0.5π, θ_1=0.25π, θ_2=π was applied to our test image Fig.8. Fig.8 contains 2 different width line in 6 directions, which represent 2 two kinds of different widths vessel such as artery and vein. Actually the fundus artery and vein have many different width and directions, so the image with 2 width and 6 direction is only a ugly test. In the case of gabor filter bank with two round Gauss kernels, the result image of testing Fig. 10 shows the line which perpendicular to filtering direction is almost vanished. The position of each Gauss kernel in the Gabor filter bank does not affect the filtering effect. A spatial domain filter is range limited. Fortunately the fundus artery and vein in actual case is not so thick, this help us putting several small size Gauss kernel in a limited filtering flat. The preprocessed fundus images are convolved with our Gabor filter bank to extract the superior temporal venule of retina and venula temporalis retinae inferior. The results show that these veins can be enhanced successfully. The bigger Gauss core parameter σ we choose, the smaller veins will vanish. We can also find out when the σ is smaller, two lines will appear on the result image, which makes the bigger vessel looks like 2 vessels.

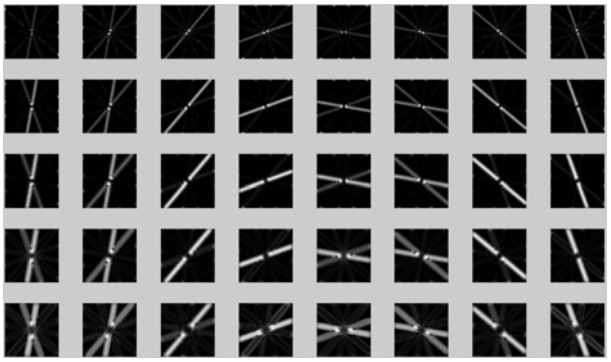


Fig.9. Result image of Gabor filter of complex sinusoid angel 0-π with para. *Sigma*=5.

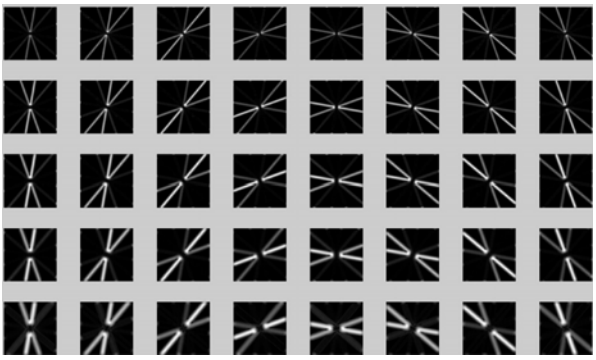


Fig.10. Result image of Gabor filter of complex sinusoid angle 0-π with para. *Sigma*= 3.14.

2.3 Preprocessing

Our method is based on the 2D Gabor filter which enables us to select a certain width of a line by adjusting its parameters. Since the fundus artery and vein have different width, it is possible to extract each type of the vessel individually without using typical spatial image processing method. The preprocessing steps include: Color to gray image conversion, image reversal, and contrast enhancement.

(1) Color to gray image conversion

Because the retinal image provided by the hospital is color, the weighted average method is used to transfer RGB color image into gray image by the formula:

$$Gray = 0.299R + 0.587G + 0.114B$$

(2) Image reversal

Then we have an image reversal by the MATLAB realization:

$$B=255-A$$

Here A is the gray matrix; B is the output gray matrix.

(3) Contrast enhancement

The original retinal image is color (R, G, B) image, which is transferred into gray matrix using the RGB2 gray function provided by MATLAB. Then the contrast linear transformation is applied to adjust the contrast by a formula described as below:

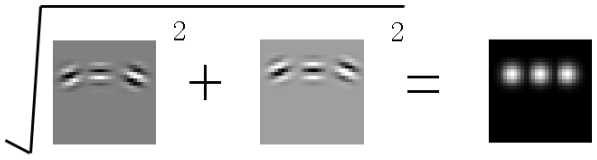
$$D_B = f(D_A) = f_A D_A + f_B$$

Where D_A represents the input gray image, and D_A has a relatively large contrast in order to optimize the experiment effect. D_B represents the gray image output, f_A is the gradient of the linear function, and f_B is the intercept of the linear function in the Y axis.

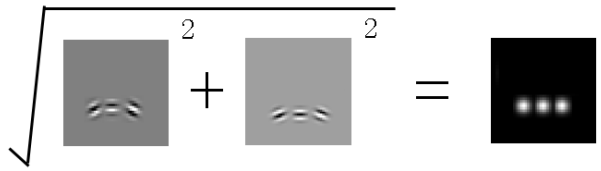
Generally speaking, Gabor transform can only deal with 1- channel image data, and Gabor transform is based on Fourier transform. A color picture should be degenerate into 0-255 gray pixel image. A Fourier transform mainly focuses on the source energy which refers to a picture's gray level. The original retinal image is color (R, G, B) image, vein is in a dull color compare with the fundus environment and account for a small proportion of image pixels. We use the method (1) and (2) raising the gray level of vein meanwhile lower the environment pixel gray level to almost zero. This procedure will improve the efficiency of Fourier filtering in frequency domain and denoise in the following step.

3 Experimental Results

In order to verify the effect of this method, a number of retinal images had been selected by a physician Z. Zhang M.D. and tested in our experiment. After preprocessing the gray image, the Gabor filter bank is used to select the vessels with different widths. The choice of filter parameters depends on the resolution of the image samples. The resolution of our datasets is 584*565. According to the anatomical structure of fundus, the superior temporal venula of retina comes out of optic disc with its main branch spreading at 45 degrees, then turns to left and goes downward. While the venula temporalis retinae inferior comes out of optic disc with minus 45 degrees turning on the left, then keep spreading horizontally.



(a) σ=0.2, θ=pi, 0.75pi, 0.25pi



(b) σ=0.1, θ=pi, 0.75pi, 0.25pi

Fig.11. Gabor filter bank parameters.

Here we use the parameters setting with: $\sigma_1=\sigma_2$, scale factor 0.2. Gabor filter bank contains 3 complex Gabor function defined as the product of 2D Gauss function modulated by 2D complex sinusoidal function with its orientation parameter θ of pi, 0.75pi, and 0.25pi. The filter template size is 31*31. The preprocessed fundus images are convolved with our Gabor filter bank to extract the superior temporal venule of retina and venula temporalis retinae inferior. The results show that these veins can be enhanced successfully. The bigger Gauss core parameter σ we choose, the smaller veins will vanished. We can also find out when the σ is smaller, two lines will appear on the result image, that makes the bigger vessel looks like 2 vessels. Fig.11(a) gives the filter template graphic symbol generated by Matlab. Two lines phenomenon is shown in Fig.12~Fig.14.

The characteristic of fundus artery is smaller and fuzzier comparing with vein. Therefore another bank of Gabor filter is designed for the extraction with the parameters of: $\sigma_1=\sigma_2$, scale 0.1; θ of pi, 0.75pi, 0.25pi; filter template size 31*31. Fig.11(b) gives the filter template graphic symbol. After the same processing to

vein, the artery can also be enhanced separately. However, in the process of extraction of retinal artery, we find that it hardly makes the venae disappeared no matter how we change the parameters of the Gabor filter bank. There are always two linear artifacts left due to the existence of venae, which makes an imperfect extraction to the retinal artery. Fig.15 shows the process of vessel extraction in ocular fundus images.

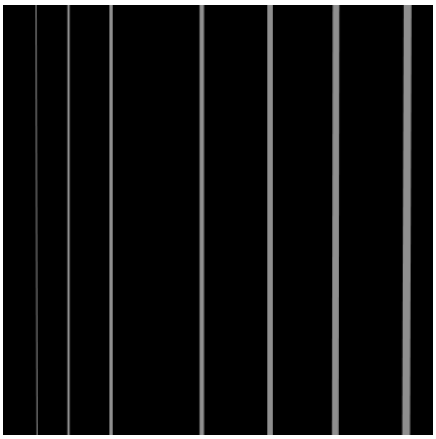


Fig.12. Original pattern.

Fig.13. The finest line is very fuzzy on the result as bigger Gauss core parameter σ is chosen.

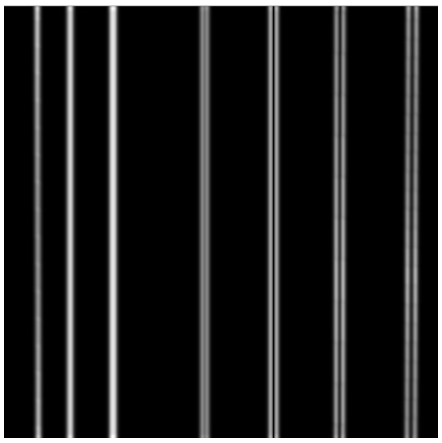


Fig.14. The result shows two linear phenomenon appears on the thicker line as smaller Gauss core parameter σ is chosen.

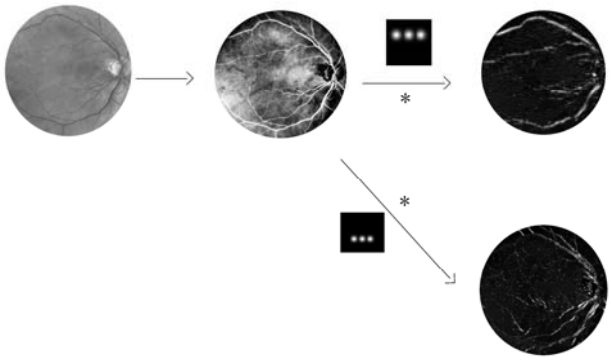


Fig.15. Process of vessel extraction in ocular fundus images.

Other 19 results are shown in Fig.16, on which the first row shows the result of fundus vein extraction, and the second row shows the result of fundus artery extraction.

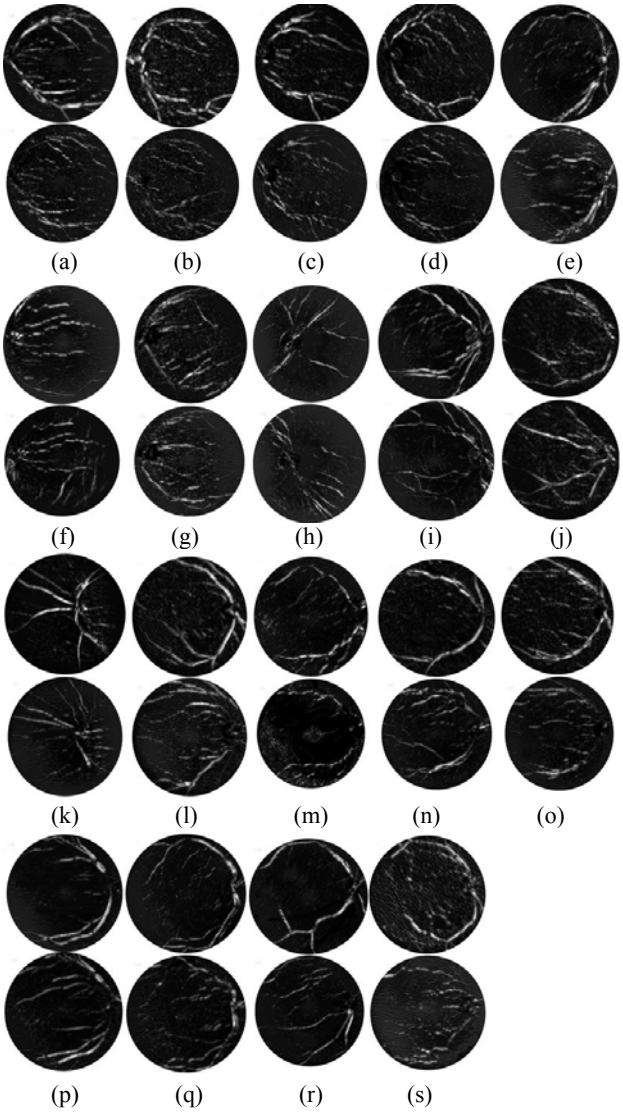


Fig.16. Extraction results of fundus vein and artery.

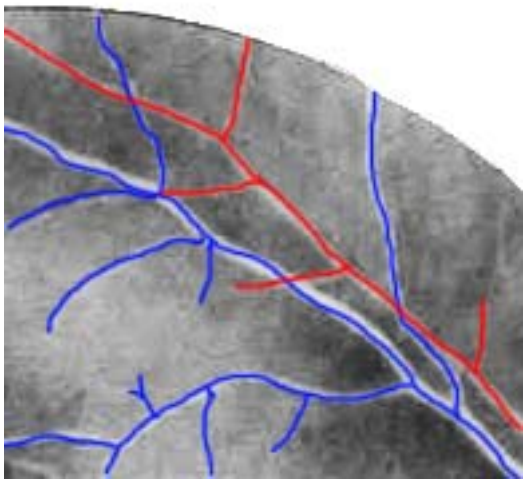
It should be noticed that there are a lot of noises in the filter result images (Fig.17(a)) which should be

removed in the next step. The binarization of result image in Fig.17(b) is labeled into each isolated part of the white area, from which the quantity of pixels is counted as the area of labels. Given each component a number, many noises could be removed if the isolated part of the area is less than 50 pixels, and the vessels can be extracted after corroding the whole image as shown in Fig.17(c).

(a) (b) (c)

Fig.17. Elimination of noises.

4 Evaluation of Performance and Discussion



(a) manually extracted vessels

(b) manually extracted vein

(c) manually extracted artery

(d) vein extracted by our method

(e) artery extracted by our method

Fig.18. Performance of vessel extraction by our method vs. gold standard.

In order to evaluate the performance of our method, manual drawing the main vessels is undertaken by a physician with the guidance of ophthalmology. We

have compared our experiment result with the manual extracted gold standard (GS) as shown in Fig.18(a)~ Fig.18(c). After Xor operation between the experiment result and GS, the accurate rate of extraction can be calculated, by using 20 cases from Guangxi medical university, the difference between the GS and our result is about 10% in vein and 18% in artery tests.

Comparing with other references, our method is the only way that can extract fundus arteriovenous separately (Table 1). For extracting temporal retinal arterioles and venules, we can roughly extract the arteries and veins of coarse contour, especially to the bigger veins, a better effect can be achieved satisfactorily. Although the thinner arteriolae temporalis retinae inferior could also be extracted, the artifact of venula always accompanies in many cases. The arteriolae could not be extracted separately in one image because of the characteristics of the filter. The extraction result is rough at current stage, we cannot accurately depict vascular edges by setting the parameters of Gabor filter. And this method ignores the vascular cross phenomenon. Table 1 also includes accurate rates and the time complexity of comparative methods.

As a matter of fact, if the vessels are extracted by a single scale of the filter, when the vessel thickness change is not big, for example, the brain vessel, our method could extract the vessel accurately. But for most of the vessel image, like our ocular fundus image, the size of vessel trunk and branch is obviously different, then use single scale filter to process the image will lead to a wrong result. If the scale is too big for the parameters of extracting the vessel, the result signal will be too diffuse to recognize. If the scale is too small, there will be no effects between the result image and the original image. Therefore the multiscale detection is needed. When the parameters' scales and directions are the most suitable for the vessel's scales and directions, the response to the filter will be the strongest. These scales correspond to the actually vessels width, the directions can be described where blood vessels go approximately, and this is why the Gabor filter bank is selected.

Table 1. Comparison of the Method Provided by the Other Document as Follows.

Methods	[4]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]	ours
MA	F	F	F	F	F	F	F	F	F	F	F	T
MB	F	F	F	F	A	F	A	F	F	F	F	F
MC	F	F	F	F	F	F	A	F	F	F	F	F
MD	F	T	T	F	A	F	A	A	T	A	A	F
ME	F	F	A	F	A	F	A	A	T	T	T	A
ACC	91%	94%	85%	*n*/a	94%	*n*/a	98%	*n*/a	*n*/a	*n*/a	*n*/a	90%
TC	$O(n^2)$	$O(n^2)$	$O(n^2)$	$O(n^2)$	$O(n^2)$	$O(n\log n)$	$O(n\log n)$	$O(n)$	$O(n^2)$	$O(n^2)$	$O(n\log n)$	$O(n^2)$

Remark: [4-16]: method provided by document in references; MA: arterial and venous segmentation separately; MB: macular; MC: optic disc; MD: vessel cross; ME: capillaries; F: False; T: True; A: Average; ACC: Accuracy; TC: Time complexity.

5 Conclusions

An automatic method of vessel extraction on ocular fundus image by using Gabor filter bank is proposed. A manual extraction is accurate but time-consuming, while our automatic method can do extraction promptly Evaluation results show an accurate rate of 90% in vein and 82% in artery tests. After preprocessing steps that include gray-scale transform, gray value inversion and contrast enhancement, the Gabor filter bank is applied to the extraction of the artery and vein in the ocular fundus images. Finally these two different width type of vessels are selected by post-processing methods such as labeling, corrosion, binarization, etc. Although the accurate rate of our method lower than some comparative methods, a Gabor filter base on Fourier transform is easy supported by hardware, processing time will fall sharply. Experimental testing has shown that our method is a robust algorithm that provides much easier way to extract artery and vein separately for the others vessel extraction processing.

References

[1] Zhou P, Wang M, Cao H. Research on Features of Retinal Images Associated with Hypertension and Diabetes [C]// *27th Annual International Conference of the Engineering in Medicine and Biology Society*, 2005. IEEE-EMBS 2005: 6415-6417.

[2] Hoyt WF, Meshel LG, Lessell S, et al. Malignant optic glioma of adulthood [J]. *Brain*, 1973, 96(1):121-132.

[3] Knox DL, Hanneken AM, Hoollows FC, et al. Uremic optic neuropathy [J]. *Arch Ophthalmol*, 1988, 106(1):

50-54.

[4] Tang Y, Zhou P. Segmentation of Retinal Blood Vessel by Merging Shape, Region and Edge Information [C]// *International Conference on Biomedical Engineering and Biotechnology*, 2012: 888-891.

[5] Nemati, RJ, Javed, MY, Fingerprint verification using filter-bank of Gabor and Log Gabor filters [C]// *International Conference on Systems, Signals and Image Processing*, 2008: 363-366.

[6] Zhao W, Jeong-Sun Park, Sang-Woong Lee. Fully automatic face detection and facial feature points extraction using local Gabor filter bank and PCA [C]// *International Conference on Machine Learning and Cybernetics*, 2011: 1789-1792.

[7] Chang C, Lin C, Pai P, et al. A Novel Retinal Blood Vessel Segmentation Method Based on Line Operator and Edge Detector [C]// *Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, 2009: 299-302.

[8] Kondo T. Gradient Orientation Based Feature Detection: An Application for Extracting Retinal Blood Vessels [C]// *Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing*, 2004: 194-197.

[9] Lin T, Zheng Y. Adaptive image enhancement for retinal blood vessel segmentation [J]. *Electronics Letters*, 2002, 38(19): 1090-1091.

[10] Fraz MM, Remagnino P, Hoppe A, et al. An Ensemble Classification-Based Approach Applied to Retinal Blood Vessel Segmentation [J]. *IEEE Transactions on Biomedical Engineering*, 2012, 59(9): 2538-2548.

[11] Sumathy B, Poornachandra S, Retinal Blood Vessel Segmentation using Morphological Structuring Element and Entropy Thresholding [C]// *Third International Conference on Computing Communication & Networking Technologies*, 2012: 1-5.

[12] Shanmugam V, Banu W. Retinal Blood Vessel Segmentation using an Extreme Learning Machine Approach [C]// *2013 IEEE Point-of-Care Healthcare Technologies (PHT)*, 2013: 318-321.

[13] Rouchdy Y, Cohen LD. Retinal blood vessel segmentation using geodesic voting methods [C]// *Biomedical Imaging (ISBI), 9th IEEE International Symposium on Biomedical Imaging*, 2012: 744-747.

[14] Yu H, Mou X, Cai Y. Adaptive Vessel Tracking Algorithm Based on Constraint of Directional Information Measure [J]. *ACTA ELECTRONICA SINICA*, 2002, 30(5): 643-646.

[15] Pan L, Wang L. Extracting Blood Vessels in Retinal Images by Adaptive Thresholding [J]. *Journal of Image and Graphics*, 2006, 11(3): 310-316.

[16] Paripurana S, Chiracharit W, Chamnongthai K et al. Retinal Blood Vessel Segmentation Based on Fractal Dimension in Spatial-Frequency Domain [C]// *Communications and Information Technologies (ISCIT)*, *2010 International Symposium*, 2010: 1185-1190.

SU Ming-jian was born in 1983, is currently a graduate in School of Computer, Electronics and Information, Guangxi University, Nanning. His current research interests include Computer Aided Diagnosis (CAD) system, image segmentation, pattern recognition, visualization in medicine. He is now following the professor Zhang Xue-jun learning computer skills. He can reached by 352912730@qq.com.

ZHANG Xue-jun received the B.S. degree in Physics from Guangxi University, China, in 1991; the M.S. degree in Electronics and Information Systems Engineering from Gifu University, Japan, in 2001; and the Ph.D. degree in Electronics and Information Systems Engineering from Gifu University, Japan, in 2004. From 2004-2007, he continued his research in Computer-Aided Diagnosis as a postdoctoral researcher at Department of Intelligent Image Information, Division of Regeneration and Advanced Medical Science, Graduate School of Medicine, Gifu University. He was invited as a visiting professor of Center for Medical Diagnostic Systems and Visualization, University of Bremen, Bremen, Germany at 2006. Since 2007, he was a Professor of School of Computer and Electronic Information, Guangxi University, Nanning, Guangxi, China. Now he is a visiting scholar in University of Southern California. His research interests include computer-aided diagnosis system, image segmentation, pattern recognition, visualization in medicine. He has published over 70 papers in Journals, Proceedings, Book chapters and Scientific Magazines.

A novel method for Measuring Mass by Image Processing

Xiaomin Tan[1,a], Jintian Lei[1,b], Hongmei Bi[1], Tianpeng Liu[1], Xuejun Zhang[1,2,c] and Ximing Wang[2]

[1]School of Computer and Electronic Information, Guangxi University, Nanning, Guangxi 530004, P. R. China

[2] IPILab, Biomedical Engineering Dept, University of Southern California, Los Angeles, CA 90033, USA.

[a]ashleytan22@hotmail.com, [b]gxuljt@163.com, [c]xjzhang@gxu.edu.cn

Keywords: Thin Plate Spline (TPS) Fast Fourier Transform (FFT) Mass Spectrum Bending Energy Computer Vision

Abstract. Image processing in terms of grid deformation was applied to measure the mass of objects. In spatial domain, Thin Plate Spline (TPS) method was employed to provide the value of minimum bending energy produced during grid deformation, and quantized mass of object; in frequency domain, Fast Fourier Transform (FFT) algorithm was to calculate the changed value of spectrum both before and after deformation in a power spectrum region. After drawing the relation curve between spatial/frequency feature and corresponding mass, mass of unknown objects can be measured on images before and after deformation. Results showed that, as the acting force that objects withstood became larger, values of bending energy increased, effects for spectrum spreading became more obvious as well. Objects having good flexibilities were compared and analyzed through an image experiment and it showed that tensile belt fitted the experiment pretty well. Different values of acting force were applied to the belt and its deformation effect was remarkable. The result was linear distribution and in accordance with the theoretical expectation.

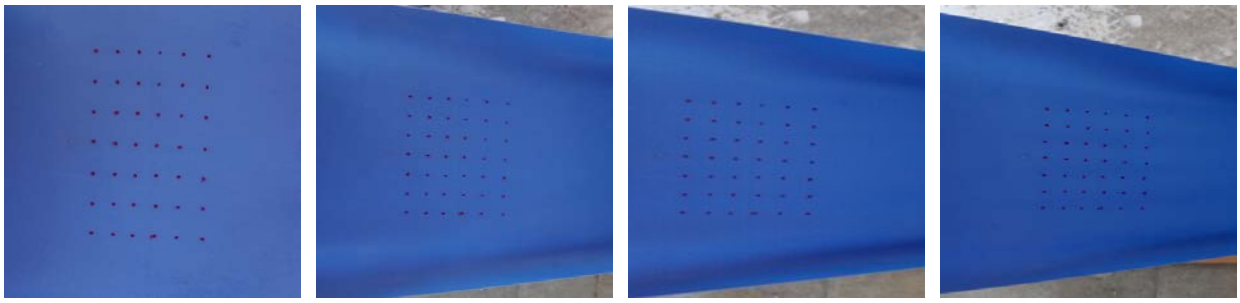
Introduction

For human being, more than 80 percents of information comes from vision and other information is obtained through tough feeling, auditory sense and gustation. More information will be produced if vision and touch feeling are combined together, but it will increase the complexity in a robot system as more sensors are required. Robot vision covers the shortage of sensors (touch feeling) thanks to the rapid development of image processing.

Measurement and analysis for object mass or its hardness using image processing are uncommon so far. Some studies were reported about the measurement of the stiffness of the internal organs such as heart [1] or liver [2] used MR tagging image. In our study, a method differs from usual hardware solution was employed to measure the mass of objects that was hanged with a soft belt having good flexibility, such as tensile belt. The measurement is undergone on a pair images with grids on belt before and after the placement of weight sets. This study is to investigate the possibility of measuring mass by quantifying the grid deformation.

Experimental material

Grid dots are drawn in the center area of tensile belt with red color as showed in Fig.1(a). Tensile belt is fixed horizontally and the dotting rectangle known as "drawing points" coincides with the region appearing deformation easily. A series of weights with different values hang under the drawing points, resulting different deformation degree of tensile belt, which showed in Fig.1(b)-Fig.1(d).



(a) (b) (c) (d)

Fig.1 Dotting rectangle(a) and the deformed grids with 550g(b),1200g(c),2000g(d).

TPS Method and bending energy (in spatial domain)

Thin Plate Spline (TPS) [3] is often used in spatial deformation of an image. It is a mapping from control points to the correspondent points in floating image. TPS is actually a interpolation with multi variables[4]. In 2-dimision spatial, Energy function is:

$$I_f = \iint_{R^2} \left(\left(\frac{\partial^2 f}{\partial x^2} \right)^2 + 2 \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 f}{\partial y^2} \right)^2 \right) dx dy. \tag{1}$$

The minimum bending energy value of I_f can be calculated by:

$$I_f = V(L_n^{-1} K L_n^{-1}) V^T. \tag{2}$$

where

$$P = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \cdots & \cdots & \cdots \\ 1 & x_n & y_n \end{bmatrix}, V = \begin{bmatrix} x_1' & x_2' & \cdots & x_n' \\ y_1' & y_2' & \cdots & y_n' \end{bmatrix}, K = \begin{bmatrix} 0 & U(r_{12}) & \cdots & U(r_{1n}) \\ U(r_{21}) & 0 & \cdots & U(r_{2n}) \\ \cdots & \cdots & \cdots & \cdots \\ U(r_{n1}) & U(r_{n2}) & \cdots & 0 \end{bmatrix} \tag{3}$$

and

$$L = \left[\begin{array}{c|c} K & P \\ \hline P^T & O \end{array} \right]. \tag{4}$$

A target landmark was chosen in the dotting rectangle and two sets of coordinate values (P and P') are obtained. P is the coordinate value without hanging weights while P' is with weights. These values are imported in the TPS program as a pair landmark to calculate minimum bending energy. Flow chart of TPS processing is shown in Figure 2.

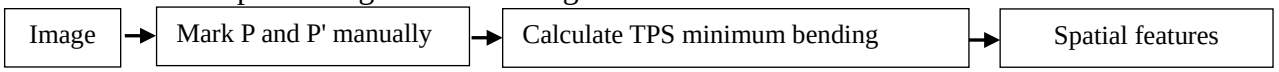


Fig.2 Flow chart of TPS method

FFT method (in frequency domain)

Fourier transform is an important tool in image processing that enables to obtain and analyze image information in frequency domain. Fourier transform is very time consuming as Discrete Fourier transform (DFT) is applied to 2-dimension digital image in frequency domain. In order to reduce the calculation time, Fast Fourier transform (FFT) [5] is proposed as a fast vision of DFT, and can be expressed by the formula:

$$F(u) = \frac{1}{2M} \sum_{x=0}^{2M-1} f(x) W_{2M}^{ux}$$

$$= \frac{1}{2}\{\frac{1}{M}\sum_{x=0}^{M-1} f(2x)W_{2M}^{u(2x)} + \frac{1}{M}\sum_{x=0}^{M-1} f(2x+1)W_{2M}^{u(2x+1)}\}. \quad (5)$$

where u=0,1,2,...,M-1.

In order to make a clean background, the red dots on the belt are firstly extracted. Then the image is compressed to 256*256 for frequency processing. Finally, FFT is applied to the compressed image to calculate the spectrum values. The flowchart of FFT processing is shown in Figure 3.

To extract the red dots from blue background, color image without deformation in Fig. 1a is transferred into gray value image by selecting the pixels with red color. Red dots are chosen on the tensile belt and only their pixel values is greater than a threshold are remained in a pre-processed image, from which FFT is used to make a power spectrum image in Figure 4. Red dots in all captured images are extracted using the same threshold.

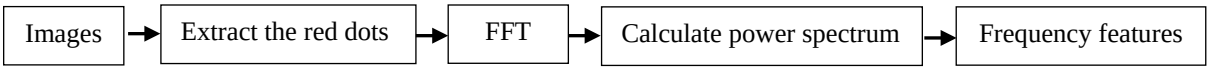


Fig.3 Flow chart of FFT method

There is only a cross in Figure 4(a) and(c), which indicate the background color and belt edge may impact the result of spectrum values. Figure 4(b) and (d) show the spectrum images become more clearly, reflecting that more information of deformation are presented in the spectrum image.

