NEXT GENERATION REPORTING AND DIAGNOSTIC TOOLS FOR HEALTHCARE AND BIOMEDICAL APPLICATIONS

SHARMILI ROY
(MSc(Engg.), Indian Institute of Science, 2006)

A DISSERTATION SUBMITTED FOR THE DEGREE OF

Doctor of Philosophy

in

SCHOOL OF COMPUTING

NATIONAL UNIVERSITY OF SINGAPORE

SINGAPORE, 2014

Declaration

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Signature: _____ **Date:** _____

Acknowledgements

I would like to take this opportunity to express my gratitude to all who have contributed towards the completion of this thesis.

First and foremost, I am extremely thankful to my advisor, Dr. Michael S. Brown. Dr. Brown has been a constant source of ideas for all my projects. I have gained a lot from his clarity of thought, his eye for the minutest details and his ability to modularize large problems into smaller problems. He has always encouraged me to attend conferences and summer schools and has always taken initiatives to form research collaborations outside the National University of Singapore (NUS). I have never had a communication gap with him which I think has been critical in making my PhD a joyful experience.

I would like to extend my sincere gratitude to Dr. George L. Shih for giving his insightful ideas and continuous feedback on our project on radiological reporting despite his busy schedule.

My earnest thanks are due to Dr. Asanobu Kitamoto for giving me the opportunity to work under him in Japan and exposing me to new problems in the field of biomedical image analysis.

I am indebted to Dr. Liu Jimin and Dr. Yanling Chi for accepting me as an intern at the Agency for Science, Technology and Research and offering me a chance to work on healthcare problems in Singapore.

At last, I would like to thank my parents for always being there for me. I owe my PhD to my husband, Anmol Sethy. He was instrumental in convincing me to apply and join the NUS PhD program. His constant encouragement has been the source of motivation behind this PhD thesis.

Abstract

Virtually all fields of healthcare and biomedical research now rely on imaging as their primary data source. Though more and more data is being generated in the imaging centers, research shows that most of this data is discarded in routine practice. Further, certain routine practices in healthcare and biomedical research, such as radiological reporting and gene-to-physiology mapping, still represent relics of the pre-digital era that underutilize the available data and today's computational technologies. The aim of this thesis is to use modern computer vision, image processing and computer graphic technologies to design reporting, analysis and diagnostic tools for healthcare and biomedical applications that not only better utilize existing, otherwise discarded, data but also uses modern techniques to enhance some of the archaic methodologies.

More specifically, using discarded radiological annotations, we aim to enhance traditional radiological reporting by proposing animated visual reports that highlight and position clinical findings in a three-dimensional volumetric context as opposed to the historic text-based white paper reports. In a second application on diagnosis of hepatic tumors, we employ already diagnosed cases of liver tumors to propose a fast content-based image retrieval system that assists experts in tumor diagnosis by retrieving similar confirmed cases from a database based on visual similarity of tumor images. As a third application we target the low efficiency age-old histological methodology of gene-to-physiology mapping and propose a defect detection framework that automatically identifies physiological defects in micro-CT images of transgenic mice.

Contents

List of Figures

List of Tables

List of Algorithms

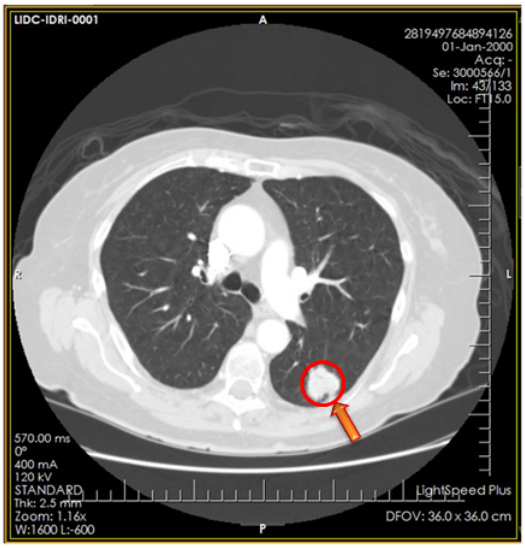
Chapter 1

Introduction

1.1 Overview

Imaging is the elementary step in virtually all fields of biomedical and healthcare research. Today, medical scans produce thousands of images and tera bytes of data for a single patient in mere seconds. The global size of data in healthcare is estimated to be 150 exabytes in 2011 and is increasing at between 1.2 and 2.4 exabytes a year (1 exabyte = 250 million DVDs of data). More data should mean that care providers – from nurses and public health officials, to specialists – have more insight into helping solve their patients' problems. Unfortunately however, research shows that healthcare providers discard 90% of the data they generate. How can this existing discarded data be utilized to design better healthcare and biomedical tools?

The second question to ponder upon is how elegantly this biomedical data is analyzed and utilized in routine practices? Biomedical image processing has been an active field of research for more than 30 years. Significant success has



(a)

Date: 13Nov 2012@0913

RADIOLOGIC EXAMINATION REPORT

Patient: XYZ

.....

HISTORY: Hemoptysis

TECHNIQUE:
CT of the chest was performed. No prior exams are available for comparison.

FINDINGS:
In the left lung, a solitary pulmonary nodule is seen. Right lung is clear.

IMPRESSION:
1. SOLITARY PULMONARY LESION IN THE LEFT LUNG.

Transcription Data/Time: 13 Nov 2012@0913

(b)