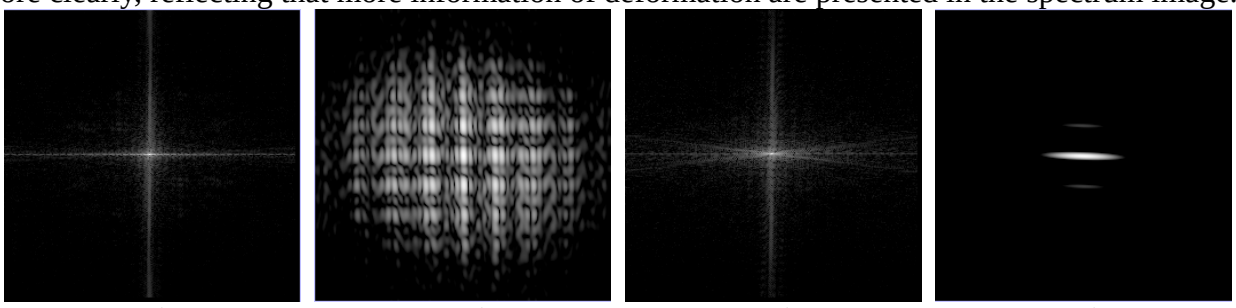


Fig.4 Spectrum images with background color and without weight (a); Spectrum images without background color and without weight (b); Spectrum images with background color and with 300g weight(c); Spectrum images without background color and with 300g weight (d).

Through the spectrum images, changes are centralized in the center. In other words, target region is centralized near the origin. However, considering the spectrum spreading, circle with the origin as center, with 3 unit pixel as radius, is chosen. All the power spectrum values add together in this circle to decrease the effect of spectrum spreading. Weights are sequenced by size and the corresponding sum of power spectrum values can be obtained. Scatter diagram are displayed with this data.

Together with feature from TPS minimum bending energy, the sum of power spectrum value is combined to be sorted and further analyzed for their correlation with mass.

Results and Discussion

Results obtained from the two methods should be verified to test its validity and feasibility. According to the linear relation from the experiments, the theory value and the actual value of mass are compared to calculate their deviation ratio, where Deviation Ratio= (theory value - actual value)/ actual value. The smaller the value of the ratio, the higher the validity and feasibility are.

TPS minimum bending energy increases as the weights on belt increases, and black trend lines in Fig.5 (both left and right) can be drawn according to their corresponding bending energy represented by red points. Experimental conditions, such as the location of capturing pictures, lead to the difference between these two sets of data. In principle, two experimental results reflect the proportional linear relation between bending energy and weight in accordance.

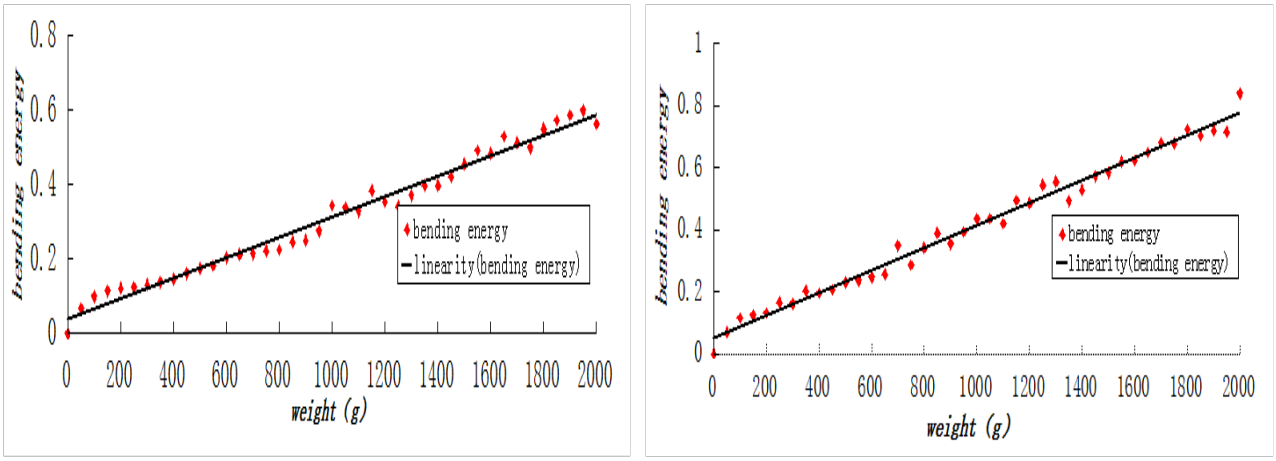


Fig.5 Experiment data distribution in spatial domain

The power spectrum value decreases as the weights on belt increases, shown with red points and black trend line in Fig.6 (both left and right). When forces being applied under the tensile belt, the distance of a grid dot changes, that corresponding to the change of single frequency of grid to multi-frequency and causing the spread of power spectrum value to distribute in a wide region. Generally speaking, according to the difference of power spectrum value before and after deformation, the mass can be found out by its relationship to power spectrum.

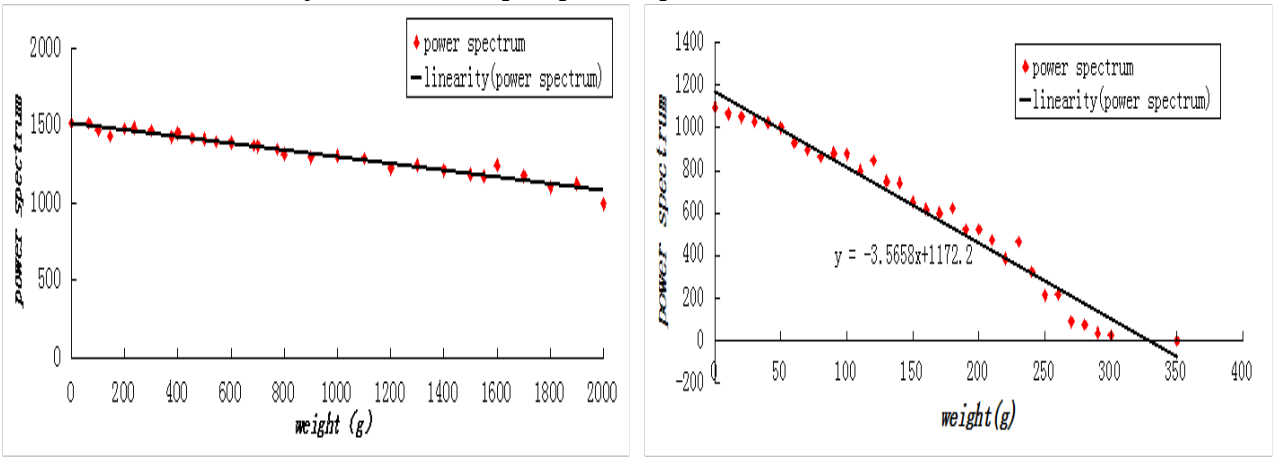


Fig.6 Experiment data distribution in frequency domain

Two sets of data in Fig.5 obtained from different mass (m_1) are combined and averaged, shown with blue points in Fig.7 (left). The black line is the trend line of data distribution. The points on the trend line can be chosen to obtain the mathematical expression: y = 0.000375x. According to its relation, data of bending energy in Fig.7 (left) is calculated to predict the mass (m_2). Then the error ratio between m_2 and m_1 can be calculated by formula (6) and yields 3.27%, that means the deviation value is small in our experiment.

Deviation Ratio (DR)=(m_1- m_2)/ m_2. (6)

Similarly, data distribution in frequency domain is shown in Fig.7 (right) with the deviation ratio 0.0121429% in average. Rectangle points in Fig.7 (right) represent the values of mass selected randomly except for the training data of mass (diamond points), which indicates the linear reciprocal relationship more generally.

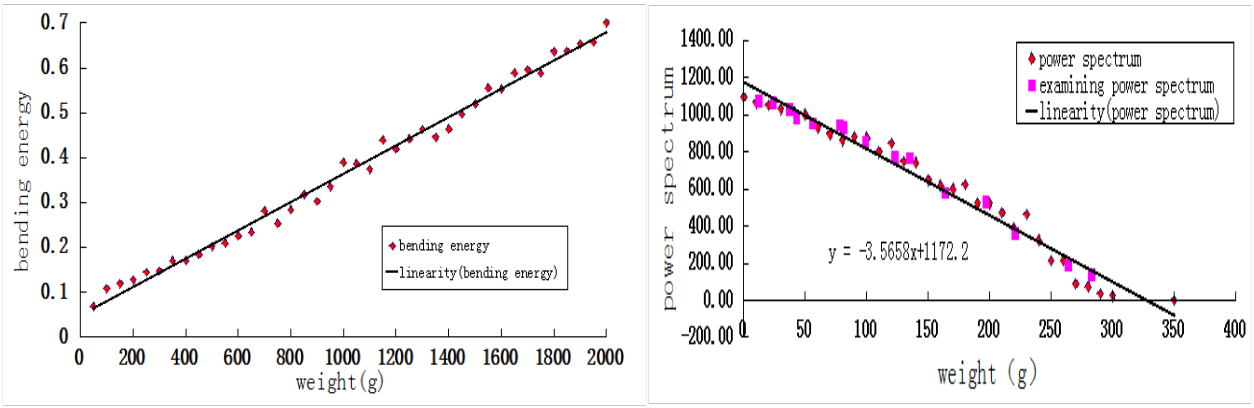


Fig.7 Experiment data distribution in spatial domain(left) and in frequency domain(right)

From the approximate curve we can notice that: In spatial domain, a linear proportional relationship between object mass and TPS minimum bending energy is shown for the material we selected, and the value of bending energy increases as the deformation becomes more remarkable. In frequency domain, it showed a linear reciprocal relationship between object mass and the sum of power spectrum value within the center of spectrum image: the power spectrum value decreases as the deformation becomes more remarkable. It is significant that TPS and FFT algorithm can measure the mass hanging on tensile belt accurately in both spatial and frequency domain. The algorithms can be easily extended to measure the mass of other objects. Furthermore, this method is now being applied to the measurement of the hardness of an object. If a fixed force pressing on two different materials, a soft one will have a greater deformation. On the contrary, smaller change of grids will be occurred on a hard object. By comparing the bending energy or power spectrum value of an unknown target with previous measured reference material, it is possible to quantify the hardness of an object we want to test. This is very useful for a Robert who with camera but no touching sensors wants to explore the hardness of an object it meeting.

Conclusion

This paper described the methods and experiments for object mass and hardness measurement based on image processing. Data with grid points is selected manually in spatial domain while automatically in frequency domain. The results show that there is a linear relationship between deformation degree and bending energy, as expected to the theory of prediction. Value of power spectrum within the center area also has linear relationship with deformation. To summary, object mass can be measured with spatial and frequency parameters in grid image sets, and it is expected that hardness can be also quantified by our proposed method.

Acknowledgment

This work was supported in part by two research support from the National Natural Science Foundation of China (No. 61262027 & 45627390), and in part by a research foundation in Guangxi University for the students' experimental skills and scientific and technological innovation ability training project. The author Xuejun Zhang gratefully acknowledges financial support from China Scholarship Council (201308455018).

References

[1] E.A. Zerhouni, D.M. Parish, W.J. Rogers, A. Yang, and E.P. Shapiro, “Human heart: Tagging with MR imaging – a method for noninvasive assessment of myocardialmotion,” Radiology, vol.169, no.1, pp.59-63, Oct. 1988.

[2] Xuejun Zhang, Satoshi Goshima, Xiangrong Zhou, Takeshi Hara, Masayuki Kanematsu, Hiroshi Fujita, Takahiro Furukawa, "Quantitative Staging the Hepatic Fibrosis: Computer-aided Diagnosis

by Shape, Texture, Volume, and Elasticity Analysis", Proc. of Radiological Society of North America Scientific Assembly and Annual Meeting Program, Radiological Society of North America (RSNA), LL-INE1211, 2011.
[3] Franke, R.. “Smooth Interpolation of Scattered Data by Local Thin Plate Splines”, Computer and Mathematics with Applications. Vol. 8, No. 4, pp.273–281,1982.
[4] L. Bookstein, “Principal warps: Thin-plate splines and the decomposition of deformations,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.11, no.6, pp.567-585, 1989.
[5] Heideman, M. T.; Johnson, D. H.; Burrus, C. S., "Gauss and the history of the fast Fourier transform". IEEE ASSP Magazine 1 (4): 14–21, 1984..

RESEARCH ARTICLE

Journal of Medical Imaging and Health Informatics
Vol. 5, 1–5, 2015

Selection of Optimal Shape Features for Staging Hepatic Fibrosis on CT Image

X. J. Zhang[1, 2, 3, *], B. Zhou[1], Kevin Ma[2], X. H. Qu[1], X. M. Tan[1], X. Gao[3], W. Yan[4], L. L. Long[4], and Hiroshi Fujita[5]

[1]*School of Computer and Electronic Information, Guangxi University, Nanning, Guangxi 530004, P. R. China*
[2]*IPILab, Biomedical Engineering Dept, University of Southern California, Los Angeles, CA 90033, USA*
[3]*Department of Medical Imaging, Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou, Jiangsu 215163, P. R. China*
[4]*Department of Radiology, First Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi 530021, P. R. China*
[5]*Department of Intelligent Image Information, Graduate School of Medicine, Gifu University, Gifu, Gifu 501-1193, Japan*

Computer-aided diagnosis of hepatic fibrosis is playing an increasing role in clinical visits, while the study based on shape features diagnosis is still in the exploratory stage. In order to improve efficiency of interpretation of medical images and diagnostic accuracy, a novel method for informatics feature selection from hepatic shape is proposed in this paper. First, the contour profile of liver is extracted from contrast-enhanced CT images. Then a difference edge curve is obtained from substation of the profile with the polynomial fitting curve, from which ten surface shape features are calculated for the input parameters of the SVM classifier. Finally, feature selection algorithm with cross-validation leave-one-out method is applied to check each combination of all features to provide the accurate rate of staging the fibrosis degree. The result shows that the optimal number of features ranges from two to six among all ten features; statistical analysis shows that maximum roughness depth (R_{max}), maximum profile valley depth (R_p), maximum profile peak height (R_m), mean spacing of the profile irregularities (S_m) and mean spacing of local peaks of the profile (S) have the greater weight than the other features. The experimental result indicates that the accuracy rate of shape feature is considerably higher than other types of features.

Keywords: Computer-Aided Diagnosis, Hepatic Fibrosis, Waviness of Surface, SVM Classification, Feature Selection.

1. INTRODUCTION

With the development of medical imaging and computer-aided technology, medical imaging technology has become one of the fastest growing fields in medicine. In liver studies, various features can be extracted from CT or MRI images to help radiologists diagnose and analyze liver lesions quantitatively. Among them, imaging features based on texture and shape have attracted more and more attention by experts and scholars. Gao et al.[1] proposed a texture image recognition method based on the correlation coefficient. The recognition rate of the diaphragmatic peritoneal ultrastructure under electron microscope texture image reached an accuracy of 92%. The multiple sclerosis (MS) research team[2] applies texture analysis technology to the field of microscopic characteristics of MS in cerebral white matter, realizing the classification prediction between the normal cerebral white matter and pathological groups. Guo et al.[3] developed a recognition method between normal livers and abnormal livers based on Gabor wavelet texture features to reach a higher recognition rate of 81.5%. Krusinska et al.[4] used shape features as a distinction analysis method to predict the histopathology of liver biopsy specimens. Bai et al.[5] used the peripheral morphological characteristics that divide edge of liver tumors into three categories, indicating that CT perfusion imaging tumor edge can reveal histopathological features and indirectly reflect the change of liver cancer angiogenesis. Computer aided diagnosis can effectively help doctors analyze the patient condition, and seeking effective diagnosis parameters may improve the performance of CAD algorithms. Elimination of redundant or irrelevant features, as well as optimization of the number of features, is a big challenge in medical image processing.

At present, there is no unified shape model and feature selection method due to the diversity of different morphological features and their combination probability. In order to find an effective and accurate method to diagnose liver fibrosis, we propose a quantitative approach for selecting effective shape features

*Author to whom correspondence should be addressed.

from the roughness parameters of micro surface that is widely used in mechanical engineering.[6]

2. MATERIALS AND METHODS

In this experiment, ten shape parameters calculated from segmented liver surface are inputted into SVM for the classification of hepatic fibrosis into two groups, from which accurate rate (AR) is derived by comparing to the gold standard with leave-one-case-out validation method. Then optimal feature numbers, optimization of feature combination and ranks of feature contribution are calculated by statistical analysis of AR results.[7]

2.1. Medical Image Materials

All of the datasets used in the experiments is provided by the radiology department of First Affiliated Hospital in Guangxi Medical University. Abdominal CT images are acquired from January 2010 to February 2011. A total of 48 cases with different liver-related diseases are collected, including cases confirmed by liver biopsy, patients with chronic hepatitis B, typical cirrhosis patients, and patients without liver biopsy and no history of liver-related diseases. Based on standard of liver fibrosis classification, the data set is divided into normal group S0, mild liver fibrosis S1 & S2, severe liver fibrosis group S3 & S4 and cirrhosis group CIR.

2.2. Liver Contour Extraction

This experiment focus on subtle changes of shape features. We use liver profile drawn manually by radiologists as the gold standard of this study. For the selection process of liver profiles, as the left liver lobe is more likely to be deformed and the formation of nodes is easier to be observed and identified compared with right lobe, we choose the partial profile in three liver areas, namely inner lower segment of left lobe of liver, inferior segment of left lateral liver and superior segment of left lateral liver, as shown in Figure 1.

2.3. Calculation of Shape Features

Shape features are intuitive and can be observed and detected easily, and thus it plays a profound role in image analysis and application. Ten most representative features[8–10] are chosen as the profile features of liver based on the roughness of micro-mechanical surface in mechanical engineering, which are

(1) roughness average (R_a);
(2) root mean square roughness (R_q);
(3) maximum roughness depth ($R_{\max}$);
(4) maximum profile peak height (R_m);
(5) maximum profile valley depth (R_p);
(6) mean spacing of the profile irregularities (S_m);
(7) mean spacing of local peaks of the profile (S);
(8) ten-point average roughness (R_z);
(9) peak density of the outline (D);
(10) profile bearing length ratio (T_p).

The measurements of the shape features are described by the following formula and demonstrated in Figures 2–4.

Roughness average R_a is a basic parameter of average profile surface roughness that is the first-order linear measurement of the degree deviating from the centerline.

$$R_a = \frac{1}{l}\int_0^l |y|\,dx = \frac{1}{n}\sum_{i=1}^{n}|y_i| \tag{1}$$

Where l stands for the sampling length of profile that is a unit measuring features; y is a mathematical function of profile; n is a set of points in uniform distribution.

The root mean square roughness R_q is the second-order linear measurement of the degree deviating from the centerline. It is based on arithmetical mean deviation, and it is a standard for the degree of profile deviating from the centerline.

$$R_q = \sqrt{\frac{1}{l}\int_0^l y^2(x)\,dx} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} y_i^2} \tag{2}$$

The values of R_a and R_q represent the average roughness of profile. Profile shows an apparent surface roughness as the values increase, and conversely, profile shows smoothness.

The maximum roughness depth $R_{\max}$ is the vertical distance between peak and valley, which describes the most prominent part of the profile surface.

$$R_{\max} = y_{p\cdot\max} + y_{v\cdot\min} \tag{3}$$

Correspondingly, maximum profile peak height R_m and maximum profile valley depth R_p is the maximum measure for deformation on the centerline in the positive direction and negative direction, respectively. R_m and R_p represent the maximum deformation in these two directions.

$$R_m = \max(y_i) \quad i = 1, 2 \ldots v \tag{4}$$

$$R_p = \max(y_i) \quad i = 1, 2 \ldots p \tag{5}$$

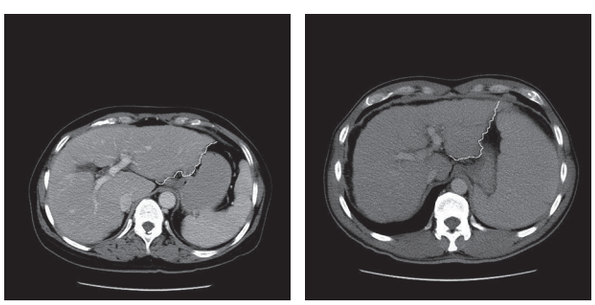


Fig. 1. Liver contour curves used in the experiment.

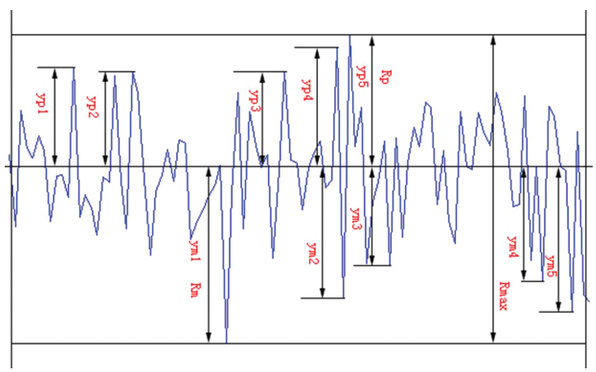


Fig. 2. Feature definition of $R_{\max}$, R_m, R_p, and R_z.

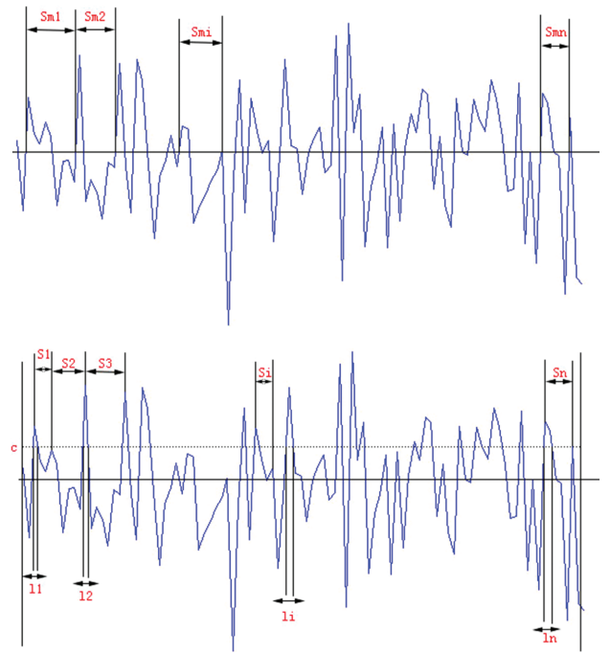


Fig. 3. Feature definition of S_m, S and T_p.

Average point height of irregularities R_z, otherwise called ten-point average height, is a comprehensive description of higher peak and lower valley on the profile.

$$R_z = \frac{1}{5}\left(\sum_{i=1}^{5} y_{p_i} + \sum_{i=1}^{5} y_{v_i}\right) \tag{6}$$

Where y_{p_i} stands for the maximum peak of the first i, and y_{v_i} stands for the minimum valley of the first i. $R_{\max}$, R_m, R_p and R_z describe profile in the vertical direction, and they are measurements of roughness in the vertical height direction.

Mean spacing of the profile irregularities S_m is an average of distances of the micro-roughness in sampling length. It was obtained by measuring the zero crossing density on the profile. The spacing of the unimodal profile represents a length projected by the distance between the peaks of two-neighbor single peak. The mean spacing of local peaks of the profile S is the arithmetic mean value of the spacing of the unimodal profile. The peak density of the profile D is a measurement of S_m in a unit length of a profile.

$$S_m = \frac{l}{n_0} = \frac{l}{\text{num}(x(y=0))-1} \tag{7}$$

$$S = \frac{1}{((p+v)/2)+1} \sum_{i=1}^{\max(p,v)} |p_i - v_i| \tag{8}$$

$$D = \frac{l}{S_m} \tag{9}$$

Where, p is the peak of profile; v is the valley of profile, n_0 is the number of zero crossing point. S_m, S and D describe profile in the transverse direction, and they are measurements of roughness in the spacing.

The profile bearing length ratio T_p is the ratio of bearing length to sampling length. Letting the cutting depth be 15% from the peak to line c, which is parallel to the mean line of profile. The fraction of the line which lies within the profile $l1, l2, \ldots, ln$ are called bearing length, which can be expressed by:

$$\begin{aligned} l_i &= x_{i+1} - x_i \\ y_i, y_{i+1} &= R_p \cdot c \end{aligned} \tag{10}$$

T_p describes the profile in both the vertical and transverse direction, reflecting the microcosmic shape characteristic.

$$T_p = \frac{\eta_p}{l} = \frac{l_1 + l_2 + \cdots l_n}{l} \tag{11}$$

2.4. Experimental Flowchart of Feature Selection

For each edge curve, ten roughness features are extracted on the basis of above formulations as the input of the SVM classifier.[7] We have chosen the radial basic core function (RBF), which is widely known and used in classifications, as the kernel function of the SVM classifier in this experiment. Feature selection includes two sections: optimizing number of input vectors and optimizing weight of each shape features. Leave-one-out cross-validation method is used to evaluate each input datasets providing classification results with accuracy rate (AR). Maximum average accuracy and the choice of corresponding features can be obtained on the basis of the AR value. The experiment flow chart is shown in Figure 4.

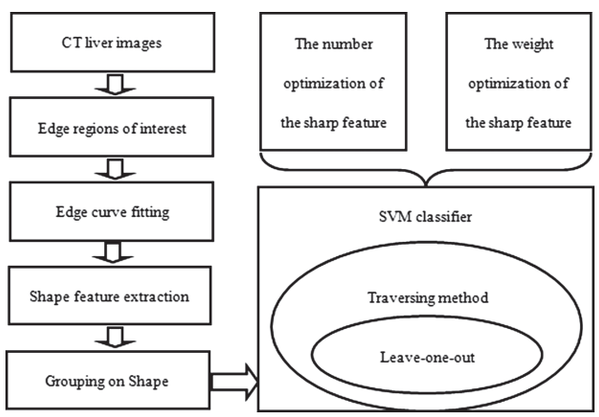


Fig. 4. Flow chart of feature selection.

3. STATISTICAL ANALYSIS OF INPUT NUMBER OF FEATURES

Based on the current stage of computer-aided diagnosis on medical images, selection of characteristic parameters is still in the exploratory stage, and there is no standard method of selecting features for reference. In order to make the statistics of maximum average accuracy corresponding to different numbers of shape feature, each number of feature sets in the classification results are grouped in $P_n(k)(k = 1, 2 \ldots 10)$, wherein n represents different input groups, k is the amount of the number of

Table I. Accuracy rate and its corresponding number of features.

Feature number	1	2	3	4	5	6	7	8	9	10
AR	0.87	0.91	0.94	0.93	0.91	0.92	0.88	0.82	0.79	0.78

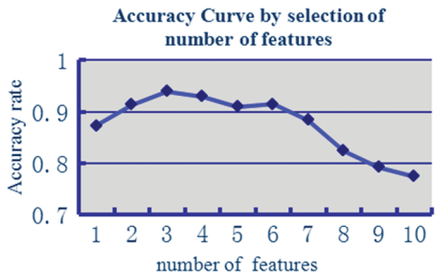


Fig. 5. Accuracy curve by the selection of number of features.

selected features. For example, $P_1(1)$ represents the highest average accuracy rate when selecting one feature as the input on a first class input grouping. It is the criteria of judging the accuracy rate of 10 types of shape in corresponding input groups. The higher the value P, the better the classification result with the currently selected feature quantity. Finally the mean value of $P_n(k)$ in each group is averaged to be the basis of shape feature quantity optimization:

$$\tilde{P}(k) = \frac{1}{7}\sum P_n(k) \quad (k = 1, 2 \ldots, 10, n = 1, 2, \ldots, 7) \qquad (12)$$

According to the results obtained with different $\tilde{P}(k)$ values in Table I, the AR rate has different performance with the number of the shape features using as an input set as shown in Figure 5.

According to Figure 5, SVM classifier obtains better classification when selecting 2–6 shape features as inputs. If the number of features is greater than 6, the classification results become worse with an increase in the input features; on the contrary, if the number of selected shape feature is too small, the insufficient information will make the classifier hard to classify the samples with less effective inputs. Therefore too many or too few features cannot effectively diagnose liver fibrosis.

4. FEATURE IMPORTANCE ANALYSIS

The weight of a feature refers to the degree of influence that a feature affects the accuracy of classification results. The greater the weight, the greater the feature influences the results of the classification experiment; on the contrary, a smaller weight means a smaller impact of this feature on the classification. According to the probability and statistics theory: A collection of events has occurred for S, $S = \{A_1, A_2, A_3 \ldots A_{n-1}, A_n\}$, the number of A_1 occurred in m replications is $mA1$, if the number of test times m is very large, the frequency $mA1/m$ stable swing near a certain value p, and with the increase of the number of tests m, its magnitude swing becomes smaller. p is called the probability of random events A_1:

$$P(A_1) = \frac{mA_1}{m} \qquad (13)$$

In order to effectively evaluate the weight of various features, we rearrange all classification results according to the classification accuracy for each combination. From all 2^{10} combination

Table II. Feature weights in the classification of hepatic fibrosis.

Shape feature	R_a	Rq	R_{max}	R_m	R_p	S_m	S	R_z	D	T_p
Feature weight	0.15	0.59	0.62	0.64	0.84	0.65	0.78	0.38	0.27	0.37
Normalized weight	0	0.64	0.69	0.71	1	0.73	0.91	0.34	0.18	0.32

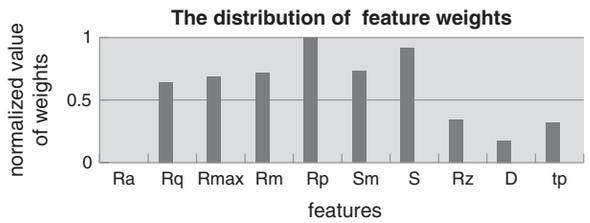


Fig. 6. The distribution of normalized value of weights for ten shape features.

datasets, the frequency of 10 features appeared in all classification combination is statistically counted as the weight of feature that can be expressed by:

$$P(k) = \frac{1}{N}\sum_{1}^{N} n \quad (k = 1, 2, \ldots, 10, n = 0, 1) \qquad (14)$$

Where k represents a number of feature, N is the number of classification test, and in our experiment $N = 1023$. n represents the appearance of k feature in each of the combinations validation: if k appears then $n = 1$, otherwise $n = 0$. After calculating the appearing probability of each feature separately, we can obtain the weight value of each feature. In order to visualize the relationships among the weights of individual features as well as the entire weight relationships in the experiment, min–max standardization[11] is applied to mapping the sample weights to a range between [0–1] using linear transformations as follows:

$$W(k) = \frac{P(k) - P(k)_{\min}}{P(k)_{\max} - P(k)_{\min}} \qquad (15)$$

Transform function normalizes the feature weights between [0, 1]. If $W(k)$ is closer to 0, that indicates a feature has a small weight and less important to the classification result, whereas $W(k)$ closer to 1 indicates a more significant feature corresponding to a higher weight. The experimental result is illustrated in Table II, and the distribution of normalized value of weights is shown in the Figure 6.

Experimental results show that the number of 3, 4, 5, 6, 7 features, namely $R_{\max}$, R_p, R_m, S_m, S, have larger weights than the other five features, meaning their larger contribution to the classification in this experiment. In other words, they play a more important role in the diagnosis of liver fibrosis with the capability of reflecting the severity of fibrosis more accurately. In mechanical engineering, the representative significance and generally accepted parameters in the measurement of microscopic surface irregularities are R_a, $R_{\max}$, R_p, R_m and T_p, that is highly identical with our results. It is also verified that the results of our classification in feature selection has a theoretical basis.

Comparing with other methods for staging the liver cirrhosis or fibrosis.[12–14] The overall performance calculated by the average sum of maximum AR value of all types number of features is 85.74% by shape features, while 66.83% by texture and 75% by volume. The result implies the efficiency of shape features to the diagnosis of liver fibrosis.

5. CONCLUSION

This study investigates the role of shape features in the classification of hepatic fibrosis by selecting the optimal parameters for

building a better Computer-aided Diagnosis system. Ten surface shape features are extracted from a standardized profile of liver. Each combination of these features is selected as input subsets to be checked by using a support vector machine (SVM) with the leave-one-case-out method to differentiate fibrosis into normal or abnormal. The result shows that the optimal number of features is ranged from 2 to 6 among all 10 features; statistic analysis shows that maximum roughness depth ($R_{\max}$), maximum profile valley depth (R_p), maximum profile peak height (R_m), mean spacing of the profile irregularities (S_m) and mean spacing of local peaks of the profile (S) have the greater weight than the other features. The experiment result indicates that the accuracy rate of using shape features for analysis is considerably higher than using other types of features.

Acknowledgment: This work was supported in part by two research support from the National Natural Science Foundation of China (No. 61262027 & 45627390). The author Xuejun Zhang gratefully acknowledges financial support from China Scholarship Council.

References and Notes

1. Y. S. Gao and J. Yang, A computer automatic identification method of medical texture image. *Journal of Zhejiang University School of Medicine* 3, 101 **(1992)**.
2. Y. Y. Zhao, L. Z. Tong, et al., Texture characteristic comparison between patients with multiple sclerosis and those with Alzheimer's disease based on MRI imaging. *Journal of Capital Medical University* 2, 199 **(2009)**.
3. Y. Guo, et al., On gabor wavelet texture feature of liver image. *Computer Applications and Software* 11, 44 **(2008)**.
4. E. Krusinska, et al., Influence of "outliers" on the association between laboratory data and histopathological findings in liver biopsy. *Methods of Information in Medicine* 32, 388 **(1993)**.
5. R. Bai, et al., A correlation of computed tomography perfusion and histopathology in tumor edges of hepatocellular carcinoma. *Hepatobiliary and Pancreatic Diseases International* 13, 612 **(2014)**.
6. X. F. Qiang, Q. G. Mao, et al., A Collection of Translated Essays in the Surface Roughness and Its Measuring, Chinese Metrology Press **(1987)**.
7. Y. F. Zeng, Computer-Aided Diagnosis of Hepatic Fibrosis on Medical Images, Guangxi University Press **(2013)**.
8. O. A. Gorlenko, Assessment of surface roughness parameters and their interdependence. *Precision Engineering* 3, 105 **(1981)**.
9. R. S. Sayles and T. R. Thomas, Micro Topometry of Engineering Surfaces, Mechanical Engineering Dept., Teesside Polytechnic **(1978)**.
10. D. J. Whitehouse, Modern methods of assessing the quality and function of surface texture, S.M.E. Technical Paper IQ72-206 **(1972)**.
11. H. G. Xiao and C. Z. Cai, Comparison study of normalization of feature vector. *Computer Engineering and Application* 45, 117 **(2009)**.
12. Y. F. Zeng, X. J. Zhang, Y. Wen, L. L. Long, Y. K. Huang, J. X. Shi, T. Liang, and Y. H. Huang, Computer aided interpretation of fibrous texture in hepatic magnetic resonance images. *Advanced Materials Research: Biomaterial and Bioengineering* 647, 325 **(2013)**.
13. H. Kato, M. Kanematsu, X. J. Zhang, et al., Computer-aided diagnosis of hepatic fibrosis: Preliminary evaluation of MRI texture analysis using the finite difference method and an artificial neural network. *American Roentgen Ray Society* 189, 117 **(2007)**.
14. X. J. Zhang, S. Goshima, X. R. Zhou, T. Hara, M. Kanematsu, H. Fujita, and T. Furukawa, Quantitative staging the hepatic fibrosis: Computer-aided diagnosis by shape, texture, volume, and elasticity analysis, *Proc. of Radiological Society of North America Scientific Assembly and Annual Meeting Program, Radiological Society of North America (RSNA), LL-INE1211* **(2011)**, p. 72.

Received: 10 April 2015. Revised/Accepted: 25 May 2015.

Abdominal Imaging

Published online: 25 August 2015

Abdom Imaging (2015) 40:3168–3174
DOI: 10.1007/s00261-015-0531-8

CT prediction of the Fuhrman grade of clear cell renal cell carcinoma (RCC): towards the development of computer-assisted diagnostic method

Hannu Huhdanpaa,[1] Darryl Hwang,[1] Steven Cen,[1] Brian Quinn,[1] Megha Nayyar,[1] Xuejun Zhang,[2] Frank Chen,[1] Bhushan Desai,[1] Gangning Liang,[3] Inderbir Gill,[3] Vinay Duddalwar[1]

[1]Department of Radiology, University of Southern California, 1500 San Pablo St, 2nd floor imaging, Los Angeles, CA 90033, USA
[2]Viterbi School of Engineering, University of Southern California, Los Angeles, USA
[3]Department of Urology, University of Southern California, 1441 Eastlake Ave, NOR 7416, Los Angeles, CA 9003, USA

Abstract

Purpose: There are distinct quantifiable features characterizing renal cell carcinomas on contrast-enhanced CT examinations, such as peak tumor enhancement, tumor heterogeneity, and percent contrast washout. While qualitative visual impressions often suffice for diagnosis, quantitative metrics if developed and validated can add to the information available from standard of care diagnostic imaging. The purpose of this study is to assess the use of quantitative enhancement metrics in predicting the Fuhrman grade of clear cell RCC.

Materials and methods: 65 multiphase CT examinations with clear cell RCCs were utilized, 44 tumors with Fuhrman grades 1 or 2 and 21 tumors with grades 3 or 4. After tumor segmentation, the following data were extracted: histogram analysis of voxel-based whole lesion attenuation in each phase, enhancement and washout using mean, median, skewness, kurtosis, standard deviation, and interquartile range.

Results: Statistically significant difference was observed in 4 measured parameters between grades 1–2 and grades 3–4: interquartile range of nephrographic attenuation values, standard deviation of absolute enhancement, as well as interquartile range and standard deviation of residual nephrographic enhancement. Interquartile range of nephrographic attenuation values was 292.86 HU for grades 1–2 and 241.19 HU for grades 3–4 (p value 0.02). Standard deviation of absolute enhancement was 41.26 HU for grades 1–2 and 34.66 HU for grades 3–4 (p value 0.03). Interquartile range was 297.12 HU for residual nephrographic enhancement for grades 1–2 and 235.57 HU for grades 3–4 (p value 0.02), and standard deviation of the same was 42.45 HU for grades 1–2 and 37.11 for grades 3–4 (p value 0.04).

Conclusion: Our results indicate that absolute enhancement is more heterogeneous for lower grade tumors and that attenuation and residual enhancement in nephrographic phase is more heterogeneous for lower grade tumors. This represents an important step in devising a predictive non-invasive model to predict the nucleolar grade.

Key words: Renal cell carcinoma—Quantitative imaging—Computer-assisted diagnosis—Computed tomography (CT)

Renal cancer accounts for more than 2% of cancers in humans worldwide [1]. In the United States, the annual incidence of renal cancer increased yearly by 1.6% over the past decade, with over 63,000 new cases in 2014 [2]. A majority of renal tumors are incidentally diagnosed on medical imaging, being often asymptomatic, small in size, and early stage [3, 4].

IRB Statement: This was institutional review board-approved, Health Insurance Portability and Accountability Act-compliant retrospective study.

Correspondence to: Hannu Huhdanpaa; *email:* huhdanpa@uw.edu; hannu.huhdanpaa@med.usc.edu

The importance of the work presented here is in the context of active surveillance (AS) [4]. If a lower grade tumor can be radiologically diagnosed with confidence, the impact on clinical management decision-making in these patients will be significant. It may substitute the need for a percutaneous renal biopsy, if considered, but will also provide more information that will be useful if active surveillance is considered as a management option [5], particularly for patients with small renal masses (<4 cm) and elderly patients or patients with significant comorbidities. As AS requires strong patient commitment, as well as imaging follow-up, having this information while management options are discussed with patients is very helpful. Additionally, it will also help in providing prognostic information and therefore help in building a personalized cancer model for the patient.

There are distinct quantifiable features characterizing renal cell carcinomas (RCC) on contrast-enhanced CT (CECT) examinations, such as peak tumor enhancement, tumor heterogeneity, and percent contrast washout [6–13]. Qualitative impressions based on visual inspections of the images are frequently sufficient for making major clinical management decisions.

Prior studies have demonstrated the degree of CT enhancement as a potentially valuable parameter in differentiating between RCC subtypes [6]. The purpose of this study is to focus on a specific subtype, clear cell RCC, and assess whether quantitative attenuation, enhancement, and washout metrics may be used to predict the Fuhrman grade of clear cell RCC.

The Fuhrman grading system is widely used for histologic grade stratification of RCC on the basis of nuclear size and shape, as well as the prominence of nucleoli [14]. The original grading system had 4 grades: grade 1 with small, round, uniform nuclei, inconspicuous nucleoli; grade 2 (40% of tumors) with slightly irregular nuclei, nuclear diameter of 15 microns and open chromatin; grade 3 (30%–40% of tumors) with very irregular nuclei, nuclear diameter of 20 microns and open chromatin; and grade 4 (15% of tumors) with mitoses, bizarre, multilobulated, pleomorphic cells, and macronucleoli, in addition to grade 3 features.

Several studies [15–20] have shown that the original moderate intra- and interobserver agreement among pathologists is improved to substantial agreement when the Fuhrman grading system is collapsed to two categories where Fuhrman grades 1 and 2 are considered together as low grade and Fuhrman grades 3 and 4 are considered together as high grade.

More recently, with an aim to improve interobserver variation, International Society of Urological Pathology (ISUP) grading system of RCCs has been developed, based on tumor cell nucleolar prominence [21–23]. ISUP also has four grades: grade 1 with nucleoli inconspicuous or absent, grade2 with nucleoli not prominent but clearly visible at high-power magnification, grade 3 with prominent nucleoli, easily visualized at low-power magnification, and grade 4 with tumor giant cells present and/or marked nuclear pleomorphism [21–23].

Materials and methods

Patients

In this institutional review board-approved, Health Insurance Portability and Accountability Act-compliant study, we retrospectively queried our IRB approved and prospectively maintained surgical database for post-nephrectomy patients who had pathology proven ccRCC and who had preoperative multiphase CECT of the abdomen between June 2009 and June 2011. A total of 65 multiphase renal CT examinations with clear cell RCCs were utilized. 48 of the patients were male, and 17 were female. The mean age of the patients was 61.5 ± 12.1 years. There were 44 tumors with Fuhrman grades 1 or 2 (considered together) as well as 21 tumors with grades 3 or 4 (considered together). There were two primary reasons for considering grades 1 and 2 together, and grades 3 and 4 together: first, as discussed above, several studies have shown significant improvement in intra- and interobserver agreement among pathologists when the grading system is thus collapsed, and secondly, the distribution of cases in our sample was such that there were very few grade 1 and grade 4 cases.

CT examination

All CT examinations were performed with a 64-detector row helical CT scanner (Brilliance, Philips Healthcare, CT). The CT scans were obtained during patient breath-holding with the following parameters: 120 kVp, variable tube current, slice thickness of 0.5 mm with reconstruction interval of 2 mm. Non-contrast, arterial, nephrographic, and excretory phase images of the abdomen were obtained. Approximately 100–150 mL of non-ionic intravenous contrast material (Isovue 350; Bracco Imaging) dosed to weight was administered with a power injector at a rate of 5 mL/sec. Time delay to scanning for arterial phase images, nephrographic phase images, and excretory phase images were 25 s, 90 s, and 5 min, respectively. Arterial phase images were obtained rather than corticomedullary phase images to create a true arterial map for surgical planning. The arterial phase used in our study is essentially similar, however, a bit earlier than a corticomedullary phase. In our routine diagnostic workflow, we did not detect a difference in diagnostic quality with studies performed with arterial phase vs. corticomedullary phase for the patients who had both as comparisons.

Data processing

Multiphase CT acquisitions were transferred to a dedicated Synapse 3D workstation (Fujifilm Medical Systems

U.S.A., Stamford, CT). Renal tumor voxels were segmented manually, slice by slice, from surrounding voxels using Synapse 3D as 3D ROIs (Fig. 1), where segmentation refers to defining the area of interest voxel by voxel so that quantitative analysis is possible. Nephrographic phase of imaging was used for segmentation as it provided the clearest tumor demarcation. In only a few cases where the nephrographic phase did not provide adequate demarcation between the tumor and the normal parenchyma, other phases as well as the embedded edge detection software available in Synapse 3D were used. Additionally, the majority of tumors were at least partially exophytic, a feature which was also used to obtain the optimal segmentation. The affected kidney and tumor were segmented out in all phases to facilitate co-registration. The entire 3D tumor volume was chosen as the ROI—as opposed to subjectively selecting smaller subareas of the tumor in a 2D plane. There were several reasons for this approach, with easier automated imaging processing workflow, accounting for tumor heterogeneity, simpler ROI definition, and improved reproducibility being the most important ones.

The DICOM-formatted CT images were converted into NIfTI (**N**euroimaging **Inf**ormatics **T**echnology **I**nitiative) volumes. NIfTI (http://nifti.nimh.nih.gov/) is a NIH-sponsored imaging initiative initially created to speed the development and enhance the utility of informatics tools related to neuroimaging. NIfTI is a file format which contains the entire imaged volume as well as a transformation which orients the voxels in three-dimensional scanner space. Given that it is just a file format representing voxels with no inherent knowledge about what the voxels contain, it is not constrained to a particular body part and can be used to convert any multi-file DICOM grayscale image into a single volumetric file.

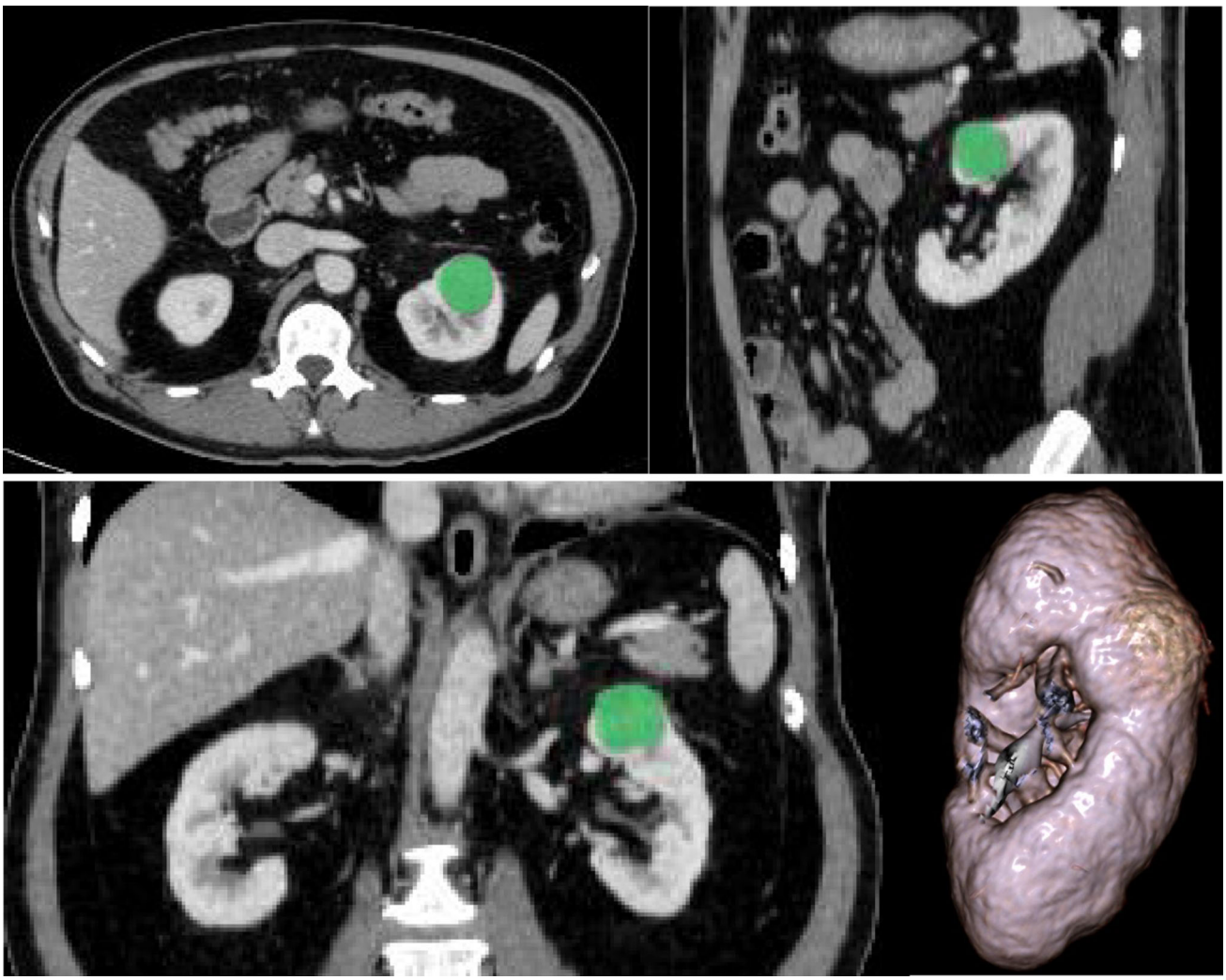


Fig. 1. The whole renal tumor is selected as a ROI. Reconstructions in axial, sagittal, and coronal planes show that the complete tumor is used for quantification. The rendered image display provides a visual reference and also references the total volume of the tumor analyzed.

DICOM keeps the images as individual slices, while NIfTI treats images as multidimensional volumes. The series of images were then co-registered using a Normalized Mutual Information (NMI) cost function implemented in Statistical Parametric Mapping SPM8 software package (The Wellcome Trust Centre for Neuroimaging at UCL, London, UK). NMI is a commonly used image similarity measure in image co-registration. A visual inspection of co-registered images was used to spot check for gross mis-registration. Custom MATLAB (MathWorks, Natick, MA) code was used to extract voxel data corresponding to the ROI.

Subsequently, histogram analysis of voxel data distribution was performed. Histogram mean, median, skewness, kurtosis, standard deviation, and interquartile range were extracted, using custom MATLAB analysis framework, for comparison between the 2 tumor grade groups for the following parameters: attenuation in each phase, absolute enhancement (arterial—pre-contrast), two washout parameters (arterial—nephrographic, arterial—excretory), and residual enhancement in nephrographic phase (nephrographic—pre-contrast) (Tables 1, 2).

While mean and median of parameter values are easily understood, the other histogram distribution parameters need further explanation. They are primarily ways to evaluate variance within a dataset. A more heterogeneous tumor would be expected to have greater variance, in other words, higher standard deviation and interquartile range, and lower kurtosis. Kurtosis is a measure of "peakedness" of histogram distribution, skewness is a measure of asymmetry of histogram distribution, interquartile range is a measure of variability which is not as sensitive to outliers as standard deviation, and lastly standard deviation is a measure of degree of dispersion of attenuation values in a dataset.

Statistical analysis

In this work, Fuhrman grades 1 and 2 were considered together as low grade and Fuhrman grades 3 and 4 were considered together as high grade.

Statistical analysis was performed using SAS 9.4 (SAS Institute Inc., Cary, NC).

Data distribution was examined using D'Agostino-Pearson test, Anderson–Darling test, and histogram-based visual inspection.

Independent t test was used for normally distributed parameters; otherwise, Wilcoxon rank sum test was used. p values less than 0.05 were considered to indicate statistical significance.

Results

Tables 1 and 2 show the results of the histogram analysis of the parameter values discussed above for the lower grade tumors vs. the higher grade tumors.

Statistically significant difference was observed in four measured imaging parameters between grades 1–2 and grades 3–4: interquartile range of nephrographic phase attenuation, standard deviation of absolute

Table 1. Histogram analysis of attenuation values by phase of imaging, comparison between low-grade tumors and high-grade tumors

Parameter	Grade 1, 2, $N = 44$	Grades 3, 4, $N = 21$	p value
Non-contrast			
Kurtosis	2.38 ± 9.35	0.47 ± 2.33	0.7
Mean	27.44 ± 11.63	25.36 ± 6.55	0.72
Median	26.07 ± 10.73	24.26 ± 6.11	0.75
Interquartile range	126.45 ± 121.67	89.48 ± 47.88	0.18
Skewness	0.75 ± 0.99	0.45 ± 0.6	0.08
Standard deviation	15.8 ± 8.38	13.4 ± 5.1	0.26
Arterial phase			
Kurtosis	2.05 ± 9.2	1.01 ± 2.09	0.26
Mean	88.23 ± 35.27	84.76 ± 35.68	0.71
Median	87.76 ± 39.23	83.76 ± 37.4	0.7
Interquartile range	283.16 ± 130.78	254.1 ± 82.81	0.28
Skewness	0.47 ± 1.08	0.42 ± 0.53	0.81
Standard deviation	41.36 ± 13.8	35.6 ± 11.68	0.1
Nephrographic phase			
Kurtosis	1.58 ± 6.88	0.6 ± 1.33	0.51
Mean	114.51 ± 39.76	105.05 ± 29.42	0.35
Median	115.01 ± 44.07	103.43 ± 31.85	0.32
Interquartile range	292.86 ± 112.77	241.19 ± 39.22	0.02*
Skewness	0.19 ± 0.97	0.25 ± 0.64	0.52
Standard deviation	42.05 ± 9.09	36.88 ± 9.04	0.08
Excretory phase			
Kurtosis	31.49 ± 45.88	24.4 ± 43.41	0.56
Mean	77.67 ± 19.02	78.08 ± 21.44	0.94
Median	74.65 ± 20.08	72.9 ± 20.74	0.75
Interquartile range	523.52 ± 500.8	553.57 ± 592.48	0.83
Skewness	2.4 ± 3.03	2.29 ± 3.26	0.73
Standard deviation	34.75 ± 24.54	40.16 ± 31.48	0.53

Table 2. Histogram analysis of enhancement and washout parameters, comparison between low-grade tumors and high-grade tumors

Parameter	Grades 1, 2, $N = 44$	Grades 3, 4, $N = 21$	p value
Arterial—non-contrast			
Kurtosis	1.84 ± 8.97	1.14 ± 2.42	0.21
Mean	64.91 ± 40.12	62.89 ± 34.62	0.84
Median	63.77 ± 43.02	60.9 ± 35.79	0.79
Interquartile range	285.68 ± 111.43	252.52 ± 91.28	0.3
Skewness	0.46 ± 1.05	0.55 ± 0.57	0.16
Standard deviation	41.26 ± 11.56	34.66 ± 11.22	0.03*
Arterial—nephrographic			
Kurtosis	2.11 ± 4.13	3.32 ± 5.47	0.13
Mean	6.45 ± 42.95	16.93 ± 32.54	0.51
Median	3.07 ± 41.02	12.93 ± 31.25	0.48
Interquartile range	251 ± 137.79	206.95 ± 117.45	0.3
Skewness	0.86 ± 0.86	1.17 ± 0.83	0.12
Standard deviation	32.38 ± 12.68	27.03 ± 12.34	0.11
Arterial—excretory			
Kurtosis	8.1 ± 17.36	8.26 ± 20.63	0.69
Mean	30.49 ± 38.96	30.96 ± 39.43	0.96
Median	29.16 ± 36.83	29.05 ± 37.03	0.99
Interquartile range	443.95 ± 513.93	373.95 ± 519.48	0.16
Skewness	0.01 ± 2.17	0.24 ± 2.24	0.23
Standard deviation	41.57 ± 25.09	35.85 ± 24.91	0.16
Nephrographic—non-contrast			
Kurtosis	1 ± 4.68	0.44 ± 1.34	0.73
Mean	89.38 ± 41.7	81.44 ± 30.61	0.53
Median	89 ± 45.77	78.95 ± 32.86	0.48
Interquartile range	297.16 ± 109.47	235.57 ± 48.17	0.02*
Skewness	0.22 ± 0.78	0.37 ± 0.6	0.19
Standard deviation	42.45 ± 9.05	37.11 ± 7.86	0.04*

* p values are statistically significant (≤ 0.05)

enhancement (arterial—pre-contrast), as well as interquartile range and standard deviation of residual nephrographic enhancement (nephrographic—pre-contrast).

Standard deviation of absolute enhancement was 41.26 HU for grades 1–2 and 34.66 HU for grades 3–4 (p value 0.03) (Fig. 2).

Interquartile range of nephrographic phase attenuation was 292.86 HU for grades 1–2 and 241.19 HU for grades 3–4 (p value 0.02) (Fig. 3).

Interquartile range was 297.12 HU for residual nephrographic enhancement for grades 1–2 and 235.57 HU for grades 3–4 (p value 0.02), and standard deviation of the same was 42.45 HU for grades 1–2 and 37.11 HU for grades 3–4 (p value 0.04).

Discussion

Determining the Fuhrman nucleolar grade of a clear cell RCC has important management implications, especially in the context of active surveillance. Determining the Fuhrman grade and its evolutionary counterpart, the ISUP nucleolar grade, on pathologic evaluation, is the most commonly used method to prognosticate renal masses. The development of a non-invasive surrogate would add an additional tool in the management of patients with RCC. However, the development and validation of these metrics to classify the nucleolar grade is complicated given the interobserver variability of pathologic evaluations. In addition, tumors may be heterogenous, and therefore, a single tumor may have loci of more than one tumor grade. Our results suggest that absolute enhancement is more heterogeneous for lower grade tumors and that residual enhancement in nephrographic phase is more heterogeneous for lower grade tumors.

Zhu et al. [24] retrospectively evaluated 255 patients with clear cell RCC and concluded that there was an inverse association between tumor enhancement and nuclear grade of RCC, with low tumor enhancement in the corticomedullary phase an independent predictor of high tumor grade. Our results demonstrated the same trend; however, it was not statistically significant. This may be partially due to sample size, but Zhu's study differs from ours in that they used subjectively selected smaller ROIs from 2D images, while in our study, the entire 3D tumor volume was defined as ROIs. As discussed above, our technique is subject to less interobserver variability and also accounts for heterogenous tumors. Other minor differences were that they used slower rate of contrast injection (3 vs. 5 mL/sec), and they used corticomedullary phase (30 s delay) vs. true arterial (25 s delay) as used in our study.

The negative correlation found between tumor grade and CT enhancement for clear cell RCC [24, 25] as well as for other tumors such as intrahepatic cholangiocarcinoma [26] has been reported to correlate with microvascular density of the tumors [24, 26]. Wang et al. established in their study of 24 cases of RCC that areas

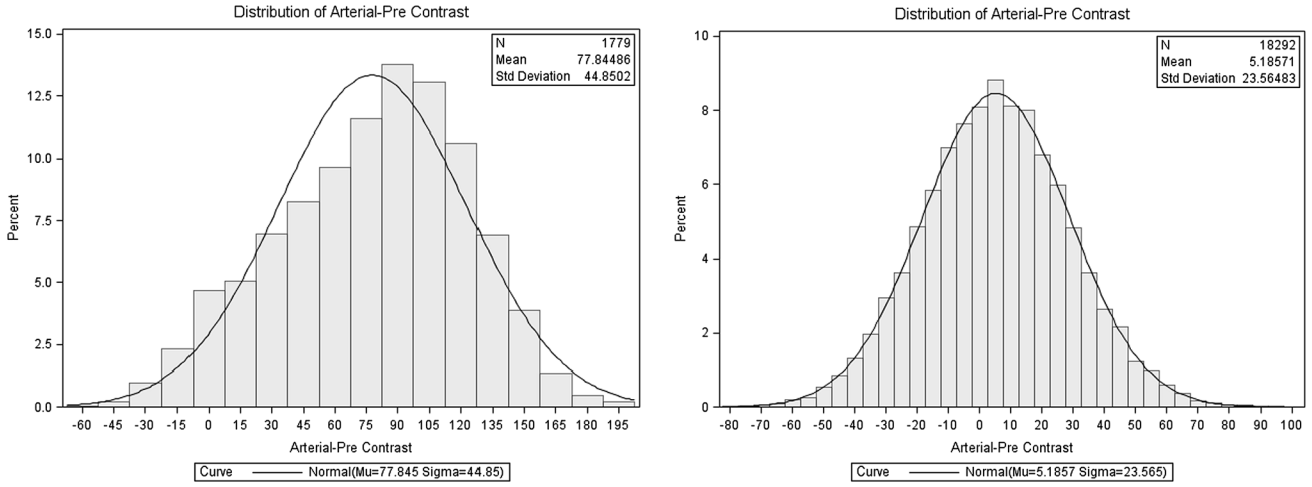


Fig. 2. Statistically significant difference was observed in standard deviation of absolute enhancement (arterial—pre-contrast). *Left graph* an example grade 2 RCC, SD 45 HU, *right graph* an example grade 3 RCC, SD 24.

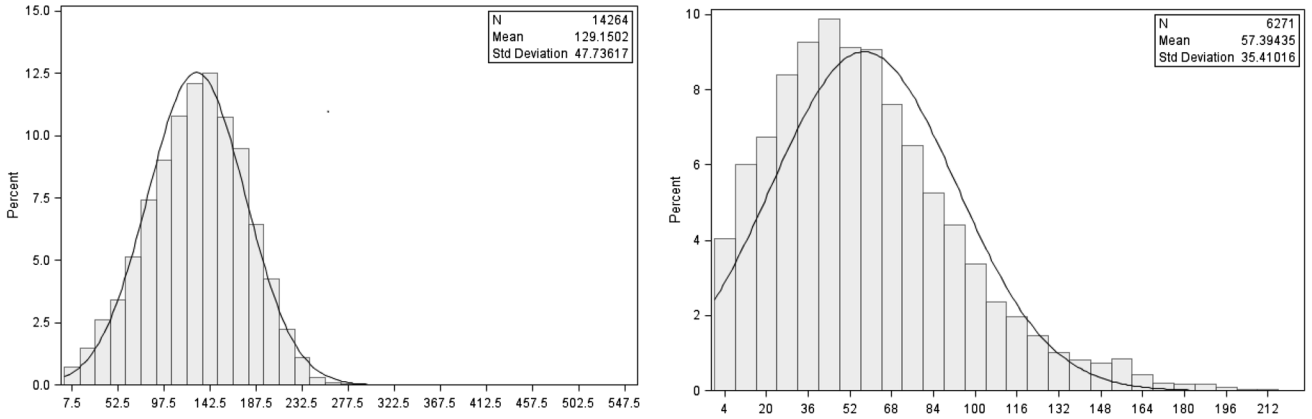


Fig. 3. Statistically significant difference was observed in interquartile range for nephrographic phase attenuation. *Left graph* an example grade 2 RCC, nephrographic phase interquartile range 549, *right graph* an example grade 3 RCC, nephrographic phase interquartile range 219.

of higher microvascular density resulted in higher CT enhancement [27]. Combining the results of Wang et al. and Zhu et al., a possible but unproven explanation is that lower grade tumors may have higher microvascular density. It is conceivable this higher microvascular density of lower grade tumors then results in more heterogeneous absolute enhancement as well as more heterogeneous residual enhancement in nephrographic phase, as suggested by our results. These possible explanations, however, need further validation. Given the importance of developing quantitative models to predict Fuhrman grade of RCC, this is an important first step in devising a predictive model. Quantifiable physical parameters including a shape and volume analysis as well as texture analysis could provide more diagnostic information.

There are several limitations to this study. This was a single-center retrospective study, and the results ideally need to be validated in a prospective multicenter trial. However, the straightforward definition of the ROI used in this study, viz. the entire 3D tumor volume, should make reproduction of the results at another center possible. Another limitation of this study was that we did not have a quantitative algorithm available to account for the degree or amount of necrosis; we are in fact evaluating such a feature currently. It is known that the amount or degree or necrosis does correlate with a higher grade of tumors and worse prognosis [28, 29]. Lastly, the sample size of our study was relatively small.

Conclusions

This study demonstrated statistically significant difference in four of the parameters characterizing contrast enhancement and washout between lower grade and higher grade tumors, viz. interquartile range of nephro-

graphic phase attenuation, standard deviation of absolute enhancement, as well as interquartile range and standard deviation of residual nephrographic phase enhancement.

Our results suggest that absolute enhancement and residual enhancement in nephrographic phase are both more heterogeneous for lower grade tumors.

Given the importance of developing quantitative model to predict the nucleolar grade, this is an important step in coming up with a non-invasive model.

Acknowledgments. This project has received funding from the Whittier foundation. The project described was supported in part by Award Number P30CA014089 from the National Cancer Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

Author contributions. Hannu Huhdanpaa—primary author of the paper, data processing and analysis; Darryl Hwang—development of Matlab algorithms for data processing; Steven (Yong) Cen—statistical analysis; Brian Quinn—segmenting and registering data; Megha Nayyar—segmenting and organizing data; Xuejun Zhang—Algorithm development and review; Frank Chen—reviewing data and image analysis; Bhushan Desai—drafting, coordination, and statistical review; Gangning Liang—pathological correlation; Inderbir Gill—clinical correlation and analysis; and Vinay Duddalwar—primary radiology faculty mentor, overall guidance, coordination, and review.

Disclosure. Hannu Huhdanpaa, Darryl Hwang, Steven (Yong) Cen, Brian Quinn, Megha Nayyar, Xuejun Zhang, Frank Chen, Bhushan Desai, Gangning Liang, Inderbir Gill, Vinay Duddalwar have nothing to disclose.

References

1. Cho E, Adami HO, Lindblad P (2011) Epidemiology of renal cell cancer. Hematol Oncol Clin N Am 25:651–665
2. Gandaglia G, Ravi P, Abdollah F, et al. (2014) Contemporary incidence and mortality rates of kidney cancer in the United States. Can Urol Assoc J 8:247–252
3. Hock LM, Lynch J, Balaji KC (2002) Increasing incidence of all stages of kidney cancer in the last 2 decades in the United States: an analysis of surveillance, epidemiology and end results program data. J Urol 167:57–60
4. Gill IS, Aron M, Gervais DA, Jewett MA (2010) Clinical practice. Small renal mass. N Engl J Med 362:624–634
5. Donat SM, Diaz M, Bishoff JT, et al. (2013) Follow-up for clinically localized renal neoplasms: AUA guideline. J Urol 190:407–416
6. Sheir KZ, El-Azab M, Mosbah A, El-Baz M, Shaaban AA (2005) Differentiation of renal cell carcinoma subtypes by multislice computerized tomography. J Urol 174:451–455 ((discussion 455))
7. Zhang J, Lefkowitz RA, Ishill NM, et al. (2007) Solid renal cortical tumors: differentiation with CT. Radiology 244:494–504
8. Herts BR, Coll DM, Novick AC, et al. (2002) Enhancement characteristics of papillary renal neoplasms revealed on triphasic helical CT of the kidneys. AJR Am J Roentgenol 178:367–372
9. Jinzaki M, Tanimoto A, Mukai M, et al. (2000) Double-phase helical CT of small renal parenchymal neoplasms: correlation with pathologic findings and tumor angiogenesis. J Comput Assist Tomogr 24:835–842
10. Kim JK, Kim TK, Ahn HJ, et al. (2002) Differentiation of subtypes of renal cell carcinoma on helical CT scans. AJR Am J Roentgenol 178:1499–1506
11. Press GA, McClennan BL, Melson GL, et al. (1984) Papillary renal cell carcinoma: CT and sonographic evaluation. AJR Am J Roentgenol 143:1005–1009
12. Ruppert-Kohlmayr AJ, Uggowitzer M, Meissnitzer T, Ruppert G (2004) Differentiation of renal clear cell carcinoma and renal papillary carcinoma using quantitative CT enhancement parameters. AJR Am J Roentgenol 183:1387–1391
13. Lee-Felker SA, Felker ER, Tan N, et al. (2014) Qualitative and quantitative MDCT features for differentiating clear cell renal cell carcinoma from other solid renal cortical masses. AJR Am J Roentgenol 203:W516–W524
14. Fuhrman SA, Lasky LC, Limas C (1982) Prognostic significance of morphologic parameters in renal cell carcinoma. Am J Surg Pathol 6:655–663
15. Bektas S, Bahadir B, Kandemir NO, et al. (2009) Intraobserver and interobserver variability of Fuhrman and modified Fuhrman grading systems for conventional renal cell carcinoma. Kaohsiung J Med Sci 25:596–600
16. Al-Aynati M, Chen V, Salama S, et al. (2003) Interobserver and intraobserver variability using the Fuhrman grading system for renal cell carcinoma. Arch Pathol Lab Med 127:593–596
17. Ficarra V, Martignoni G, Maffei N, et al. (2005) Original and reviewed nuclear grading according to the Fuhrman system: a multivariate analysis of 388 patients with conventional renal cell carcinoma. Cancer 103:68–75
18. Lang H, Lindner V, de Fromont M, et al. (2005) Multicenter determination of optimal interobserver agreement using the Fuhrman grading system for renal cell carcinoma: assessment of 241 patients with >15-year follow-up. Cancer 103:625–629
19. Sun M, Lughezzani G, Jeldres C, et al. (2009) A proposal for reclassification of the Fuhrman grading system in patients with clear cell renal cell carcinoma. Eur Urol 56:775–781
20. Qayyum T, McArdle P, Orange C, et al. (2013) Reclassification of the Fuhrman grading system in renal cell carcinoma-does it make a difference? Springerplus 2:378
21. Delahunt B, McKenney JK, Lohse CM, et al. (2013) A novel grading system for clear cell renal cell carcinoma incorporating tumor necrosis. Am J Surg Pathol 37:311–322
22. Delahunt B, Egevad L, Montironi R, Srigley JR (2013) International Society of Urological Pathology (ISUP) consensus conference on renal neoplasia: rationale and organization. Am J Surg Pathol 37:1463–1468
23. Delahunt B, Cheville JC, Martignoni G, et al. (2013) The International Society of Urological Pathology (ISUP) grading system for renal cell carcinoma and other prognostic parameters. Am J Surg Pathol 37:1490–1504
24. Zhu YH, Wang X, Zhang J, et al. (2014) Low enhancement on multiphase contrast-enhanced CT images: an independent predictor of the presence of high tumor grade of clear cell renal cell carcinoma. AJR Am J Roentgenol 203:W295–W300
25. Villalobos-Gollás M, Aguilar-Davidov B, Culebro-García C, et al. (2012) Pathological implications of areas of lower enhancement on contrast-enhanced computed tomography in renal-cell carcinoma: additional information for selecting candidates for surveillance protocols. Int Urol Nephrol 44:1369–1374
26. Nanashima A, Abo T, Murakami G, et al. (2013) Intrahepatic cholangiocarcinoma: relationship between tumor imaging enhancement by measuring attenuation and clinicopathologic characteristics. Abdom Imaging 38:785–792
27. Wang JH, Min PQ, Wang PJ, et al. (2006) Dynamic CT evaluation of tumor vascularity in renal cell carcinoma. AJR Am J Roentgenol 186:1423–1430
28. Sengupta S, Lohse CM, Leibovich BC, et al. (2005) Histologic coagulative tumor necrosis as a prognostic indicator of renal cell carcinoma aggressiveness. Cancer 104:511–520
29. Knight DA, Stadler WM (2007) Prognostic factors in localized renal cell cancer. BJU Int 99:1212–1216

SPIE 2016 PRE-PRINTS

Multi-Site Evaluation of a Clinical Decision Support System for Radiation Therapy

Ruchi Deshpande[a], John DeMarco[b], Kerstin Kessel[c], Brent J. Liu[a]
[a]Image Processing and Informatics Laboratory, Department of Biomedical Engineering, University of Southern California, 1042 Downey Way, DRB 265, Los Angeles, CA, USA 90089;
[b]Department of Radiation Oncology, University of California Los Angeles, 200 Medical Plaza Drive, Los Angeles, CA
[c]Department for Radiation Oncology, Technische Universität München (TUM), Klinikum rechts der Isar, Ismaninger Str. 22 D-81675 München

ABSTRACT

We have developed an imaging informatics based decision support system that learns from retrospective treatment plans to provide recommendations for healthy tissue sparing to prospective incoming patients. This system incorporates a model of best practices from previous cases, specific to tumor anatomy. Ultimately, our hope is to improve clinical workflow efficiency, patient outcomes and to increase clinician confidence in decision-making. The success of such a system depends greatly on the training dataset, which in this case, is the knowledge base that the data-mining algorithm employs. The size and heterogeneity of the database is essential for good performance. Since most institutions employ standard protocols and practices for treatment planning, the diversity of this database can be greatly increased by including data from different institutions. This work presents the results of incorporating cross-country, multi-institutional data into our decision support system for evaluation and testing.

Keywords: Decision Support, Radiation Therapy, Radiation Oncology, IMRT, Imaging Informatics, Head and Neck Cancer, DICOM RT

1. INTRODUCTION

The biggest challenge in utilizing Radiation Therapy to treat cancer is targeting the tumor with maximum possible radiation dose, while sparing the surrounding normal Organs At Risk (OARs) as much as possible. This is especially difficult to accomplish when the OARs are small and situated very close to the tumor, as in the case of head and neck cancer. Since it is often not possible to spare the OARs in entirety, it is advisable to limit their dose exposure as much as possible. However, the lowest possible, yet practically achievable dose to vital OARs cannot be computed quantitatively, and so clinicians must rely on experience, evidence-based guidelines and trial-and-error methods to arrive at close approximations during treatment planning. This is where computational data mining techniques can help by determining best practices from previous patients with similar anatomical tumor-OAR configurations, to use as templates for the current patient. Our ultimate objective is to build a decision support system that assists clinicians in identifying good OAR dose end-points for patients, and further, in determining treatment-planning parameters that lead to these optimal dose end-points. In order to accomplish this, we must also build a number of essential system components to support the decision making engine by facilitating data collection and management. The following sections describe our workflow analysis and data model, system components, strategies and techniques to facilitate multi-institutional data sharing, and results of evaluating the decision support algorithm with multi-institutional data.

1.1 Workflow

Before designing a clinical decision support system, or any other medical imaging informatics system, it is essential to analyze the clinical workflow. The next step is to determine where the new system fits into the clinical workflow, and to analyze the effect it will have on normal clinical operations. Figure 1 shows the clinical workflow in Radiation Therapy Treatment Planning [1]. It assumes that the patient has already been enrolled for Radiation Therapy, and outlines the steps that follow. These steps are summarized below:

- CT Simulation and Portal Image: Acquisition of the CT images that are used for treatment planning. These images provide a three-dimensional model of the patient's anatomy that helps clinicians locate tumors with precision.
- ROI Contouring: All relevant Regions Of Interest are contoured slice-by-slice on the CT images, with specially designed software from the Treatment Planning System (TPS).
- Initial Parameters: Selection and placement of fields, number and direction of beams, etc., which form the initial parameters for forward treatment planning. In inverse treatment planning for Intensity Modulated Radiation Therapy, clinicians set dose constraints for ROIs as objectives for the plan optimization algorithm.
- Plan Optimization: The Treatment Planning System calculates radiation beam settings (Multi-Leaf Collimator settings, etc.) that satisfy the previously specified dose constraints, and generates the resulting dose distribution.
- Plan Evaluation: Review of Dose Volume Histograms, Isodose contours, etc. in order to ensure plan quality.
- Plan Approval: The radiation oncologist either approves or rejects the plan based on the evaluation results
- Re-adjustments and fine-tuning: Further adjustments are made to the dose constraints or other initial parameters to resolve inadequacies (if any) found in the evaluation. Figure 1. [1, 2]

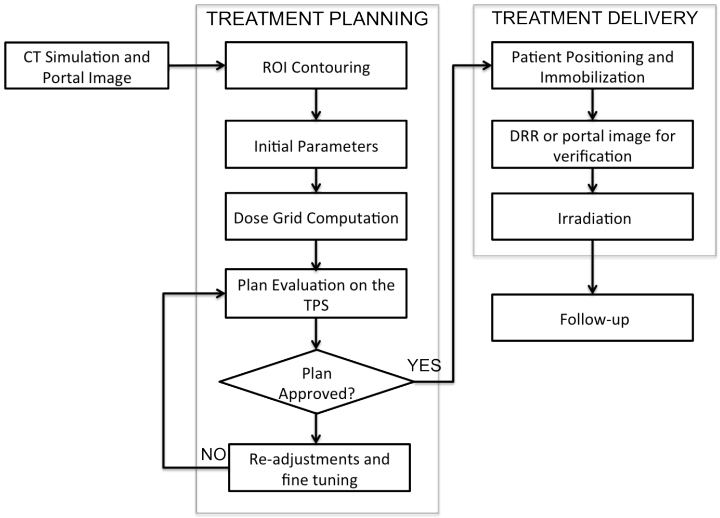


Figure 1. Clinical Workflow in Radiation Therapy Treatment Planning

The decision support system benefits the clinical workflow by potentially reducing the number of iterations of the treatment evaluation loop. This helps to increase workflow efficiency, and to facilitate development of treatment plans that are practical as well as more optimal, which in turn may improve patient outcomes.

1.2 The Data Model

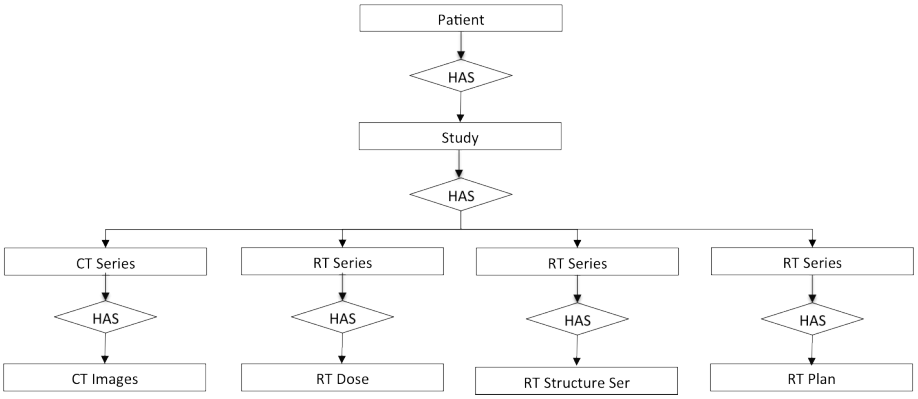


Figure 2. DICOM RT objects in the DICOM data model: Dose, Structure Set and Plan

We have chosen to use the DICOM [3] standard to ensure vendor-neutrality of our system. Good organization and optimal data accessibility are essential in making the machine learning and data analytics tools feasible. An understanding of the structure and the relationship between the principal data elements facilitates proper implementation. The four types of data objects that the system deals with are: DICOM RT Structure Set, DICOM RT Dose, DICOM RT Plan, DICOM CT Images. Figure 2 shows where each of these objects fit into the data model. The Structure Set object defines the various structures of relevance, or 'Regions Of Interest' (ROIs) such as the radiation target (the tumor) as well as surrounding Organs at Risk (OARs). It also provides the coordinates of the contours that outline these ROIs. The Dose object contains a three-dimensional dose grid, as well as Dose Volume Histogram sequences for all the ROIs defined in the Structure Set object. The Plan object contains technical parameters and details regarding the treatment beams and fields such as shape, number, energy, etc. The CT images are the CT simulation images that are used specifically for treatment planning. The RT objects all fall under a special type of series – the RT series.

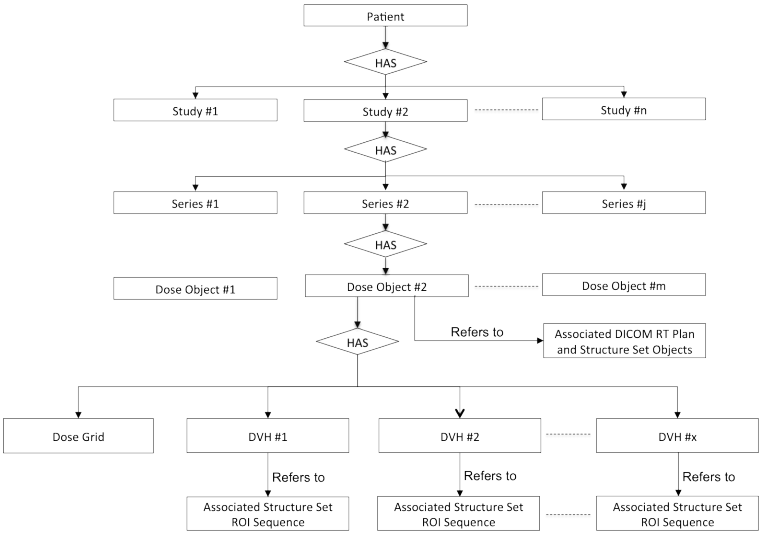


Figure 3. Data Model of the DICOM RT Structure Set Object

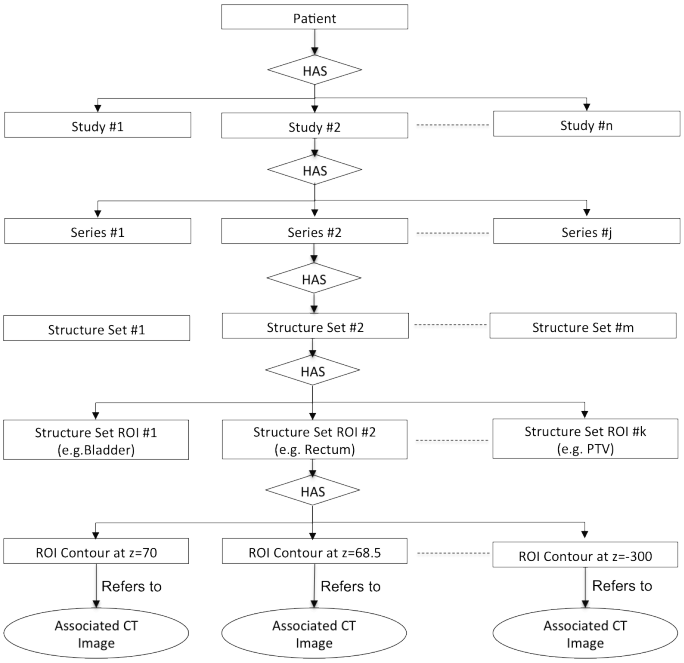


Figure 4. Data Model of the DICOM RT Dose Object

2. METHODS

The following sections describe the system components in brief.

2.1 An Overview of System Architecture

Figure 5 shows the system components and inter-connections between them. The main components are described below:

1) The DICOM Parser: The parser is responsible for reception and management of incoming DICOM data. It can operate in one of two modes. The first is server-side parsing, wherein the files are uploaded after anonymization, and then parsed on the server by a python script. The second is client side parsing, where the DICOM files do not leave the client's machine at all, and a JavaScript parser accesses these files through a browser and extracts only certain pre-defined attributes that are known to be anonymous. The parser extracts and then catalogs all the DICOM metadata that it receives into the database.

2) The Database: The database records all relevant metadata associated with various DICOM objects, as well as other data that is derived from the raw DICOM data after processing. It contains the anatomical descriptors that quantify the relationship between the Planning Target Volume and surrounding Organs At Risk [4, 5]. The decision support system uses MySQL, which is an open-source relational database management system.

3) The Feature Extraction Module: This piece is responsible for carrying out computational image processing on the DICOM datasets in order to extract features that can be used by the decision support algorithm. Currently, this is being implemented in MATLAB. A direct connection to MySQL ensures that the results of feature extraction are recorded in the non-DICOM parts of the database in an organized fashion. This module extracts anatomical features that quantify properties specific to the PTV with respect to surrounding Organs At Risk. Examples of such features are distance, overlap, shape, orientation and location. These features have been integrated into two comprehensive shape descriptors – the Overlap Volume Histogram (OVH) and the Spatial Target Signature (STS).

4) The Machine Learning module: This module utilizes the features derived by the feature extractor in order to carry out similarity matching between new patients and database patients, based on the geometric relationship of the PTV with respect to surrounding OARs. This was implemented in MATLAB, and we are in the process of porting it to Python scripts that can be integrated directly with the system. The algorithms are performed on the server-side while the results are displayed through the web-based GUI. This module uses the features extracted by the system to calculate anatomical similarities across cases with respect to PTV-OAR geometry. The dose distributions of similar database cases are then used to guide the treatment planning process for a new patient.

5) The Web-based Zero-Footprint Graphical User Interface: The GUI is responsible for presentation and visualization of the results of system analysis, as well as clinical data objects such as CT images, ROI overlays, isodose curves, etc. The GUI is written using HTML and JavaScript.

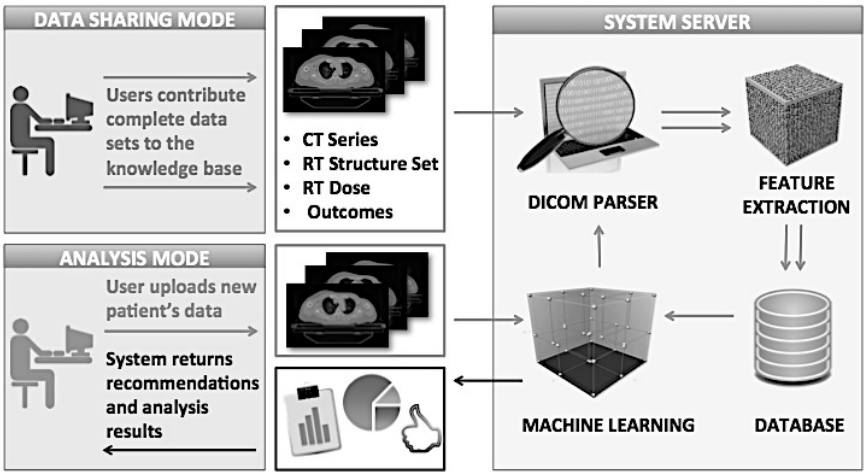


Figure 5. System Component and Architecture diagram

The system may be used in one of two modes as shown in Figure 5. In the data-sharing mode, interested collaborators may contribute treatment planning data while the analysis mode is used for clinical decision support.

2.2 Client-Side Parser

In order to make the system a collaborative initiative that encourages participants, we have enabled the system to work in a local client-side parsing mode. This protocol ensures that the user's DICOM files do not the client machine. Only fields that have been pre-selected and verified for anonymity are extracted from the DICOM files that the user selects. This extracted metadata is not identifiable, and thus fulfills HIPAA requirements.

This client-side parser is written in JavaScript. The functions of this toolkit run purely on the client's machine, facilitated by a web browser. The user interface prompts users to select a folder from their local file system. All the DICOM files associated with the treatment data that the user wishes to contribute must be available in this folder. The parser will then associate each data element with an attribute. Once this mapping is in place, another component of the toolkit will extract data elements for a list of pre-determined attributes, send them to the server.

2.3 The Decision Support Algorithm

The decision support mechanism is based on the principle of identifying anatomically similar database patients to use as reference cases for incoming patients. Anatomical similarity, in this case, refers to the spatial relationship between the tumor and surrounding organs. We have picked specific features that quantify these spatial relationships. By assessing the differences in feature values across various patients, we can derive a similarity ranking for a set of database patients with respect to a new, incoming patient. Some of these features include, but are not limited to – comprehensive distances between OAR voxels and the PTV, overlap between the OARs and the PTV, the directional orientation of the PTV with respect to the OARs, the size of the PTV, etc. These features are calculated for each OAR-ROI pair separately and then combined in a weighted average. After assigning similarity scores to all database patients, we set a similarity threshold to pick a subset of the most similar patients in order to derive practically achievable IMRT dose constraints from them.

3. RESULTS

3.1 Data Sharing Workflow

The implementation of this new client-side DICOM parsing workflow is still in progress. However, we have developed a data sharing workflow that demonstrates how this new protocol can be put to practical use.

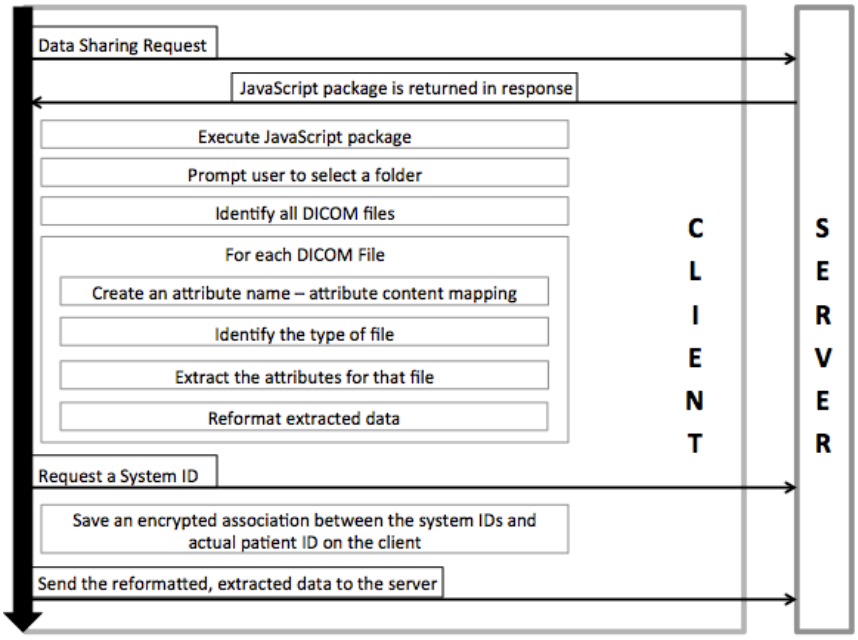


Figure 6. Data Sharing Workflow Profile

Currently, the client-side DICOM parser is being developed and tested. It has been shown to preserve the integrity of the data being extracted. It was tested with a dataset of 10 treatment plans associated with 10 different patients. Each dataset included CT images, DICOM RT Structure Set and DICOM RT Dose. A subject matter expert in Radiation Therapy performed verification of the results. DICOM RT Plan is yet to be tested. Figure 6 shows the workflow of this new data sharing protocol for research applications.

(1) The user must first initiate a data sharing protocol with the server. This is a notification to the server that a client wishes to upload data.
(2) The server then returns an acknowledgement and transfers resources back to the client in the form of a JavaScript package that is executed by the client's web browser.
(3) The user is first prompted to select a folder containing all the DICOM files to be processed.
(4) The parser identifies all the DICOM and non-DICOM files stored within sub-folders in the main selected folder. For every DICOM file encountered, the file type is first identified.
(5) Based on the type of file, the parser selects a set of pre-determined attributes to extract since a different set of attributes is associated with each file type.
(6) The values or content for those attributes are extracted and reformatted.
(7) Next, the JavaScript tool sends a request to the server for a new system ID to associate the data with.
(8) This system ID is then stored in a mapping with the true patient ID of the dataset. This mapping is stored in an encrypted format on the client's machine, and can be loaded the next time the user wants to upload follow-up information for the same patient.
(9) The de-identified, reformatted data is then transferred to the server for storage.

3.2 Data Integration

We have integrated treatment planning data from two different institutions. The first dataset was obtained from the University of California Los Angeles (UCLA), and consisted of 137 anonymized treatment planning cases. The second dataset was obtained from the Technische Universität München (TUM), and consisted of 52 anonymized treatment planning cases. The datasets were exported from three different treatment planning systems belonging to two different vendors – RapidArc and TrueBeam from Varian and TomoTherapy from Accuray. Our data gateway successfully parsed DICOM files collected from two different countries and institutions, as well as different treatment planning systems. Our data processing functions were also successfully being applied to both datasets. Since different institutions follow different protocols and naming conventions, the ROI names and definitions had to be manually reconciled and standardized across all datasets.

The decision support system was evaluated to test the algorithm's ability to retrieve retrospective database cases that are most likely to share similar dose profiles with a new prospective case. The datasets from each individual institution were evaluated separately and then together. Each evaluation produced a success rate of approximately 77% was for the parotid glands. This indicates that the decision support algorithms are able to successfully retrieve database cases that are relevant to a new case 77 out of 100 times. The performance of the system can be improved with more data and fine-tuning of the algorithms.

3.3 RT Data Sharing Research Initiatives

There have been some recent efforts at building solutions to support large-scale collection of radiation therapy data for research purposes. Two examples that stand out are the 'Transfer of Images and Data' (TRIAD) platform [6] and the 'National Radiation Oncology Registry' (NROR) [7]. TRIAD is a medical data sharing platform sponsored by the American College of Radiology (ACR) that enables exchange of DICOM and non-DICOM data from various participating sites for the purposes of clinical trials, accreditations, etc. The Radiation Therapy Oncology Group (RTOG) has chosen to use TRIAD in collecting data for its clinical trials. However, since this platform is intended to facilitate the flow of data involved in sharing and collecting files, it does not address the need to catalogue this data into databases for easy query and retrieval. This is one of the key components of the system presented here. Additionally, it requires installation of de- identification software, which further entails platform dependence and regular installation of software updates. The NROR aims to build a large-scale national registry of treatment delivery and health outcomes data to fuel a broad range of research and quality assurance activities. It functions on the basis of a Business Associate agreement between NROR and participating sites, and as such, it does not require a de-identification module since it is considered a "covered entity" by the HIPAA Privacy Act. As a result, only tried and tested decision support algorithms can be

integrated with it. However, research and validation of new decision support algorithms cannot be conducted with NROR data. This decision support system, in contrast, provides a small but robust platform for developing and testing new decision support techniques. Moreover, the modular nature of this system allows future integration with systems such as NROR. For instance, the decision support component or the client-side metadata extraction tool may be integrated into the NROR, or data from other systems may be integrated into the current decision support system.

3.4 System Evaluation

The decision support system showed a 90% success rate in detecting anatomical PTV dissimilarities. This evaluation was conducted by deliberately introducing PTV dissimilarities to retrospective data using image warping techniques. Further, the system also showed a success rate of 75%, 80%, 74% and 58% in retrieving relevant database cases for the purpose of decision support with respect to the mandible, left parotid gland, right parotid gland and brainstem respectively.

4. DISCUSSION

In summary, we have developed a system that mobilizes data from extensive treatment planning databases, cancer registries, etc.; derives knowledge from this data; and uses it to help clinicians make better informed decisions. In order to ensure compatibility and vendor neutrality, we have chosen to work with DICOM data, especially since most TPS vendors provide the ability to export planning data to DICOM RT. In addition, DICOM provides an organized and effective data model that we have used to build our database schema. We have also developed a data sharing protocol that may encourage more clinicians to contribute data, and may even be incorporated into future registries and research databases. The decision support algorithm itself is under continual testing and improvement. We have conducted informal and subjective analyses of the system's performance, and have found it's potential clinical impact to be very promising. So far, it has succeeded in identifying a number of cases where using the system could yield inputs that might lead to better dose end-points. This stage of our evaluation is currently being planned and designed.

5. CONCLUSION

We have presented the concept of a decision support system for treatment planning of head and neck cancer cases in radiation therapy. We have also outlined the system architecture, described its various components and how they fit together. The outline of our system evaluation was discussed, and preliminary results were touched upon.

REFERENCES

[1] Law M. Y. Y., Liu B. J., "Informatics in Radiology – DICOM RT and Its Utilization in Radiation Therapy", RadioGraphics 2009; 29:655-667

[2] Deshpande R., Zhou, A., Zhang, J., et al., "Role of an imaging informatics-based DICOM-RT cancer registry in evaluating treatment parameters of IMRT for prostate cancer", Proceedings of SPIE Vol. 8674, 86740O (2013)

[3] Digital Imaging and Communications in Medicine (DICOM) Part 3. Information object definitions. PS3.3-2009

[4] Kazhdan M., Simari P., McNutt T., Wu B., et al., "A Shape Relationship Descriptor for Radiation Therapy Planning", MICCAI 2009; 5762:100-108

[5] Wu B., Ricchetti F., Sanguineti G., Kazhdan M., et al., "Patient geometry-driven information retrieval for IMRT treatment plan quality control", Med Phys. 2009;36(12):5497-505

[6] Radiation Therapy Oncology Group Core Lab TRIAD. http://www.rtog.org/CoreLab/TRIAD.aspx

[7] Efstathiou JA, Nassif DS, McNutt TR, et al (2013) Practice-based evidence to evidence-based practice: building the National Radiation Oncology Registry. J Oncol Pract 9(3): e90-5.

Lesion Registration for longitudinal disease tracking in an imaging informatics-based Multiple Sclerosis eFolder

Kevin Ma[1], Joseph Liu[1], Xuejun Zhang[1], Alex Lerner[2], Mark Shiroishi[2], Lilyana Amezcua[3], Brent Liu[1]
[1]Image Processing and Informatics Laboratory, Dept of Biomedical Engineering, Univ. of Southern California, 1042 Downey Way. Los Angeles, CA 90089
[2]Dept of Radiology, Univ. of Southern California, Los Angeles, CA 90089
[3]Dept of Neurology, Univ. of Southern California, Los Angeles, CA 90089

ABSTRACT

We have designed and developed a multiple sclerosis eFolder system for patient data storage, image viewing, and automatic lesion quantification results stored in DICOM-SR format. The web-based system aims to be integrated in DICOM-compliant clinical and research environments to aid clinicians in patient treatments and data analysis. The system needs to quantify lesion volumes, identify and register lesion locations to track shifts in volume and quantity of lesions in a longitudinal study. In order to perform lesion registration, we have developed a brain warping and normalizing methodology using Statistical Parametric Mapping (SPM) MATLAB toolkit for brain MRI. Patients' brain MR images are processed via SPM's normalization processes, and the brain images are analyzed and warped according to the tissue probability map. Lesion identification and contouring are completed by neuroradiologists, and lesion volume quantification is completed by the eFolder's CAD program. Lesion comparison results in longitudinal studies show key growth and active regions. The results display successful lesion registration and tracking over a longitudinal study. Lesion change results are graphically represented in the web-based user interface, and users are able to correlate patient progress and changes in the MRI images. The completed lesion and disease tracking tool would enable the eFolder to provide complete patient profiles, improve the efficiency of patient care, and perform comprehensive data analysis through an integrated imaging informatics system.

Keywords: multiple sclerosis, brain normalization, lesion registration, data mining, DICOM

1. INTRODUCTION

This manuscript presents new advancements in a disease-centric patient record and data management system for Multiple Sclerosis patients. Previously, we have introduced the MS eFolder as a big data analysis tool to relate imaging features and clinical data in longitudinal studies. This year, we are including a lesion registration component to define lesion locations and lesion registration to track individual lesion progressions. The manuscript explains the methodology of brain normalization, lesion registration, and database/GUI modifications to store and display the newly incorporated data.

1.1. Multiple Sclerosis research

Multiple Sclerosis (MS) is an autoimmune neurological disease that affects approximately 2.5 million people worldwide, and proximately 200 new patients are diagnosed with MS each week in the United States. The body's own immune system attacked the central nervous system, causing damages and scar tissues (called lesions) in brain parenchyma, spinal cord, and optic nerves[1]. There is no known cure for MS, and thus treatments for MS include disease management, reducing number and severity of attacks, and improve patients' ability to function in daily lives[2, 3]. Therefore, longitudinal disease tracking of patients become key in MS treatment.

Magnetic Resonance Imaging (MRI) is a commonly-used tool in diagnosing and monitoring MS by visually displaying lesions[4]. In longitudinal tracking, existing individual MS lesions need to be identified and quantified for monitoring patients' responses to treatments as well as disease progress. To solve these challenges, an imaging-informatics based eFolder has been designed to store and display MS patient data with MR images and MS lesion quantification results. The benefits of the eFolder include integrated patient data repository, an automatic lesion detection and quantification system to allow disease tracking on MR, and a data mining tool for both clinical and research purposes.

1.2. The MS eFolder project

The MS eFolder is a disease-centric, imaging informatics based electronic system that allows management of patient data, imaging data, and post-processing data. The purpose is to integrate patient's neurological examinations, demographic data, and disease history with patient's radiological images to help track a patient's disease profile and disease progression. The MS eFolder system has three main design components: database, graphical user interface, and a computer-aided detection (CAD) system that can quantify lesion volume and number of lesions. The CAD system is used to detect disease changes on the imaging level, including changes in quantity and size of 3-dimensional lesions and changes in brain parenchyma ratio, and correlate MS lesion characteristics (size, number, location) with patient's demographic data for research purposes.

1.2.1. eFolder Database

The eFolder database stores text data such as patient history, MR image locations, and lesion quantification results. Database schema has been developed in MySQL The database structure is built such that one single patient has a unique data entry regarding demographics and social data, has a list of all MR studies regarding to MS, and a list of all post processing results available for that patient. The data therefore is patient-centric and allows quick access to a patient's historical data. Patient demographic data is collected and designed via physicians' survey forms. The imaging database follows the DICOM structure to store metadata from headers. The CAD results database stores quantified lesion statistics on both study and image level. The purpose of the database design is to allow patient lookup and query/retrieve of images based on disease profiles and MS lesion characteristics.

1.2.2. Computer-aided Lesion Detection (CAD) and quantification system

The MS CAD algorithm is designed to output lesion volumes, lesion locations, and total lesion load. The detailed algorithm design splits up into three parts: preprocessing, lesion voxel identification by probability thresholding, and lesion quantification. The algorithm has been prototyped in MATLAB and has been refined to increase post-processing efficiency by reducing processing time.

The CAD algorithm is designed on 3-D MRI brain images. It uses T1 and FLAIR (Fluid attenuated inversion recovery) axial sequences. The algorithm converts the series of MR images into a three-dimensional matrix for 3-D lesion analysis. Lesion voxel classification is based on Statistical Parametric Mapping (SPM) brain image analysis toolkit for MATLAB[4]. Grey matter and white matter are first segmented, and an expectation minimization algorithm for k multidimensional Gaussian mixture[5] is applied to the brain images. The estimation results are used to determine the likelihood of a lesion voxel based on whether the voxel intensity is outside the predetermined normal range. The normal range is current set at within 3 standard deviations of normal FLAIR intensities.

The results from the voxel classification algorithm is then clustered and quantified based on DICOM values, and the final output includes individual lesion volumes in 3D, lesion locations in coordinate space, and total lesion load for the study.

1.3. IHE Post-processing workflow with MS eFolder and big data model

Previously, we presented a workflow profile and simulation for MS eFolder within a clinical environment. The workflow is designed based on the Integrating the Healthcare Enterprise (IHE) post-processing workflow profile[6,7]. Figure 1 shows how the MS eFolder is hit in the clinical workflow.

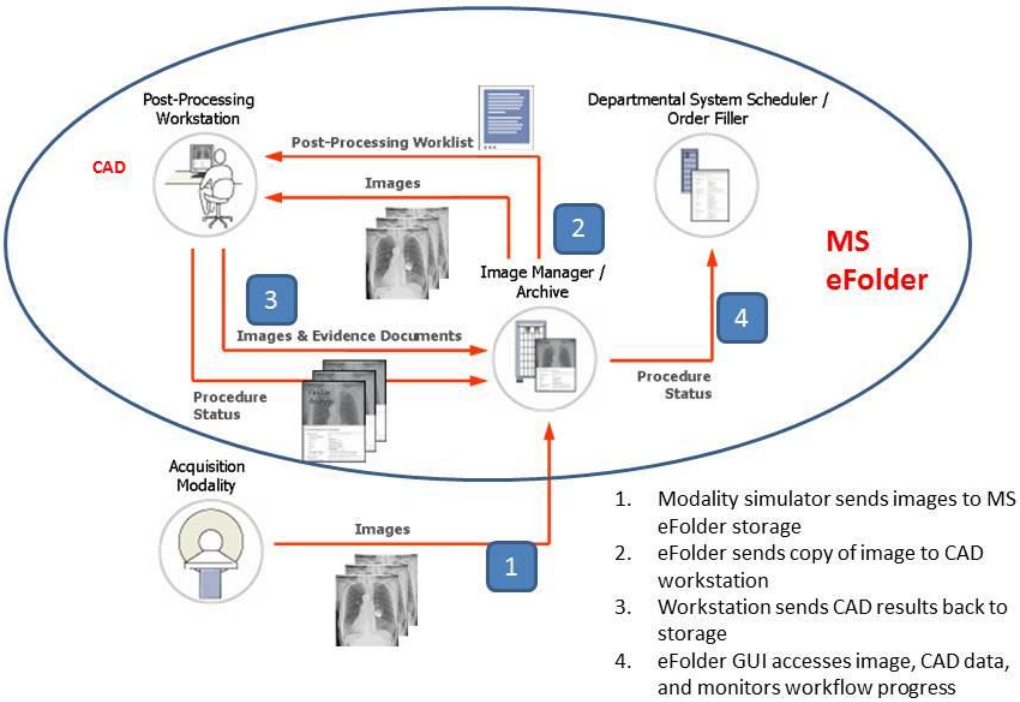


Figure 1. MS eFolder workflow diagram with IHE postprocessing profile. The blue circle indicates all of the components included in the eFolder. The steps 1 through 4 indicates the order of workflow of the demonstration.

The workflow for MS eFolder integration is defined in four steps:

1. MR images are sent from modality simulator to the eFolder server for archiving
2. The eFolder server sends a copy of the images to the CAD Workstation for postprocessing analysis
3. The CAD Workstation sends the completed CAD report back to eFolder server for archiving
4. At the completion of each of the previous steps, a status tracking tool inside eFolder displays alerts of the study progress to the user

Figure 2 shows an example of complexity of a MS patient's data, and how eFolder's data model is designed to accommodate the different types of data.

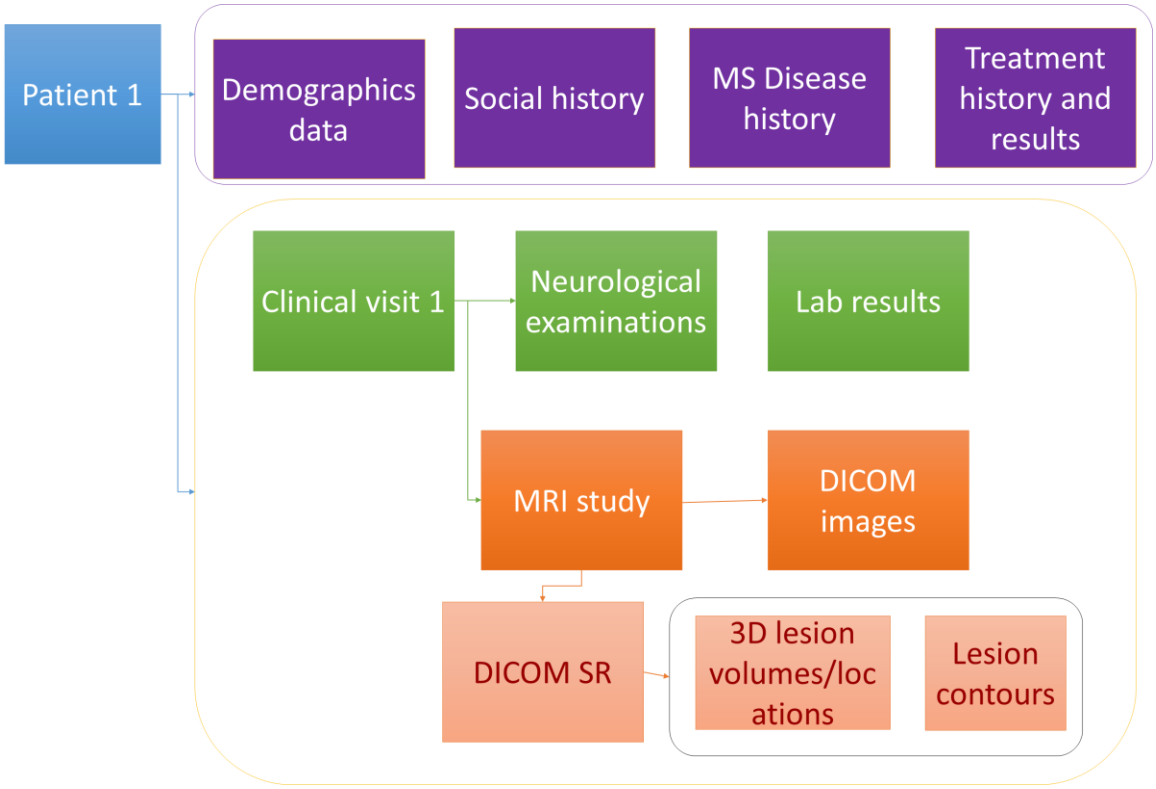


Figure 2. MS eFolder data model that showcases complexity of one patient's data. Each patient, in addition may have multiple clinical visits and scans that further expands the data model. The number of patients, and the number of studies each patient has, highlights the need for big data analysis in medical imaging informatics.

The data model allows the system to store comprehensive patient data and complex data analysis based on clinical findings and imaging features (i.e. lesion volume, brain volume, etc.). With a large enough quantity of data, users can observe disease trends in a large population of patients, disease trends for individual patients, and data mining for decision support.

1.4. The need for lesion registration

MS lesion changes in longitudinal studies have shown to be related to advancing disabilities and symptoms[10]. Furthermore, brain lesion locations is related to symptoms, types of MS, and disability scores[11]. Tracking new lesions and other lesion changes is important in understanding the disease, predicting symptoms and severity of symptoms, and creating personalized treatment options. Therefore, instead of tracking overall aggregated lesion volumes, there is a need for localizing and registering MS lesions to automatically track individual lesion volume changes and identifying new and most active lesions.

In order to incorporate the lesion registration module in the eFolder system, the computer-aided detection and quantification algorithm is modified to include:

- Brain normalization to a template
- Lesion location identification via coordinates and subcortical structures
- Algorithm of identifying and registering lesions in longitudinal studies

In this paper, we will explore the methodologies of each steps of lesion registration, show the web-based GUI results to display lesion locations and changes, and discuss normalization results and any future work to be completed.

2. METHODS

2.1. Data collection

Image and patient data used in the MS eFolder setup is the same as existing eFolder data that has been collected over 3 years. A total of 72 patients are collected: 36 Hispanic and 36 Caucasian patients. The patients of two groups are matched by gender, age (within 5 years), disease duration (within 5 years), and disease type (all are relapse-remitting). All brain MR studies are collected at University of Southern California Academic Medical Center and Los Angeles County Hospital. MR images are in DICOM format and anonymized. All studies contain noncontrast T1 and FLAIR axial slices as required by the MS CAD algorithm. In addition, 4 patients have been selected with longitudinal studies. These patients (randomly selected) have had yearly MRI scans completed at USC-LAC hospitals. Four imaging studies for each patient, taken from years 2009 to 2014, are collected to test the longitudinal study viewer.

2.2. CAD data

The CAD algorithm has been performed on all 72 studies and additional longitudinal studies. In addition, lesion contours by neuroradiologists at USC have been collected to act as gold standard for lesion detection. The contours were done manually by two neuroradiologists on Fuji Synapse 3D post-processing client in the clinical environment. Currently the manual contours are used in this project, as the CAD algorithm is still being refined for more consistent accuracy in its results. Figure 3 shows the MATLAB results of a sample data. The original DICOM images are converted to NIFTI format for processing. The output lesion contour images are saved in NIFTI format as well.

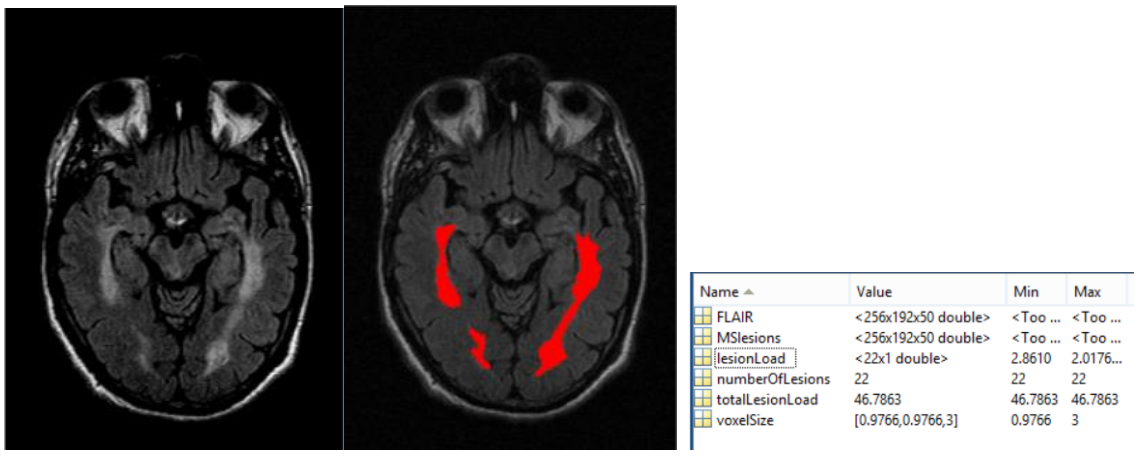


Figure 3. Raw MATLAB output of MS CAD. Left: original FLAIR axial image. Middle: lesion contour overlaid on the FLAIR image. Right: MATLAB outputs. lesionLoad includes individual lesion volumes, numberOfLesions indicates that there are 22 unique lesion bodies in this study, and totalLesionLoad is the sum of all lesion volumes.

2.3. Brain Normalization

In order to identify each individual lesions in the same patient in separate longitudinal studies, the studies have to be normalized to a template brain to map the lesions based on their locations. To accomplish this, the brain warping technique using MATLAB's Statistic Parametric Mapping toolkit is used to prototype the normalization methodology. SPM's voxel-based morphometry methodology[12,13] is able to warp the subject brain into a template brain with refinement in the subcortical structures. The template used in this algorithm is the ICBM 152 Nonlinear Atlases version 2009[14,15], which includes labeling of 152 different subcortical structures that is needed for lesion location identification. A semi-automatic methodology, which is included in the SPM toolkit, is applied to the warping of the brain images.

Firstly, the SPM12 module titled "Normalisation: Estimate and Write" is launched via MATLAB. The normalization process requires three inputs: the FLAIR axial MRI images, the lesion contour images (in registration with the FLAIR images), and the tissue probability map (the ICBM template). The FLAIR images are warped according to the tissue probability map, and the resulting warp parameter is calculated and applied to the lesion contour images, thus warmping the lesion space to the template spae. The estimation options, including bias regularization, Gaussian smoothing of bias, affine regularization, and warping regularization have been set. Figure 4 shows the MATLAB GUI that operates the normalization process.

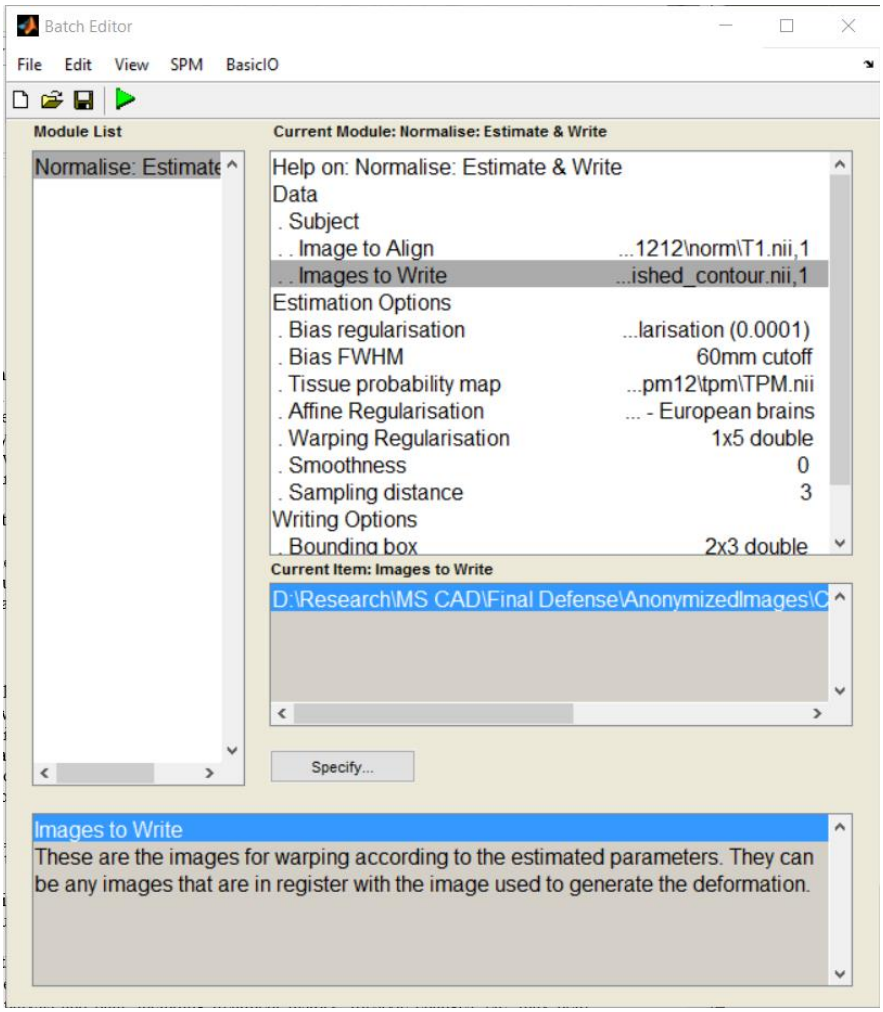


Figure 4. SPM GUI for applying brain normalization

2.4. Lesion localization and registration

After brain normalization is completed, the lesion contour images are registered to the tissue probability map, which includes a labeling map. The labels, in XML format, are included in the SPM toolkit. The coordinates of centroid for each lesion cluster is first extracted, and using the registered coordinates of the label map, the centroid location is obtained.

During the clustering procedures, each lesion is assigned an identifier. Due to different parameters in each study such as number of lesions, the identifiers are not consistent from one imaging study to the next in a longitudinal study. Therefore, a lesion registration algorithm is applied so link different lesion identifiers. The registration algorithm involves matching two normalized studies against each other. Because of lesion growth variations and contouring differences, it is possible for each lesion in subsequent studies to be registered to multiple lesions in the previous study, and one lesion in the previous study may have occupied the same coordinate spaces as multiple lesions in the subsequent studies. For our prototyping purposes, we chose the lesion from the previous study that occupies the most contour areas of the targetted lesion as the "registered lesion". In order to achieve this, for each lesion in the subsequent study, a histogram count is performed to determine the location of the lesion relative to the lesion space in the previous study. Figure 5 illustrates the lesion registration methodology.

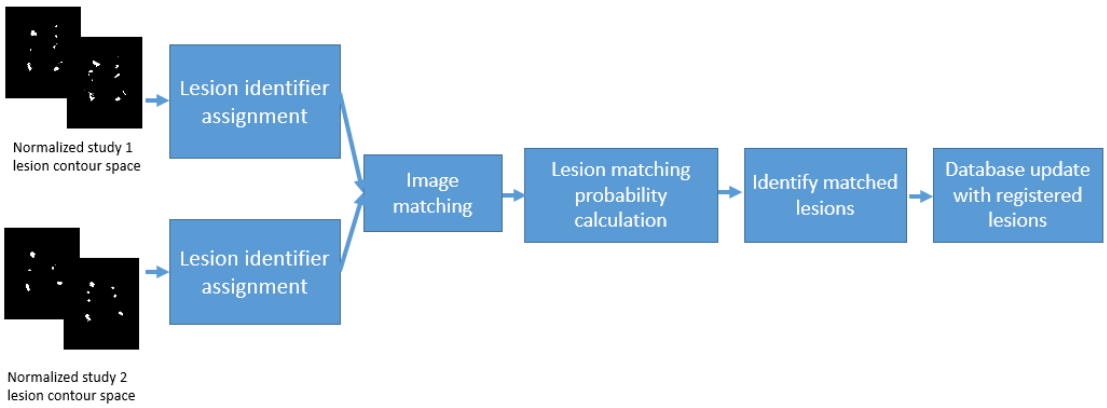


Figure 5. Workflow of MS lesion registration in longitudinal studies.

2.5. Database support and web-based GUI modification

The eFolder's web-based GUI has been modified to include lesion tracking and comparison results. Figure 6 shows an example of tracking volume change for a specific lesion.

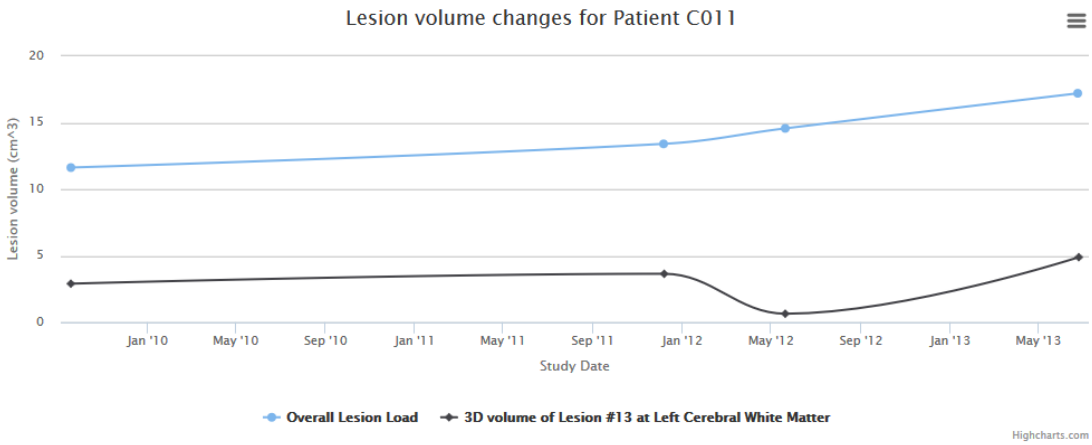


Figure 6. Graphical lesion volume tracking of longitudinal studies. The top line is the overall lesion volume change from 2010 to 2013, and the bottom line graph is the lesion volume change of a specific lesion (#13) located in left cerebral white matter. It appears that lesion 13 seemed to drop in volume. Users can correlate this phenomena to the eFolder to determine whether it was a change or response to new drug therapies or other treatment possibilities.

The dynamic lesion tracking module allows searching for specific lesions and showing lesion contour changes by overlaying images and showing in different colors, which is shown in Figure 7.

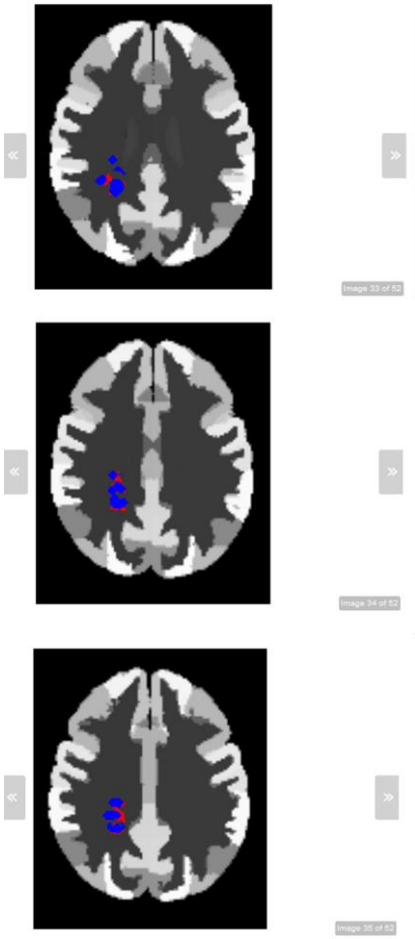


Figure 7. Lesion contour comparisons of a longitudinal study. The different-colored contours show the lesion locations and the shape changes. Lesion contour from the newer study is overlaid on lesion contour from the older study.

3. RESULTS AND DISCUSSIONS

Brain normalization and lesion registration methodologies have been prototyped and tested on four longitudinal studies. The database has been updated with lesion registration information, and the web-based GUI has been updated to display lesion tracking results. The initial normalization and registration results are effective in determining general lesion locations.

The completed registration module allows for further complex data analysis for longitudinal studies. For example, user may look up on the location of lesions that have increased in volume the most, where any new lesions have appeared, and proportion of lesion volumes in left and right cortical white matter regions. The variety of data available for physicians and clinicians has significantly increased. In the future, more complex data analysis toolkits are being developed to allow for more comprehensive patient data repository and data mining.

4. CONCLUSION

We have presented the MS eFolder system that utilizes the idea of big data storage and analysis in medical imaging and imaging informatics. The system has the capability of tracking disease progress in longitudinal studies, especially correlating symptoms and disabilities with lesion volumes in MRI. We have developed a lesion registration and localization modules to track lesion growth and appearance in longitudinal studies. The new modules allow for tracking individual lesion volume changes in 3D, identifying lesion location data, and identify new and active lesions. The lesion registration module can link symptoms and disease characteristics with lesion locations and volume changes at certain locations. The eFolder system has become a more powerful tool in both clinical care and research in the treatment and study of multiple sclerosis.

REFERENCE

[1] National Multiple Sclerosis Society http://www.nationalmssociety.org/about-multiple-sclerosis/index.aspx

[2] S Ringold, C Lynn, R Glass “Multiple Sclerosis” JAMA. 2005;293(4):514

[3] D Hafler “Multiple Sclerosis” J. Clin. Invest. 2004, 113(6):788-794

[4] Statistical Parametric Mapping http://www.fil.ion.ucl.ac.uk/spm/

[5] P Tsui “Fast EM_GM” http://www.mathworks.us/matlabcentral/fileexchange/9659 downloaded from MATLAB Central

[6] K Ma “Integration of Imaging Informatics-based Multiple Sclerosis eFolder System for multi-site clinical trials utilizing IHE Workflow Profiles”, Proc. SPIE 8674, Medical Imaging 2013: Advanced PACS-based Imaging Informatics and Therapeutic Applications, 86740A (March 29, 2013); doi:10.1117/12.2008266;

[7] IHE Post-processing Workflow http://wiki.ihe.net/index.php?title=Post-Processing_Workflow

[8] “What is Big Data?” by University Alliance at Villanova University http://www.villanovau.com/resources/bi/what-is-big-data/#.VOUo6JjF_Vo

[9] J Manylka et al. “Big data: The next frontier for innovation, competition, and productivity.” McKinsey&Company, May 2011. http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation

[10] L. K. Fisniku, P. A. Brex, D. R. Altmann, et al. “Disability and T2 MRI lesions: a 20-year follow-up of patients with relapse onset of multiple sclerosis” Brain Mar 2008, 131 (3) 808-817; DOI: 10.1093/brain/awm329

[11] Eloyan, Ani et al. “Health Effects of Lesion Localization in Multiple Sclerosis: Spatial Registration and Confounding Adjustment.” Ed. Tim Friede. PLoS ONE 9.9 (2014): e107263. PMC. Web. 25 Feb. 2016.

A handheld computer-aided diagnosis system and simulated analysis

Mingjian Su[*a], Xuejun Zhang[a,d,e], Brent Liu[b], Kening Su[c] and Ryan Louie[b]
[a]School of Computer, Electronics and Information, Guangxi University, Nanning, Guangxi 530004, P. R. China (*e-mail: 352912730@qq.com)
[b]IPILab, Biomedical Engineering Dept, University of Southern California, Los Angeles, CA 90033
[c]Anorectal section, The First Affiliated Hospital of Guangxi Univ. of Chinese Medicine, Nanning, Guangxi 530004, P. R. China
[d]Guangxi Key Laboratory of Multimedia Communications and Network Technology (Cultivating Base), Guangxi University, Nanning 530004, China
[e]Guangxi Colleges and Universities Key Laboratory of Multimedia Communications and Information Processing, Guangxi University, Nanning 530004, China

Abstract

This paper describes a Computer Aided Diagnosis (CAD) system based on cellphone and distributed cluster. One of the bottlenecks in building a CAD system for clinical practice is the storage and process of mass pathology samples freely among different devices, and normal pattern matching algorithm on large scale image set is very time consuming. Distributed computation on cluster has demonstrated the ability to relieve this bottleneck. We develop a system enabling the user to compare the mass image to a dataset with feature table by sending datasets to Generic Data Handler Module in Hadoop, where the pattern recognition is undertaken for the detection of skin diseases. A single and combination retrieval algorithm to data pipeline base on Map Reduce framework is used in our system in order to make optimal choice between recognition accuracy and system cost. The profile of lesion area is drawn by doctors manually on the screen, and then uploads this pattern to the server. In our evaluation experiment, an accuracy of 75% diagnosis hit rate is obtained by testing 100 patients with skin illness. Our system has the potential help in building a novel medical image dataset by collecting large amounts of gold standard during medical diagnosis. Once the project is online, the participants are free to join and eventually an abundant sample dataset will soon be gathered enough for learning. These results demonstrate our technology is very promising and expected to be used in clinical practice.

Keywords: Computer aided diagnosis, hadoop, cloud, pattern recognition

1. INTRODUCTION

Skin disease is a common ailment and frequently-occurring disease that greatly influence people's health, such as leprosy, scabies, fungal or bacteria infection. Once the disease is occurred, the morphology, structure and functions of the infected skin will be changed with the development of disease. The skin would get massive ulcer in the worst. In developing countries, especially in remote mountain area, the medical resources are rare that makes skin disease situation serious. With the help of digitization, cloud computing, machine learning technology and wireless communication technology, CAD (Computer Aided Diagnose) system on smartphone is changing every field in medical industry.

From longevity country Japan, Ministry of Education, Science, Culture and Sports supported a mobile phone-based communications system named “Electronic Doctor’s Bag”[1], which can easily send biological information with multiple high-definition images, and let nurse instead of doctor carry out home medical treatment. Shabnam Kia et al. [2] designed a CAD system for skin illness detection by using intelligent artificial neural network to classify the skin sonograms. Their results show the high capability of this method in diagnosis and classification of the skin diseases. “SKINVISION”[3] is a handheld skin CAD system developed by a Romania develop team. It’s the first CE certified melanoma App that helps people detect melanoma in an early stage by using fractal geometry algorithm. This App could detect 73 melanoma cases out of 100 cases, and the usage of these new tools can reduce the number of doctors needed, cost, and improve medical diagnosis efficiency.

While only a small number of skin diseases accounts for most visits to the physician, thousands of skin conditions have been described. Classification of these skin disorders is a difficult task, since the underlying etiologies and pathogenetics are often not known [4], and the shortage of medical resources is a social problem. In our experiment, we chose two common skin disease, tinea pedis and pityriasis, as our CAD classification target. This paper mainly focuses on the methodology of building a CAD expert system with the ability of self-upgraded by the expanded number of users. The experiment mainly divided into two parts: 1. Training the tinea pedis and pityriasis sample data set. 2. Distributed computing classification query task.

2. METHODOLOGY

2.1 Materials

The datasets using in our experiment is taken from the camera by cellphone or other photographic apparatus, including the images from patients with tinea pedis and pityriasis disease. The original image is pre-processed for feature extraction.

2.1.1 Tinea pedis

Tinea pedis is also called Athlete's foot, it’s a fungal infection. The infected skin appear red or ulcerative, scaly, flaky, with soft and white skin [5][6]. When the doctors take a marcophotograph to the lesion of tinea pedis and transfer to the central computer, metadata server will do the processing of incoming images. After the processing of gray, edge detecting, binaryzation, corrode, dilate, connected region labeling, dot product, the lesion area of original photo would be extracted. Then the texture of lesion area is extracted by using 2D Gabor filter bank, which could generate texture in different scales and directions depending on the Gabor kernels. The processing procedure is shown in Figure 1.

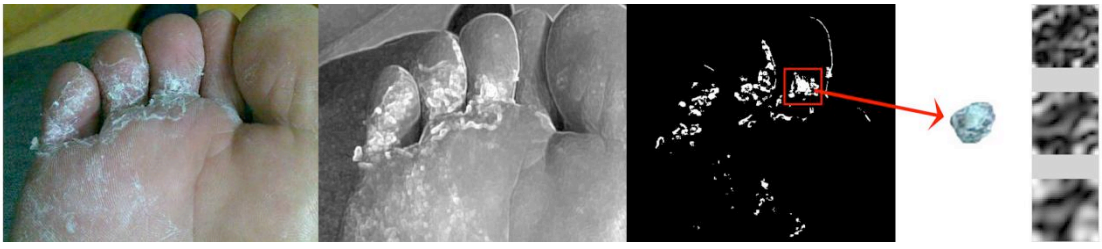


Figure 1. The procedure of extraction of tinea pedis texture

2.1.2 Pityriasis

Pityriasis is another common body skin disorder which cause by poor cleaning condition. The color between pityriasis skin and normal skin is about the same. When the body has slight itching, color changed can be found on the skin where you feel itch by looking into the mirror. The lesion photo is firstly enhanced by adjusting gamma, contrast, histogram equalization, Gaussian blur. According to the pathological characteristics of pityriasis, the infected area turns into irregular circle [7][8]. Hough transform is applied to the detection of similar area with the circle. The lesion area of pityriasis is particularly larger than the other ulcer. Figure 2 shows the processing procedure of extract the characteristic of pityriasis. This characteristic will also write into the feature table.

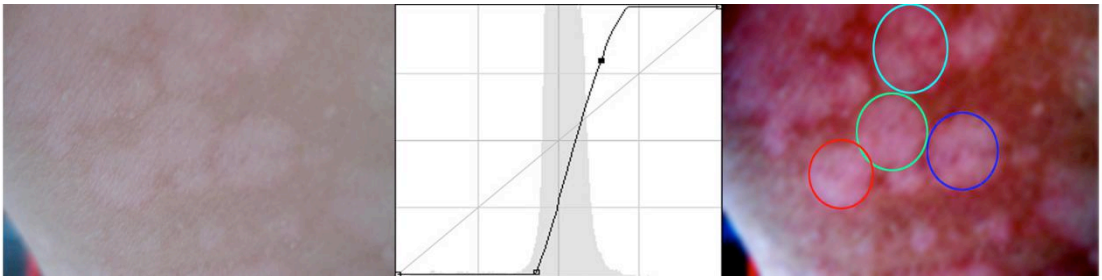


Figure 2. The processing procedure of extract the characteristic of pityriasis

2.2 Pattern recognition overview

In order to build a novel medical image data set to collect vast amounts of medical diagnosis evidence, a pattern recognition based system is proposed with the skill of self-upgraded by expanded the number of users. The overview of this expert system is shown in Figure 3.

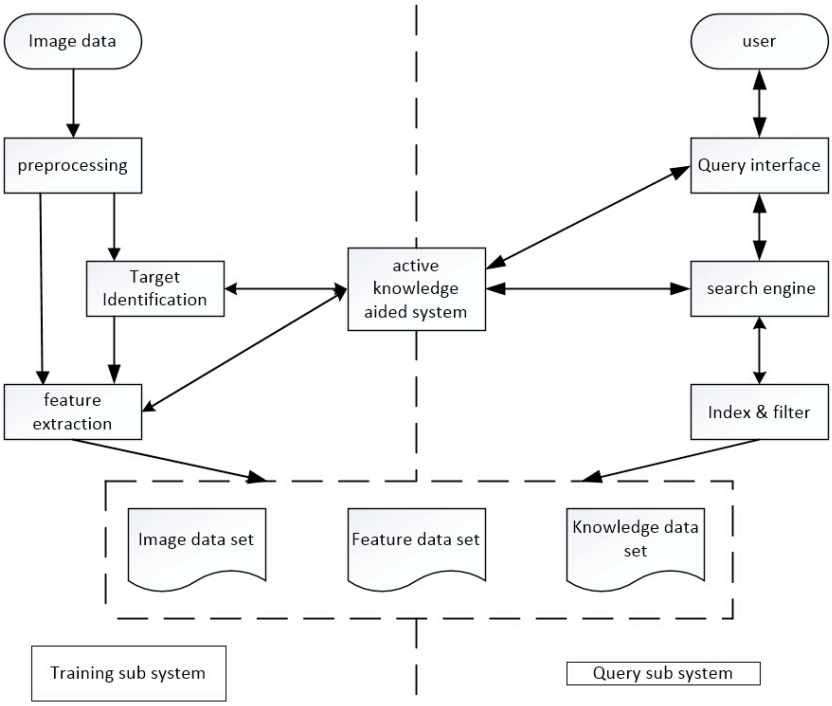


Figure 3. Pattern recognition overview

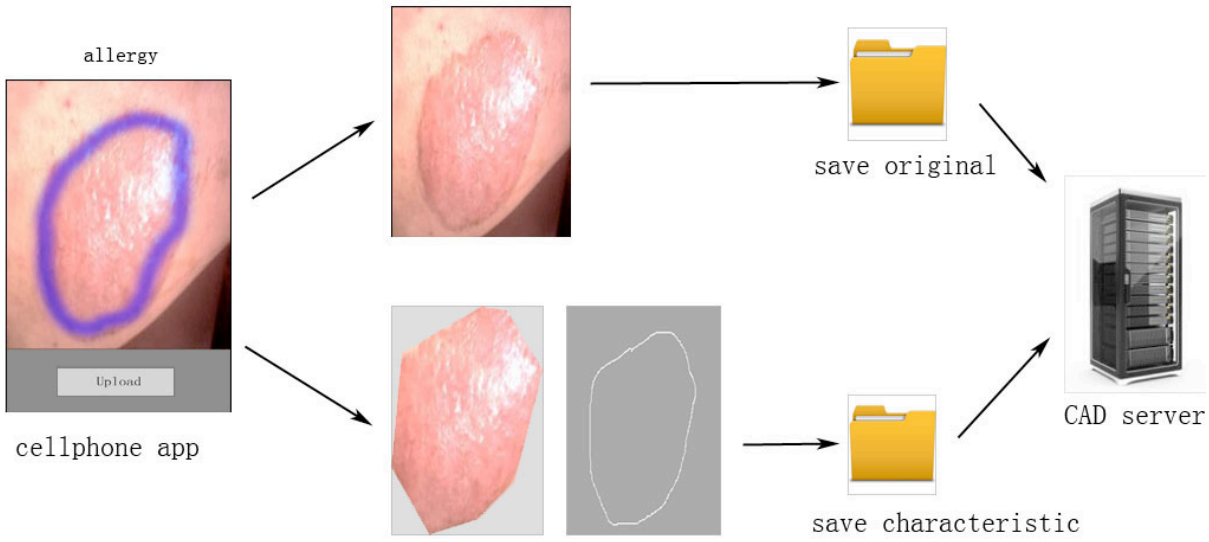


Figure 4. CAD storage processing of pathology image.

2.3 The pathological feature extraction

In the training sub system, experts (doctors) pointed out the location of the lesion on the screen, and the contour of lesion is extracted by CAD and meanwhile draw by doctors as shown in Figure 4. In our research, the characteristic of two kinds of skin disease is extracted, that are tinea pedis and pityriasis.

2.4 Feature table setup

In our approach, each image captured by the expert has a feature table. Each table record several attribute, such as textural features, Discrete Cosine Transform(DCT) characteristic, circularity, original scaling feature, text-based retrieval. DCT characteristic, textural features, original scaling feature are sampled by hash method, and the hash characteristics are generated by the following steps:

- Reducing the resolution of pictures to remove high-frequency details, that is a necessary step for saving system resources. In our approach, the picture is resized into only 64 pixels, which is 8*8 square, thus we can compare images in any size, any proportion.
- Dropping the color of small image from last step.
- Calculating the mean value of every pixel in the small image.
- Comparing the gray value of each pixel in the small image. Mark 1 if the pixel gray larger or equal to the mean value, mark 0 if below it.
- Combine the result of last step in certain sequence, forming a 64bit integer, which is the hash code of the image.

The 2nd attribute in our feature table is the DCT hash characteristic. Discrete Cosine Transform is a common transform usually used in image processing field. DCT decompose the frequency of an image and gather them into trapezoid data. 2 dimensions (image) discrete DCT transform can be represented by:

$$
\begin{aligned}
F(0,0) &= \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y) \\
F(0,v) &= \frac{\sqrt{2}}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\cdot\cos\frac{(2y+1)v\pi}{2N} \\
F(u,0) &= \frac{\sqrt{2}}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\cdot\cos\frac{(2x+1)u\pi}{2N} \\
F(u,v) &= \frac{2}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\cdot\cos\frac{(2x+1)u\pi}{2N}\cdot\cos\frac{(2y+1)v\pi}{2N}
\end{aligned}
\tag{1}
$$

The DCT inverse transformation can be represented by:

$$
\begin{aligned}
f(x,y) = \frac{1}{N}F(0,0) &+ \frac{\sqrt{2}}{N}\sum_{v=1}^{N-1} F(0,v)\cos\frac{(2y+1)v\pi}{2N} + \frac{\sqrt{2}}{N}\sum_{v=1}^{N-1} F(u,0)\cos\frac{(2x+1)u\pi}{2N} \\
&+ \frac{2}{N}\sum_{u=1}^{N-1}\sum_{v=1}^{N-1} F(u,v)\cos\frac{(2x+1)u\pi}{2N}\cdot\cos\frac{(2y+1)v\pi}{2N}
\end{aligned}
\tag{2}
$$

Where f(x,y) is image pixel, F(u,v) is the result of the coefficient matrix. The DCT hash characteristic is calculated by the following steps:

- Resize Original image to 32*32 resolution, and turn the pixels into gray color.
- Using 32*32 DCT transform, then choose 8*8 matrix in the left corner of the resulting image.
- Comparing the mean value of each pixel, record the 64bit binary code to an integer.

The result of DCT hash characteristic could roughly tell the relative proportions of mean frequency. If the contexts of the picture not change, the DCT hash will be almost the same. This procedure could lower and influenced the Gamma and histogram variability. The structure of the feature table is:

```
struct feature_table
{
int textural_features;
int DCT_characteristic
short circularity;
int Original_scaling_feature;
char text_based_retrieval[99];
};
```

In general, the table could be more complicated. Depends on the balance of efficiency and quality you setup the entire system.

2.5 Hadoop distributed computing classification environment set up

Hadoop is a distributed computing framework powered by Apache Software Foundation, which is able to realize distributed computing on common desktop computer efficiently [9]. In [10] Chu, et al. 2007, implemented a machine learning scheme in distributed computing base on principle of hadoop, variety of learning algorithms including locally weighted linear regression (LWLR), k-means, logistic regression(LR), naive Bayes (NB), SVM, ICA, PCA, Gaussian discriminant analysis(GDA), EM, and backpropagation (NN) was apply on their system. But the efficiency of algorithms can not balance to the state-of-art computational capabilities [11]. Our CAD system is built based on Hadoop framework that is mainly consist of four parts: User platform on cellphone, Namenode, Datanode and Metadata server on server racks. There are two daemon processes running on the Metadata server: JobTracker and Cellphone listener. The JobTracker plays an important role in controlling and managing the TaskTracker, which runs in the Datanode. TaskTracker is responsible for performing tasks, and must run with Datanode. Therefore a Datanode is not only a storage node, but also a compute node. The JobTracker puts the map or reduce task to the idle TaskTracker in parallel. If one TaskTracker failure, the JobTracker will switch the task to another idle TaskTracker. This feature is an advantage of the Hadoop platform.

Usually, Namenode server plays an important role in a Hadoop file system (HDFS). Namenodes put some records to the metadata such as segmentation state of datasets, and where the datasets come from, or the state of a Datanode. The Cellphone listener in charge of collecting mass of small files that sent by users; extracting the characteristic into the

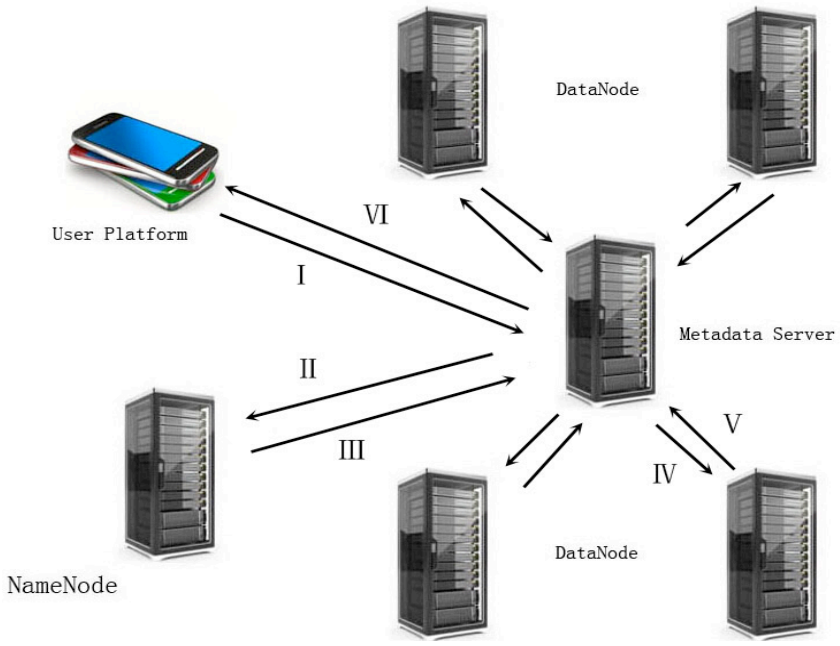


Figure 6. The operation mode of our distributed computing framework

feature table; and transfer the table into the Namenode. Mapping and Reducing is the main way for a Hadoop system to deal with information. Namenode uses Map function generate a set of key, value pair as the transition set, and transfer to the reduce function in Datanodes. The reduce function receives a set of key with related value (mostly represented the image information), then calculate with pattern recognize algorithm and generate the final result. The operation mode of our distributed computing framework is shown in Figure 6.

3. EXPERIMENTAL RESULTS

3.1 The accuracy of retrieval

According to the above mentioned system, we carry out our image retrieval task by using combination features of images. The synchronous combination retrieval [12] is applied in our approach. The similarity measurement testing can be made by combination features of images in one query. Once a certain weight is given in each test, the final distance could be calculated by weighted summation. The procedure of synchronous combination retrieval is shown in Figure 7.

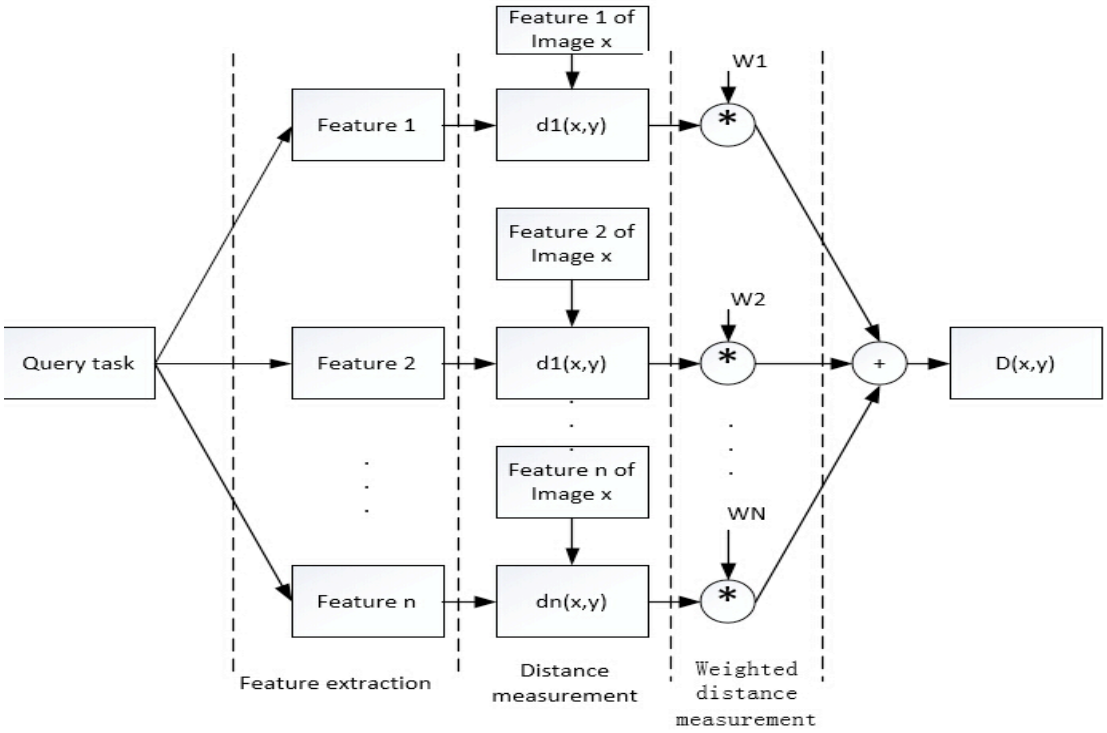


Figure 7. The procedure of synchronous combination retrieval

Where *Wi* is the weight number, *di* is the distance between the retrieval image and the image *x* on *ith* feature. In this approach, Hash values can be compared using the Hamming distance algorithm, which can be calculated by:

$$\sum_{i=1}^{n} W_i = 1, \quad \mathrm{d(x, y)} = \sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{W_i} d_i(x, y) \tag{3}$$

In order to verify the superiority of our method, the image retrieval test is done by using 3 different characteristics which is textural features, Discrete Cosine Transform (DCT) characteristic, and original scaling feature, as well as a combination test of these three characteristics. 50 tinea pedis pathological pictures and 50 pityriasis pictures are used to verify the accuracy of the results. Random rotation (R), resize (S), added pepper noise (N) was applied to these pictures. The race of result images represents the accuracy of each method. Image retrieve results are shown in the following tables.

Table 1. The race of texture retrieval on 100 pathological pictures

RT	**n=1(%)**	**n<=2(%)**	**n<=3(%)**	**n<=4(%)**	**n<=5(%)**	**n<=20(%)**
R	45.25	51.25	53.14	60.23	63.23	85.63
S	42.68	57.54	59.36	63.74	65.25	82.56
N	60.21	66.84	70.54	71.32	75.55	90.36

Table 2. The race of DCT characteristic retrieval of 100 pathological pictures

RT	**n=1(%)**	**n<=2(%)**	**n<=3(%)**	**n<=4(%)**	**n<=5(%)**	**n<=20(%)**
R	45.45	48.26	49.78	56.23	62.45	88.11
S	40.25	48.54	54.85	55.58	62.54	78.82
N	55.78	56.32	60.12	62.95	68.74	84.21

Table 3. The race of original hash characteristic retrieval on 100 pathological pictures

RT	**n=1(%)**	**n<=2(%)**	**n<=3(%)**	**n<=4(%)**	**n<=5(%)**	**n<=20(%)**
R	n/a	n/a	n/a	n/a	n/a	n/a
S	38.78	41.45	49.58	54.68	56.21	65.29
N	40.52	41.85	45.63	50.11	55.86	70.21

Table 4. The race of combination characteristic retrieval on 100 pathological pictures

RT	**n=1(%)**	**n<=2(%)**	**n<=3(%)**	**n<=4(%)**	**n<=5(%)**	**n<=20(%)**
R	74.86	75.13	75.68	77.33	77.41	83.21
S	66.68	67.25	75.54	77.51	75.23	81.25
N	70.21	71.44	75.81	79.25	80.54	83.93

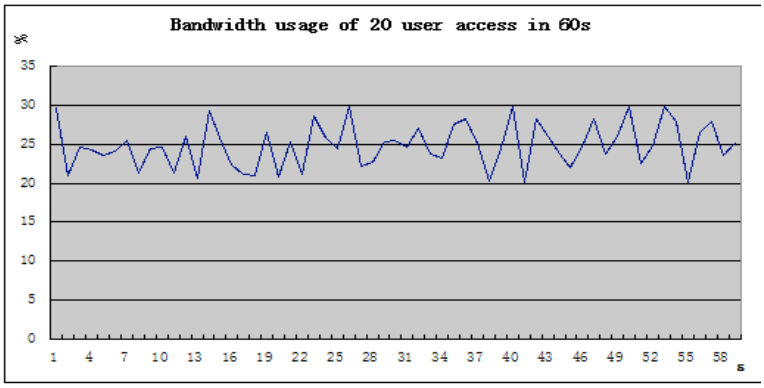


Figure 8. Bandwidth usage of 20 users online in 60s

Where *n* represent the percentage of race of match result. RT refers to retrieval type. We only take a sample of top 20 percent of retrieval result. Compare Table 4 to other 3 tables, combination characteristic retrieval could yeild a better performance than the single method, especially improving the accuracy of rotated pictures.

3.2 Hadoop architecture pressure test

This experiment is setup on six HP ProLiant DL580 G5 server, 1000M local network, four Intel Xeon X7460 CPU(2.67GHz, 16M L3) in each server, 16GB ECC-RAM. The network topology structure is shown in figure 5. We save 5,000 pathology photo of skin illness from web search engine (e.g. Baidu, Google, etc), which is used to simulate the activity of cellphone taking and sending the picture files to the Hadoop. The files were store in each Datanode with hashcode, and then a photo with anonymous skin illness is recognized by our CAD system. By the time of first retrieval, it takes about 3s to return a result set. While using the combination characteristic retrieval, it takes about 10s and hit rate is 75%. Figure 8 shows Bandwidth usage of 20 users online in 60s. Figure 9 shows the results of hit race on pathological pictures.

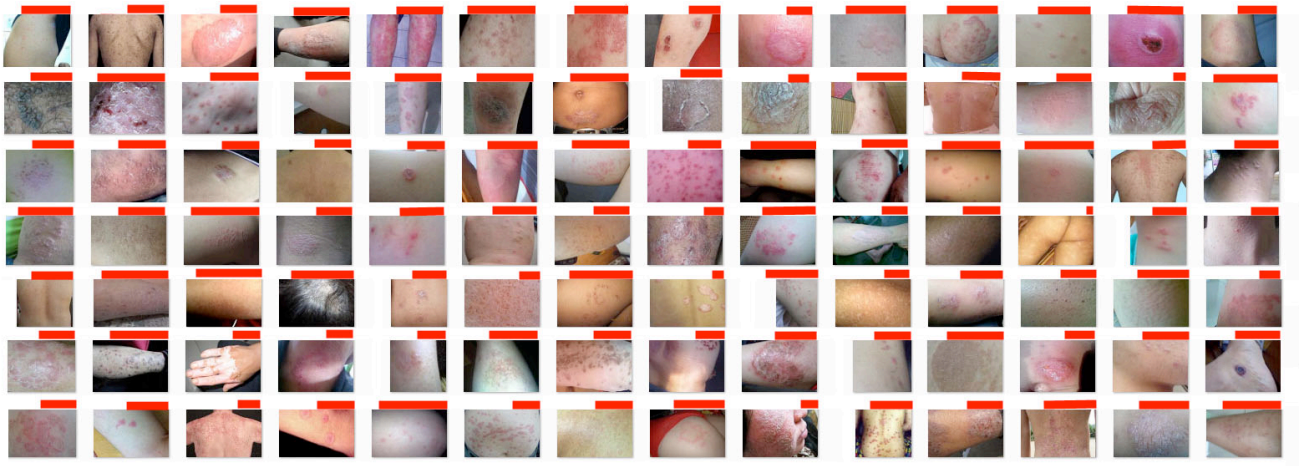


Figure 9. 98 pathology samples hit rate in server. Red line denotes a hit rate.

4. DISSCUSSION AND CONCLUSION

Nowadays, many people carry their own cellphone all day long, and the quality of photo capture reaches at a high level. Not only Japan but also many other country already begins to think much about the role of mobile phone in the mobile medical[13]. In our approach, the minimum require of users' cellphone is a smartphone with macro photography function (higher class cellphone is recommended). A photo capture by user should be close-up to lesion. Even so, the images captured from users' cellphone are in different scale, contrast, or from different angels. The result implies that the effective of combination characteristic retrieval is better than any single approach. The feature table in our approach is extensible that many kinds of characteristic can be store in the table. But large amounts of characteristics may cause system performance fall off, and setup a distributed computing environment may lead to an expansive cost. Hadoop framework could setup a distributed computing system on common personal computers, which could meet the needs of small and medium-sized enterprises. In our approach, Hadoop framework could easily expand and carry on the distributed computing task effectively.

In this paper, we present a Computer Aided Diagnosis (CAD) system on ubiquitous handheld device and simulated running on Hadoop framework. Our system has the potential help in building a novel medical image dataset by collecting large amounts of gold standard during medical diagnosis. Once the project is online, the participants are free to join and eventually an abundant sample dataset will soon be gathered enough for learning. These results demonstrate our technology is very promising and expected to be used in clinical practice.

ACKNOWLEDGMENT

This work was supported in part by two research support from the National Natural Science Foundation of China (No. 61262027 & 45627390), and in part by a research foundation project of the Guangxi Key Laboratory of Multimedia Communications and Network Technology (Cultivating Base). The author Xuejun Zhang gratefully acknowledges financial support from China Scholarship Council.

REFERENCES

[1] Yoshizawa, M., Yambe, T., et al. "A Mobile Communications System for Home-Visit Medical Services:The Electronic Doctor's Bag," 32nd Annual International Conference of the IEEE EMBS, 5496 - 5499(2010)

[2] Kia, S., Setayeshi, S., Shamsaei, M., Kia, M., "Computer-aided diagnosis (CAD) of the skin disease based on an intelligent classification of sonogram using neural network," Neural Computing & Applications, 22(6), 1049-1062(2013)

[3] https://www.skinvision.com

[4] https://en.wikipedia.org/w/index.php?title=Cutaneous_condition&redirect=no#cite_ref-Lynch_5-0

[5] Zhu, W. Y., “Color Atlas of Difficult Skin Disorders,” PEOPLE'S MILITARY MEDICAL PRESS, Beijing, China, 93(2008)

[6] Sun, Q. N., Liu, J., "Atlas OF DERMOSCOPY," PEOPLE'S MEDICAL PUBLISHING HOUSE, Beijing, China, 57(2015)

[7] Sun, Q. N., Liu, J., "Atlas OF DERMOSCOPY," PEOPLE'S MEDICAL PUBLISHING HOUSE, Beijing, China, 82(2015)

[8] Zhao, B., "COLOR ATLAS OF CLINICAL DERMATOLOGY," Phoenix Science Press, Nanjing, China, 686(2012)

[9] Shvachko, K., Kuang, H., Radia, S., Chansler, R., “The hadoop distributed file system,” Mass Storage Systems and Technologies (MSST), 2010 IEEE 26th Symposium on, 1-10(2010)

[10] Chu, C. T., Sang, K. K., Lin, Y. A., et al, “Map-Reduce for Machine Learning on Multicore,” Advances in Neural Information Processing System, 19,281-288(2007)

[11] Dave, T., Hanan, S., “A Web Database for Computer-Aided Detection and Diagnosis of Medical Image, ” Springer Berlin Heidelberg, 5116,265-272(2008)

[12] Punpiti, P., Nikitas, A., et al., "A Parallel Algorithm for Multi-Feature Content-Based Multimedia Retrieval," Seventh International Conference on Intelligent System, 1-3(2000)

[13] Bajorek, M., Nowak, J., “The role of a mobile device in a home monitoring healthcare system,” PROCEEDINGS OF THE FEDCSIS. SZCZECIN, 371-374(2011)

Comparing the role of shape and texture on staging hepatic fibrosis from medical imaging

Xuejun Zhang[*a,b,e,f], Ryan Louie[b], Brent J. Liu[b], Xin Gao[c], Xiaomin Tan[a], Xianghe Qu[a] and Liling Long[d]

[a]School of Computer, Electronics and Information, Guangxi University, Nanning, Guangxi 530004, P. R. China (*e-mail: xjzhang@gxu.edu.cn)

[b]IPILab, Biomedical Engineering Dept, University of Southern California, Los Angeles, CA 90033

[c] Department of Medical Imaging, Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou, Jiangshu 215163, P. R. China

[d]Department of Radiology, First Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi 530021, P. R. China

[e]Guangxi Key Laboratory of Multimedia Communications and Network Technology (Cultivating Base), Guangxi University, Nanning 530004, China

[f]Guangxi Colleges and Universities Key Laboratory of Multimedia Communications and Information Processing, Guangxi University, Nanning 530004, China

Abstract

The purpose of this study is to investigate the role of shape and texture in the classification of hepatic fibrosis by selecting the optimal parameters for a better Computer-aided diagnosis (CAD) system. 10 surface shape features are extracted from a standardized profile of liver; while15 texture features calculated from gray level co-occurrence matrix (GLCM) are extracted within an ROI in liver. Each combination of these input subsets is checked by using support vector machine (SVM) with leave-one-case-out method to differentiate fibrosis into two groups: normal or abnormal. The accurate rate value of all 10/15 types number of features is 66.83% by texture, while 85.74% by shape features, respectively. The irregularity of liver shape can demonstrate fibrotic grade efficiently and texture feature of CT image is not recommended to use with shape feature for interpretation of cirrhosis.

Keywords: Computer-aided diagnosis (CAD), Liver fibrosis, Texture feature, Waviness of surface, Support Vector machine (SVM), Feature selection

1. INTRODUCTION

Staging of the degree of hepatic fibrosis is an important issue in the diagnosis and therapeutic assessment of cirrhosis of the liver and chronic hepatitis. Fibrosis is caused by excessive deposition of extracellular matrix owing to histological and molecular reshuffling of various components such as collagens, glycoproteins, proteoglycans, and other

macromolecules within the extracellular matrix. These features, common to almost all patients with chronic liver disease, lead to the changes in the hepatic morphology, texture pattern, and degree of liver stiffness. Accurate assessment of hepatic fibrosis is crucial for the determination of the appropriate treatment because fibrosis is potentially a reversible process in the early stages.

The most popular and accurate way for fibrosis staging is liver biopsy, which is used for histological scoring and is still used as a reference test. However it is invasive and the bleeding always continued for two and a half minutes in average. To decrease the need for painful biopsies, non-surgical methods using CT or MRI modalities have been proposed to obtain images of internal liver, and various features can be extracted for different types of computer-aided detection/ diagnosis (CAD) systems to help radiologists diagnose and analyze liver lesions quantitatively. Among them, imaging features based on texture and shape has attracted more and more attention by researchers. Y. Gao et al. [1] proposed a texture image recognition method based on the correlation coefficient. Y. Guo et al. [2] developed a recognition method between normal livers and abnormal livers based on Gabor wavelet texture features to reach a higher recognition rate. E. Krusinska et al. [3] used shape features as a distinction analysis method to predict the histopathology of liver biopsy specimens. Huhdanpaa et al. [4] used the peripheral morphological characteristics that divide edge of renal cell carcinomas into four categories by the Fuhrman grade, indicating that contrast-enhanced CT imaging tumor edge can reveal histopathological features and indirectly reflect the change of renal cancer angiogenesis.

Our group has been investigating the performances of differentiate normal liver from cirrhotic case by texture features [5] and shape features[6]. However, quantification and classification of hepatic fibrosis between different 5 grades is an extremely difficult task and the results are still unknown. Furthermore, some studies suggusted that integrating the texture features into shape features with SVM classification of cirrhosis may improve the overall performance of cirrhosis detection. This study is focused on the effective selection of texture features together with shape features for fibrosis classification. We utilized 15 texture features and 10 shape features calculated from all of the CT datasets. Each combination of features is evaluated using support-vector-machine (SVM) with leave-one-case-out method to select the optimal feature subsets according to their performance.

2. METHODOLOGY

2.1 Materials

149 patients were scanned by CT (GE Lightspeed VCT) from June 2009 to March 2012 at the Department of Radiology, First Affiliated Hospital of Guangxi Medical University (Hospital no.1). The data set consists of 36 normal cases, 39 mild fibrosis cases, 38 severe fibrosis cases and 36 typical cirrhosis cases. The imaging protocol is as follows: quad-phase scans are made at 120kV tube voltage, 250mA tube current; the image size is 512×512 pixels. From February 2011 to March 2012, a total of 216 patients had an abdominal examination performed using a 3-T superconducting MR scanner (Intera Achieva Quasar Dual; Philips Medical Systems, Netherlands) with a six-channel torso array coil. All cases have been verified by needle biopsies as the gold standard of our experiment. Surgical specimens were retrospectively examined by a pathologist who was blinded to patient histories and radiology and surgery reports. The patients' fibrosis stages are evaluated in accordance with the Chinese Viral Hepatitis Prevention and Treatment Plan (VHPTA) [7], ranging from 0(no fibrosis) to 5(cirrhosis). In this paper, the entire dataset is divided into two groups: non-cirrhosis (containing normal cases S0 and mild liver fibrosis cases S1, S2); and cirrhosis (containing severe liver

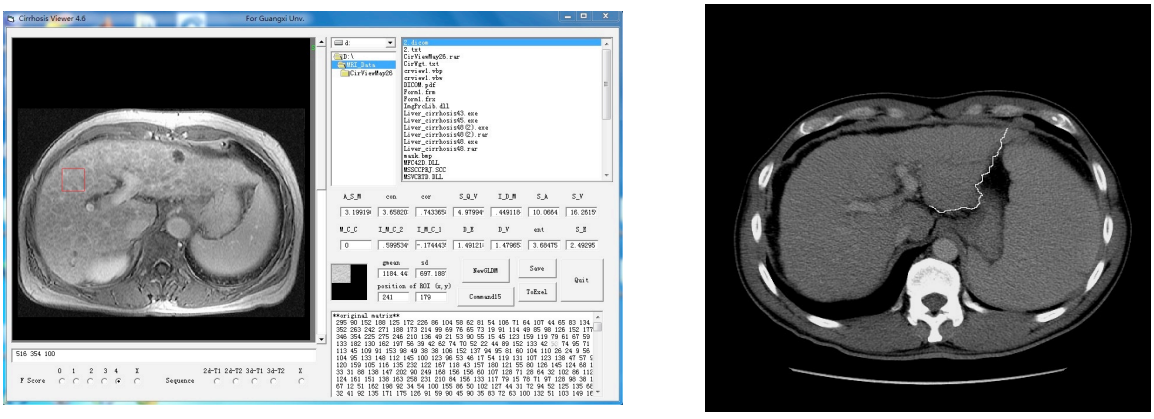
fibrosis cases S3, S4 and typical cirrhosis cases S5). Table 1 gives the detailed distribution of the datasets. This study was approved by the institutional review board at both hospitals and informed consent was obtained from all patients.

Table 1 Patients Statistic with five fibrous stages according to VHPTA System

Stage of fibrosis	Score	Description of hepatic fibrosis degree	Number of MRI cases	Number of CT cases	Groups
Normal	0	No fibrosis	45	36	Non-cirrhosis (Group 0)
Mild fibrosis	1	Fibrous expansion of some portal areas, with or without short fibrous septae.	34	26	
	2	Fibrous expansion of most areas and bridging fibrosis appeared	33	13	
Severe fibrosis	3	Most fibrous septum and lobular structure disorder	26	19	Cirrhosis (Group 1)
	4	Early cirrhosis, diffuse fiber hyperplasia	38	19	
Cirrhosis	5	Cirrhosis	40	36	

2.2 Determination of ROIs and profiles on liver for feature extraction

In this experiment, 10 ROIs in liver were manually picked in a disperse fashion from each sequenced image by a radiologist using our graphical user interface (GUI). The ROIs are shown as red squares in Fig.1a. For the selection process of liver profiles, as the left liver lobe is more likely to be deformed and the formation of nodes is easier to be observed and identified compared with right lobe, we choose the partial profile in three liver areas, namely inner lower segment of left lobe of liver, inferior segment of left lateral liver and superior segment of left lateral liver, as shown in fig. 1b. Although we have automatic program to extract liver contour, we use liver profile drawn manually by radiologists as the gold standard of this study to ensure the comparing accuracy of two methods.



(a) (b)

Fig.1 (a) The graphical user interface of our software designed for selecting ROIs by radiologists and automatic feature extraction by CAD system. The red square of ROI contains the texture patterns of liver cirrhosis. (b) liver profile drawn on the segment of left lateral liver.

2.3 Calculation of texture and shape features

The texture features are obtained by gray level co-occurrence matrix (GLCM) method. As the DICOM images have 12-bit gray levels, which is too large for establishing co-occurrence matrices, the gray levels of an image are reduced to 4-bit, which is accurate enough for this study [17], and the matrix element in row *i* and column *j* are denoted:

$$P(i,j,\delta,\theta) = \{[(x,y),(x+\Delta x, y+\Delta y)]| \\ f(x,y) = i, f(x+\Delta x, y+\Delta y) = j; x, y = 0,1,..,N-1\} \quad (2\text{-}1)$$

Where $i,j=16;\ \delta = 1;\ \theta=0^o,45^o,90^o,135^o$.

13 texture features are calculated by the method introduced by Haralick [8] based on the co-occurrence matrix: 1. Angular second moment (ASM) 2. Contrast (CNT) 3. Correlation (COR) 4. Entropy (ENT) 5. Sum of squares: Variance (SQV) 6. Inverse Difference Moment (IDM) 7. Sum Average 8. Sum Variance 9. Sum Entropy 10. Difference Variance 11. Difference Entropy 12. Information Measures1 (IMC1) 13. Information Measures2 (IMC2). In our experiment, two features based on first order statistics of the image: 14. Mean gray value (MGV) and 15. Standard deviation (SD) are also used in this study, thus there are total 15 texture features used in a ROI.

At present, there is no unified shape model and feature selection method due to the diversity of different morphological features and their combination probability. In order to find an effective and accurate method to diagnose liver fibrosis, we propose a quantitative approach for selecting effective shape features from the roughness parameters of micro surface that is widely used in mechanical engineering. Shape features are intuitive and can be observed and detected easily, and thus it plays a profound role in image analysis and application. Ten most representative features [6] are chosen as the profile features of liver based on the roughness of micro-mechanical surface in mechanical engineering, which are 1) roughness average (*Ra*); 2) root mean square roughness (*Rq*); 3) maximum roughness depth (*Rmax*); 4) maximum profile peak height (*Rm*); 5) maximum profile valley depth (*Rp*); 6) mean spacing of the profile irregularities (*Sm*); 7) mean spacing of local peaks of the profile (*S*); 8) ten-point average roughness (*Rz*); 9) peak density of the outline (*D*); 10) profile bearing length ratio (*Tp*).

2.5 Support Vector machine (SVM) with Leave- one-out cross-validation (LOOCV) method

SVM originated on the basis of statistics by *Vapnik et al.* [9] is a current general learning method. The discrimination function for linear separable problem is:

$$f(X) = \sum_{i=1}^{N} y_i a_i * k(X_i, X) + b* \quad (2\text{-}2)$$

where N is the number of support vector, X_i, y_i are the labels of corresponding support vectors, and α_i* and b* are parameter learned from training samples. The kernel *k(Xi,X)* is significant for determining the behaviour of the classifier. In this study, the Radial Basis Function (RBF) kernel function is used:

$$k(X,Y) = \exp\{|X-Y|^2 / 2\sigma^2\} \quad (2\text{-}3)$$

The |x-y| represents the distance between two vectors, where σ is a constant 1. The input vectors to SVM are 15 texture features or 10 shape features in this study while the output is probability for the presence of cirrhosis.

The Leave-one-out (LOO) [10] method is to take out a single observation from the original datasets containing M samples as the validation data, and the remaining M-1 samples as the training data to build a classification model of SVM. After validating this model, the sample will be moved back into the training data set and another sample is

selected. The procedure is repeated such that each observation in the sample is used once as the validation data. After looping a total of M times of training and testing on M samples, all of the cases in one dataset are validated by the SVM model. Among the testing results, the ratio between the number of correctly classified cases as cirrhosis and total number of cirrhosis cases is defined as true positive (TP) while the ratio between the number of correctly classified cases as non-cirrhosis and total non-cirrhosis cases is defined as true negative (TN). The average of TP and TN is defined as the accuracy rate (AR).

Feature selection includes two sections: optimizing number of input vectors and optimizing weight of each shape features. Maximum average accuracy and the choice of corresponding features can be obtained on the basis of the AR value. The experiment flow chart is shown in figure 2.

In this paper, the data samples are firstly divided into two categories in the SVM classification model: the mild liver fibrosis cases S0, S1, S2 as the negative samples group 0, while S3, S4, S5 severe liver fibrosis cases and typical cirrhosis cases as the positive samples group 1. Among sequenced datasets, each possible combination of texture/shape features is performed by the SVM model with LOOCV method to compare their accuracy. After the optimal number of features is determined, performances between different grades are further investigated.

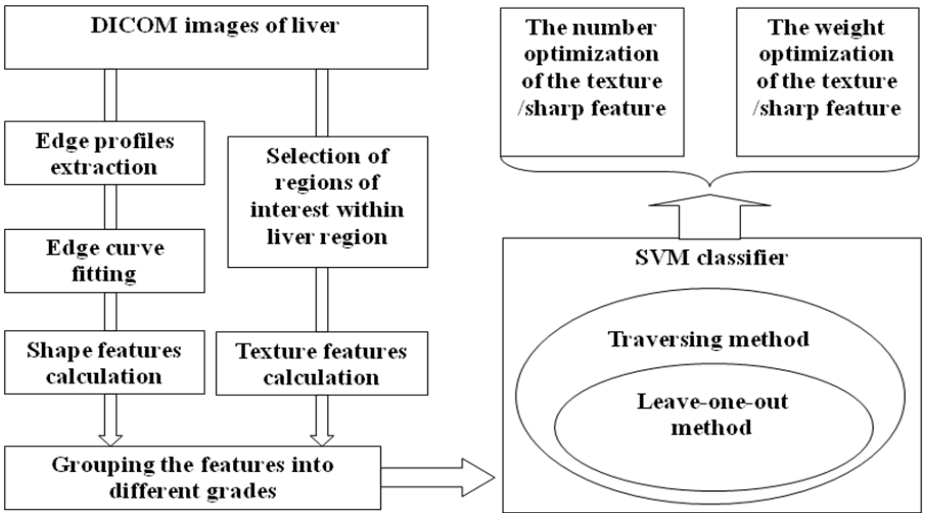


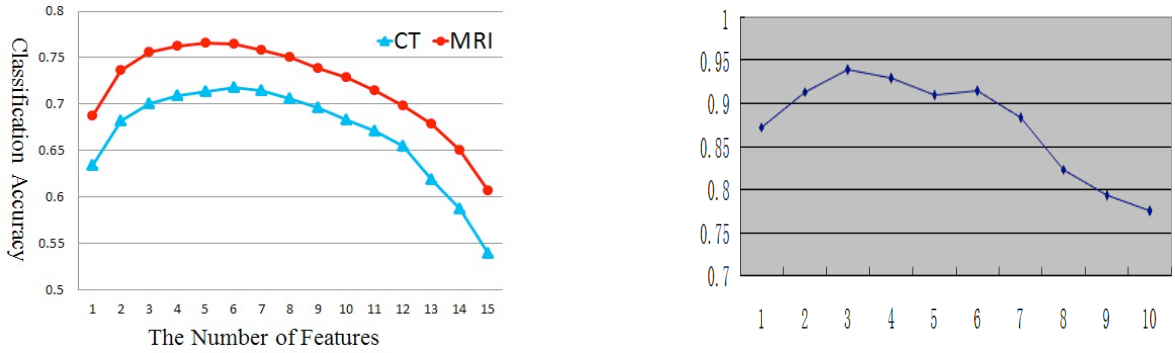
Fig. 2 Flow chart of feature selection

3. EXPERIMENTAL RESULTS

3.1 Statistical analysis of input number of features

According to the results obtained with different combination of number of features, the AR rate has different performance with the number of the texture/shape features using as an input set as shown in Fig. 3. SVM classifier obtains better classification when selecting 4-7 shape features while 2-6 shape features as inputs. If the number of features is greater than 7, the classification results become worse with an increase in the input features; on the contrary, if

the number of selected shape feature is too small, the insufficient information will make the classifier hard to classify the samples with less effective inputs. Therefore too many or too few features cannot effectively diagnose liver fibrosis. The accurate rate value of all 10/15 types number of features is 66.83% by texture, while 85.74% by shape features, respectively.



(a) texture features (b) shape features

Fig. 3 Accuracy curve by the selection of different number of texture and shape features

3.2 Feature importance analysis

The weight of a feature refers to the degree of influence that a feature affects the accuracy of classification results. Figure 4 indicates the top 7 texture features in the discrimination power histogram: mean gray value (MGV), standard deviation (SD), inverse difference moment (IDM), difference variance (DV), entropy (ENT), contrast (CON), and sum of squares variance (SQV). Among 10 shape features, *Rmax*, *Rp*, *Rm*, *Sm*, *S*, have larger weights than the other five features, meaning their larger contribution to the classification in this experiment. Root mean square deviation of the profile & maximum height of the profile irregularities are in top 2 ranks in shape features.

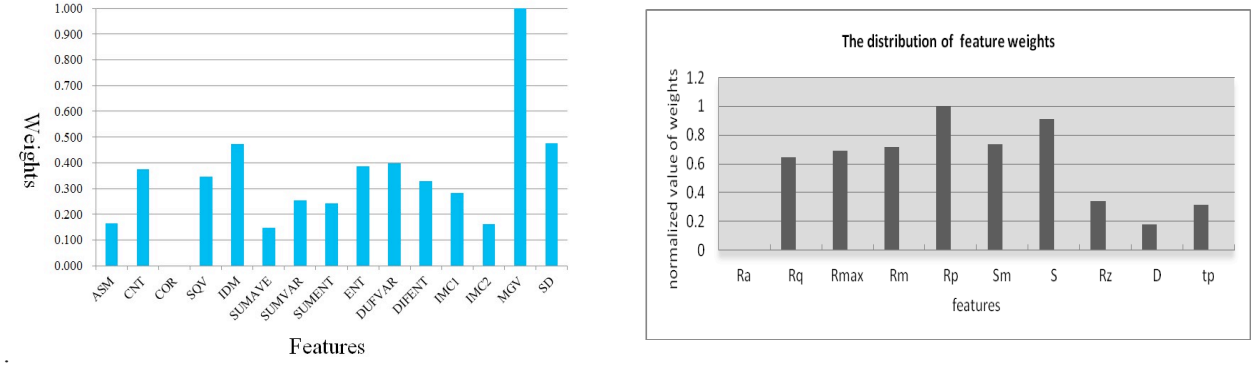


Fig. 4 The distribution of normalized value of weights for 15 texture and 10 shape features

3.4 Classification of fibrosis in different groups

Our previous studies are based on the classification results between two groups: non-cirrhosis or cirrhosis. As we known, liver fibrosis has 5 grades that can be divided into four groups: Normal (S0), Mild fibrosis (S1S2), Severe fibrosis (S3S4) and Cirrhosis (CIR), as illustrated in table 1. Datasets from different groups are picked up into a pair to calculated a AR value, Table 2 shows AR values in different groups between texture and shape features classification. We should notice that shape feature has a relative high AR than texture features (both in minimum, maximum or average), which implies shape features have better differentiation ability than texture features in diagnosis of fibrosis.

Table 2 Comparison of AR values in different groups between texture and shape features classification

Groups	texture features			shape features		
	minimum	average	maximum	minimum	average	maximum
S0_S1S2	0	39.29%	78.57%	28.57%	60.72%	92.86%
S0_S3S4	21.43%	56.14%	92.86%	42.86%	71.43%	100.00%
S0_CIR	14.29%	57.15%	100.00%	52.87%	76.44%	100.00%
S1S2_S3S4	7.14%	35.72%	64.29%	35.71%	67.86%	100.00%
S1S2_CIR	14.29%	50%	85.71%	50.00%	71.43%	92.86%
S3S4_CIR	0	32.15%	64.29%	7.14%	46.43%	85.71%
S0S1S2_S3S4CIR	40.00%	61.67%	83.33%	44.36%	67.18%	90.00%

3.3 AR analysis by combination of texture and shape features

Since we already have the optimal number of features and most important texture and shape features at the above sections, it should be a challenge to build a better system by combining 7 most effective texture and shape features: *IDM, ENT, MGV* in texture and *Rm, Rp, Sm, S* in shape. According to the group test in table 2, the optimal number of features can be concluded in table 3. The AR values are all above 0.7, indicating a better performance by mixed features than individual texture features.

Table 3 selection of optimal number of mixed features by texture and shape

number of features	1	2	3	4	5	6	7
Maximum AR	0.845	0.893	0.881	0.833	0.774	0.714	0.701

Table 4 demonstrates the weight value of 7 most effective texture and shape features, from which we can see the weight of shape features are generally larger than that of texture features indicating the shape features are more effective

than texture features.

Table 4 Weight Value of 7 most effective texture and shape features

Feature	Rm	Rp	Sm	S	IDM	ENT	MGV
Weight	0.26	0.97	0.621	0.805	0.224	0.257	0.497
Normalized Weight	0.048	1	0.532	0.779	0	0.044	0.366

4. DISSCUSSION AND CONCLUSION

This study investigates the role of shape features vs. texture features in the classification of hepatic fibrosis by selecting the optimal parameters for building a better Computer-aided Diagnosis system. 15 texture and 10 surface shape features are extracted from liver region. Each combination of these features is selected as input subsets to be checked by using a support vector machine (SVM) with the leave-one-case-out method to differentiate fibrosis into normal or abnormal. The result shows that the optimal number of features is ranged from 2 to 6 among all 10 features; The experiment result indicates that the accuracy rate of using shape features for analysis is considerably higher than using texture features and texture feature of CT image is not recommended to use with shape feature for interpretation of cirrhosis.

ACKNOWLEDGMENT

This work was supported in part by two research support from the National Natural Science Foundation of China (No. 61262027 & 45627390), and in part by a research foundation project of the Guangxi Key Laboratory of Multimedia Communications and Network Technology (Cultivating Base). The author Xuejun Zhang gratefully acknowledges financial support from China Scholarship Council.

REFERENCES

[1] Gao Yongsheng, and Yang Jie, “A Computer Automatic Identification Method of Medical Texture Image”, Journal of Zhejiang University School of Medicine, 1992(03), pp. 101-105.

[2] Y. Guo et al., “On Gabor Wavelet Texture Feature of Liver Image”, Computer Applications and Software, 2008(11): pp. 44-45, 79.

[3] Krusinska, E., et al., Influence of "outliers" on the association between laboratory data and histopathological findings in liver biopsy. Methods of information in medicine, 1993. 32(5): pp. 388-95.

[4] Hannu Huhdanpaa , Darryl Hwang, Steven Cen, Brian Quinn, Megha Nayyar, Xuejun Zhang, Frank Chen, Bhushan Desai, Gangning Liang, Inderbir Gill, Vinay Duddalwar, “CT prediction of the Fuhrman grade of clear cell renal cell carcinoma (RCC): towards the development of computer-assisted diagnostic method”, Abdominal Imaging, DOI: 10.1007/s00261-015-0531-8, Available online 25 August 2015.

[5] Zhang Xuejun, Xin Gao, Brent J. Liu, Wen Yan, Long Liling, Huang Yuhong, and Hiroshi Fujita, "Effective staging of fibrosis by the selected texture features of liver: Which one is better, CT or MR imaging?", Computerized Medical Imaging and Graphics, Volume 46, pp. 227–236, 2015.

[6] Xuejun Zhang, Bin Zhou, Kevin Ma, Xianghe Qu, Xiaomin Tan, Xin Gao, Wen Yan, Long Liling, and Hiroshi Fujita, Selection of optimal shape features for staging Hepatic fibrosis on CT image, Journal of Medical Imaging and Health Informatics, Vol.5, No.8, pp.1926-1930, 2015.

[7] Viral Hepatitis Prevention and Treatment Plan: Chinese Society of Hepatology, Chin. J. Infect. Dis., Vol. 19(1), pp.56-62, 2001.

[8] R. Haralick, ”Statistical and Structural Approaches to Texture” proceedings of the IEEE, vol.67, no.5, pp.786-804, May 1979.

[9] S. Yang, “Pattern Recognition and Intelligent Computing - Matlab technology”, Beijing, Publishing House of Electronics Industry, pp.113-137, July 2009.

[10] R. Schapire, “Leave One Out Error, Stability, and Generalization of Voting Combinations of Classifiers” Machine Learning, 55, 71–97, 2004

Development of a web based informatics system utilizing quantitative imaging features for predicting outcomes in stroke rehabilitation clinical trials

Ximing Wang, Ji Hoon Park, Jeffrey Tse, Brent Liu

Image Processing and Informatics Lab, Viterbi School of Engineering,
University of Southern California

ABSTRACT

Imaging biomarkers have been used widely in current stroke rehabilitation clinical trials as a method for tracking and evaluation of treatment effects. Previously, we presented an electronic patient record system to support imaging based stroke rehabilitation clinical trials. However, normally, the data management system supporting the stroke rehabilitation clinical trials are solely data entry systems with no analysis tools for modeling and predicting outcomes. In order to facilitate the data analysis and provide a prediction based on the model, we aim to integrate a generalized linear mixed effects model (GLMM) module to automatically investigate potential significance and coefficients based on features extracted from textual database and imaging biomarkers, and finally predict treatment outcomes based on fitted model and data from new subjects.

Based on a previous presented web-based system, a GLMM module is integrated. The user is able to select potential textual data fields and imaging biomarkers, such as lesion volumes and locations, as predictors and response variables. The module will automatically fit the data and display the significance of each variable. The p-value of the each dependent variable and fitted curves will also be presented on the web. The module will be evaluated by 40 subjects in a phase I stroke rehabilitation clinical trial.

Keywords: electronic Patient Record (ePR), Clinical Service, generalized linear mixed effects model

1. INTRODUCTION

With the development of technology, imaging biomarkers have been used widely in current stroke rehabilitation clinical trials. By computer aided detections, features can be extracted from medical images. These features, such as stroke lesion size, lesion location and rate of growth, are used as imaging biomarkers. Imaging biomarkers may have relevance with the medical outcomes and sometimes are been recorded to patient's medical reports.

Imaging biomarkers are mostly used as surrogate endpoints in clinical trials. Traditional measurement of clinical trials, such as morbidity and mortality, are sometimes very subjective. Another drawback of traditional endpoints is that measuring some of them is usually time-consuming. For example, mortality usually cost years or decades for measurement. In contrast, Imaging biomarkers are easy to measure, objective, and can be quantitative. Quantitative imaging biomarkers can be used to detect subtle changes, which are hard to determine qualitatively, and can be used for analysis like computational models. In stroke rehabilitation clinical trials, imaging biomarkers are commonly used as a method for tracking and evaluation of treatment effects.

Imaging biomarkers brings massive amounts of data to clinical trials. In the age of "big data", imaging data, as well as other multimedia data, are collected during the clinical trials. In stroke rehabilitation, quantitative imaging biomarkers can be used in statistical models for data evaluation. Computational models can be used to fit the data and discover correlations behind the big data. With a fitted data model, the model can be used to predict outcomes for future subjects. Previously, we presented an electronic patient record system to support imaging based stroke rehabilitation clinical trials.[1] However, normally, the data management system supporting the stroke rehabilitation clinical trials are solely data entry systems with no analysis tools for modeling and predicting outcomes. Ghosh et al. [2] presents that one of the current major challenges for informatics in clinical trial is to integrate a data analysis tool with the data entry system, and Sung et al[3] also mentions that an integrated platform is one of the major challenges for future informatics in clinical trials. In order to facilitate the data analysis, provide a prediction tool, we aim to develop an integrated system with a generalized linear mixed effects model (GLMM) module as a knowledge discovery tool to automatically investigate potential significance and coefficients based on features extracted from textual database and imaging biomarkers. Based on a fitted model, the module can predict outcomes for future subjects.

Once a model is fitted, the system can be used as an ePR system for real patients in practice. The knowledge discovery module is able to predict the treatment outcomes based on new patient's information and various dosages. Based on the predicted results, clinicians can choose the best treatment plan for the patient.

2. METHODS

2.1 Introduction to Generalized linear mixed model

Generalized linear mixed model (GLMM) is a powerful model in statistics and has become popular for data analysis in clinical trials. Based on a systematic review conducted by Casal et al. [4], over 400 articles published during 2000-2012 include "Generalized linear mixed model" in topics. Over 50% of the articles they chose are declared to longitudinal studies. The GLMM is an extension from linear regression by adding the random effects to usual fixed effects. It also extends the normal distribution requirements of regressors to non-normal distribution. Compare to the classic linear model, the GLMM allow for the accommodation of non-continuous, non-normal distributions, and non-linear relations between covariates and responses. In application, GLMMs are also widely used in diverse fields beyond medicine, including biology [5], psychology [6], and ecology [7]. In summary, GLMM is chosen for our knowledge discovery tool as it is able to evaluate the significance of each potential independent variable, estimates the coefficients of the model to fit the data, and finally predict outcomes based on fitted model, and widely used in the past decade.

2.2 workflow of the knowledge discovery module

The knowledge discovery module is developed based on a previously presented electronic patient (ePR) record system [1].This system is developed for a stroke rehabilitation clinical trial, "Optimizing the Dose of Rehabilitation after Stroke", which is a phase I clinical trial which aims to optimize the dosage of treatment for paresis recovery after stroke. In this clinical trial, the correlations between imaging features such as lesion volume, lesion location and treatment outcomes are to be investigated. The system allows user to measure the lesion volume and identify the vascular territory location of the lesion for each subject. This system allows radiologists to fill out electronic case report forms (eCRF) while reading images. By using the template and measurement tool provided by the system, researchers are able to collect imaging biomarkers like lesion location and lesion size. These imaging features are recorded through case report forms and stored into the database.

Based on the existing system, we developed the knowledge discovery module with GLMM model. The module allows the user to choose the regressors from the database, including imaging features and other clinical measurements, define the parameters and then the system will show fitting model and curves on the webpage. Based on the fitting coefficients, once a new subject is enrolled in the clinical trial, the system is able to show the predicted outcome variables based on the subjects information. Figure 1 illustrates how to use the knowledge discovery tool to evaluate and predict outcomes. There are seven main steps in the workflow.

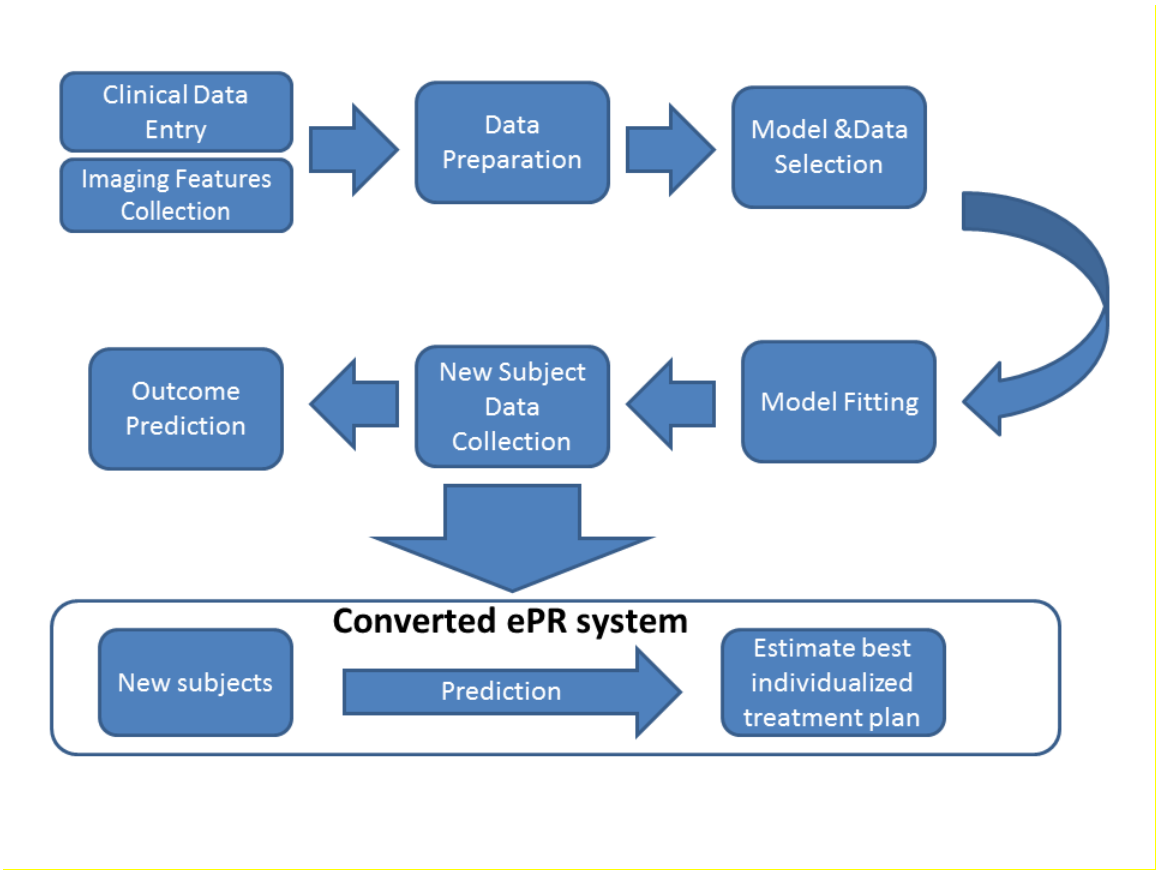


Figure 1. Workflow of the knowledge discovery tool

1. Data collection. During the clinical trials, data entry system allows users to enter all clinical data, just like classic data entry system. All the data will be stored in the relational database. Imaging features are collected by eCRF and stored into the database as well.
2. Data Preparation. In this step, the system provides a command window for user to enter SQL commands. User can select candidate variables into a view in the database for future processing. This step combines all data by one or several keys.
3. The Module provides a webpage for user to select potential data fields from the database as predictor variables and response variables. For example, in a stroke rehabilitation clinical trial, to investigate if the lesion volume affects the treatment effects, the user should select the lesion volume data field from the database as a feature, and use the treatment evaluation score as the observed outcome variables. The user is also able to set the parameters for the model, such as link function, distribution etc.
4. Once the user selects the model and data in step 3, the module will automatically fit the data and show the p-value of the each dependent variable and fitted coefficients.
5. After the model and predictor variables are selected in module, user can continue to use the data entry system to collect data.
6. Once all predictors data are collected and prepared (step 2), the module is able to show the predicted response variables and predicted confidential intervals. In stroke

rehabilitation clinical trials, the model is usually used to show the predicted evaluation scores with confidential intervals.

7. The clinical trials can be used as an ePR system for real patients. Once a new post-stroke patient was prescribed to take the physical therapy for motor recovery, the model developed in the step 6 can predict the evaluation scores based on patent's information and the treatment dosage. Based on the prediction score, the treatment dosage can be optimized. Thus a better individualized treatment plan can be estimated. The patient's feedback after the treatment plan can also be stored into the data pool for optimization of the model in future.

3. RESULTS

A ePR system for stroke rehabilitation clinical trials has been developed in linux 14.04, apache 2.4, php 7 and javascript 1.2. jQuery and HTML 5 is used for front-end programming[1]. MySQL is used for relational database management system. NanoDICOM is used for server-side DICOM processing. The knowledge discovery module is developed in a web application framework for R, the Shiny by Rstudio.[8] Twitter-bootstrap is used for front-end CSS markup.

The core of the knowledge discovery module, the GLMM model fitting algorithm, is developed based on R-package "lme4" in server side, and a graphical user interface is developed with Shiny framework in the client side. The module has been fully integrated with the existing system. An interface between the knowledge module and the central database has been built, allowing directly access of database from the knowledge discovery module. Due to the confidentiality of the stroke rehabilitation clinical trial, all data presented here are synthetic data. Some of the screenshots are mockup work and are used for demonstration purpose only.

In step 1, the system allows user to collect clinical data like classic data entry system. The system also provides a side-by-side integrated viewer with eCRF. (See figure 2.) In this viewer, a templates tool allows user to define the lesion location, and a lesion quantification tool allows user to measure the lesion volume. These results will be stored into the database by the eCRF.

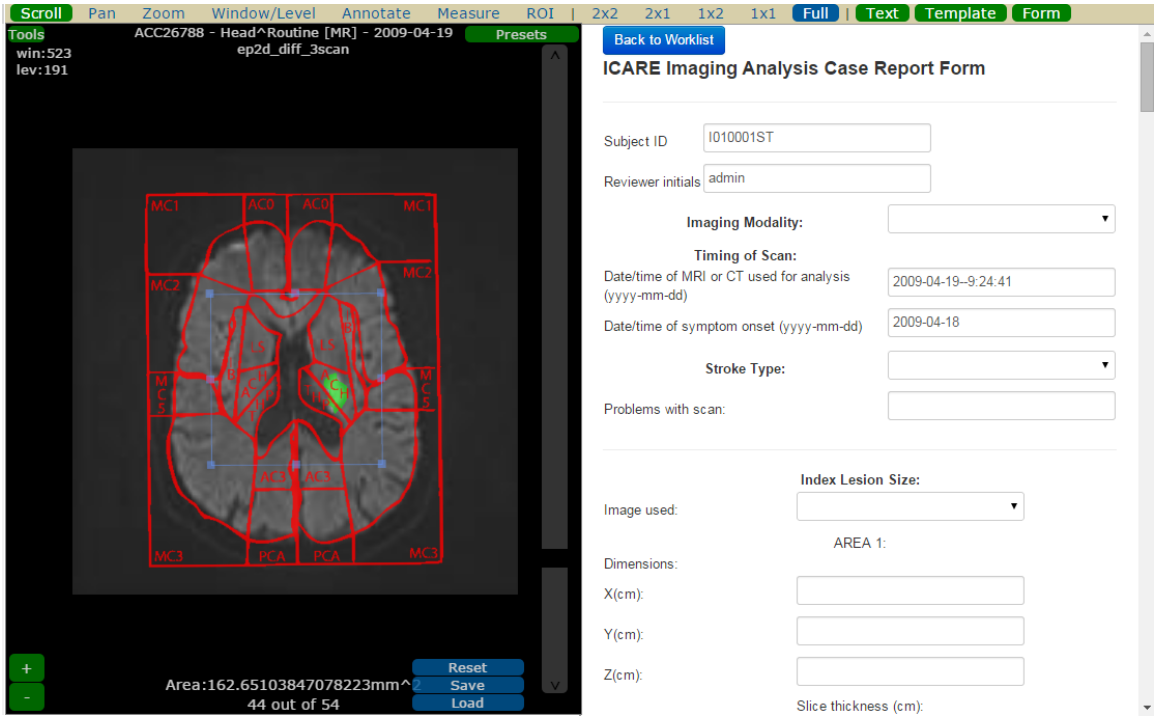


Figure 2. Imaging features Collection. (Cited from "Characterizing stroke lesions using digital templates and lesion quantification tools in a web-based imaging informatics system for a large-scale stroke rehabilitation clinical trial", Proceedings of SPIE Vol. 9418, 94180E (2015).) User is able to use tools to collect imaging features like lesion location, and lesion volume. These data will be recorded into eCRF on the right side of the viewer.

In step 2, user need to use SQL command to join all the features data by keys such as patient ID and visit number. A viewer for data mining will be created in the database.

In step 3, the knowledge discovery module allows user select the model parameters and the variables.(see figure 3) The module allows user to select the distribution of the data, choose suitable link function, and finally the user can specify potential model formula. Like other GLMM software, the formula needs to be specified in Wilkinson notation. On the right panel of the figure 3, user is able to select candidate variables. These variables must be in the view created in step 2. The variable can be categorical or numeric. The user needs to assign each variable a name for the specification of model formula.

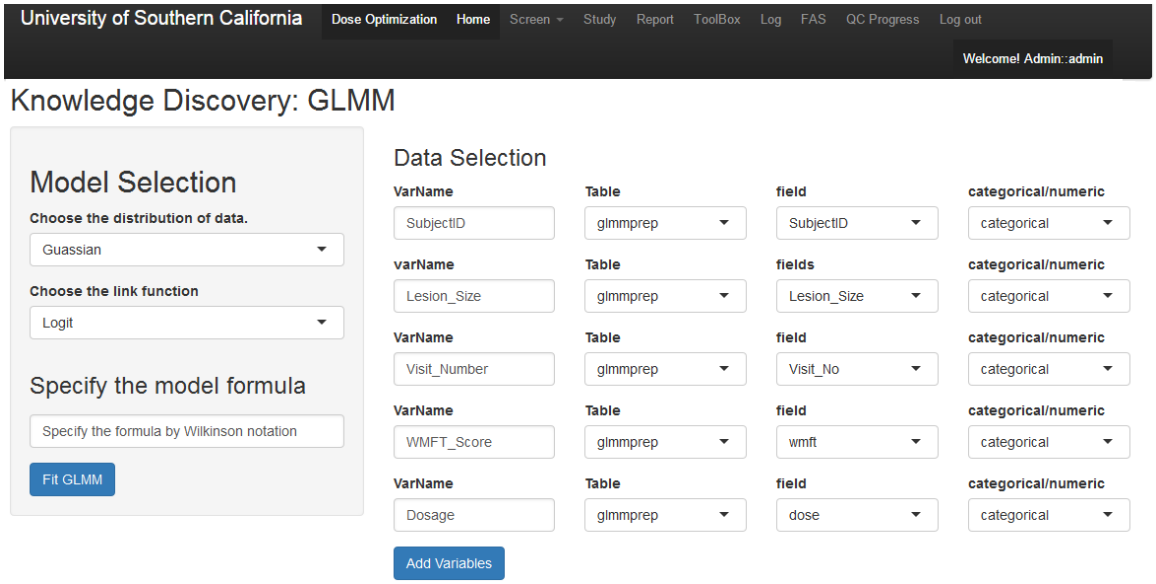


Figure 3. Model and data selection. Left panel allows user to select model parameters, including distribution family, link function and specify the formula. The right panel allows user to select candidate variables.

After the user selects the data and model parameters, the knowledge discovery module is able to automatically fit the data and shows significance of each variable. Figure 4 demonstrates the fitting results of the model. Based on the results, the user can change the model or variables to find the best model.

At last, figure 5 illustrates the prediction outcomes for new subjects. If a new subject is enrolled, the system is able to use the predictor variables and the model defined in step 4 to predict the treatment outcome, which is defined by the score of an evaluation test. Figure 5 shows the prediction based on the model and new subject's information (predictor variables). The system estimates the treatment outcome after each visit, with confidential intervals. The figure on the right panels shows the mean and error bars of the evaluation score.

Figure 5 also shows the prediction result of a new patient in the converted ePR system (step 7 in figure 7). Based on the patient's baseline evaluation score, imaging features, and preliminary treatment effects, the fitted model gives prediction of treatment outcomes depend on the treatment dosage. The clinician can use this outcome to choose the best dosage for patient.

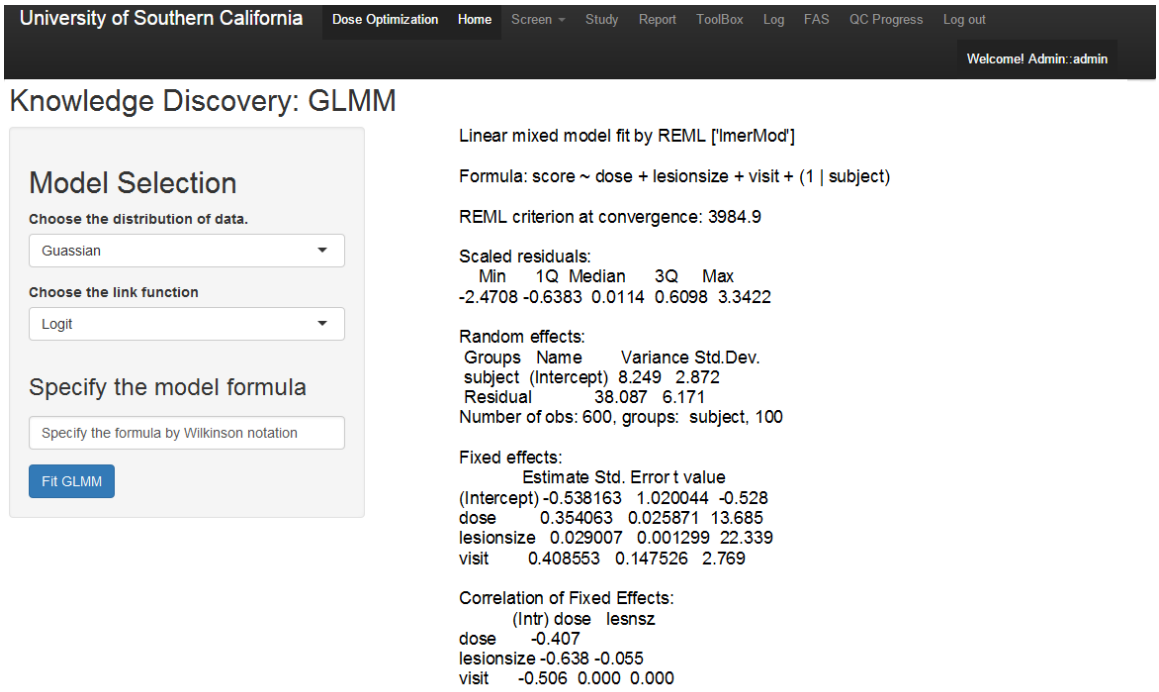


Figure 4. Model fitting results. The right panel shows the fitting results and coefficients of each variable. The user can update parameters or variables to get better fitting results.

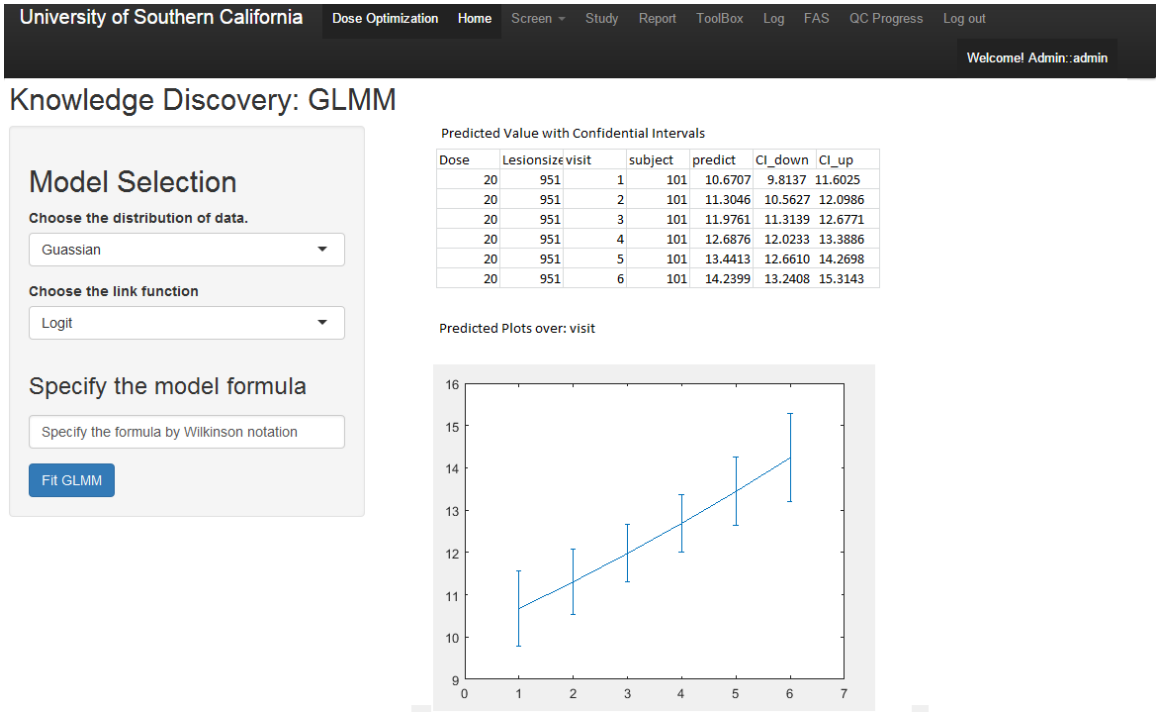


Figure 5. Treatment outcome prediction. The top half of the right panel shows predicted value with confidential intervals for 6 treatments. The bottom half of the right panel shows the plot of predicted value with error bars.

4. DISCUSSION and FUTURE WORK

In classic systems for imaging based clinical trials, imaging data are clinical data are not integrated in the data entry system. The classic way for processing and analyzing data is to transfer data through spreadsheet, and process the data by third-party software such as Matlab. However, Ghosh et al[2] and Sung et al[3] proposed that the integrated platform with data analysis tool will be the future of the informatics systems in clinical trials. The system presented in this paper shows the effort as a tentative solution for this integrated platform for imaging based clinical trials by integrating imaging data viewer, clinical data entry system and data analysis tools.

With the integration of the knowledge discovery module, the system is able to evaluate the treatment outcome based on a variety of measurements including quantitative imaging features. After fitting the model, the knowledge module is able to make predictions of treatment outcomes based new subjects' information. Moreover, the module can be adapted as a decision support tool for an ePR system of future stroke rehabilitation clinical trials. Once an accurate GLMM model is identified from the clinical trial, the module is able to predict outcomes for new patients based on their conditions and used as a treatment planning tool for optimization of the treatment plans.

The stroke rehabilitation clinical trial "Dose optimization after stroke" has enrolled over 40 subjects. The system will be evaluated by the 40 subjects and the beta version will be developed based on the evaluation feedbacks.

5. CONCLUSION

A generalized linear mixed model (GLMM) module is developed and integrated with an electronic patient record system to support a stroke rehabilitation clinical trial. This module can facilitate preliminary evaluation of correlation of clinical measurements including quantitative imaging features, and predict outcomes for new subjects based on an identified model. This module can also be adapted and used as a treatment planning tool for ePR system. The system is currently being used in a stroke rehabilitation clinical trial and the module will be evaluated by 40 subjects in the clinical trial.

Acknowledgement This work is supported by NIH/NICHD R01HD065438

6. REFERENCE

[1] Ximing Wang, Matthew Edwardson, Alexander Dromerick, et al., "Characterizing stroke lesions using digital templates and lesion quantification tools in a web-based imaging informatics system for a large-scale stroke rehabilitation clinical trial", Proceedings of SPIE Vol. 9418, 94180E (2015)

[2]Ghosh S, Matsuoka Y, Asai Y, Hsin K, Kitano H: Software for systems biology: from tools to integrated platforms. Nat Rev Genet 12:821–32, 2011

[3] Sung NS, Crowley WF, Genel M, Salber P, Sandy L, Sherwood LM, et al: Central challenges facing the national clinical research enterprise. JAMA 289(10):1278–87, 2003

[4] Casals, Mart ínez, Montserrat Girabent-Farr és, and Josep L. Carrasco. "Methodological quality and reporting of generalized linear mixed models in clinical medicine (2000–2012): a systematic review." PloS one 9.11 (2014): e112653.

[5] Thiele, Jan, and Bo Markussen. "Potential of GLMM in modelling invasive spread." CAB Reviews 7, no. 016 (2012): 1-10.

[6] Moscatelli, Alessandro, Maura Mezzetti, and Francesco Lacquaniti. "Modeling psychophysical data at the population-level: the generalized linear mixed model." Journal of vision 12, no. 11 (2012): 26-26.

[7] Bolker, Benjamin M., Mollie E. Brooks, Connie J. Clark, Shane W. Geange, John R. Poulsen, M. Henry H. Stevens, and Jada-Simone S. White. "Generalized linear mixed models: a practical guide for ecology and evolution." Trends in ecology & evolution 24, no. 3 (2009): 127-135.

[8] Shiny, a web framework for R. http://shiny.rstudio.com/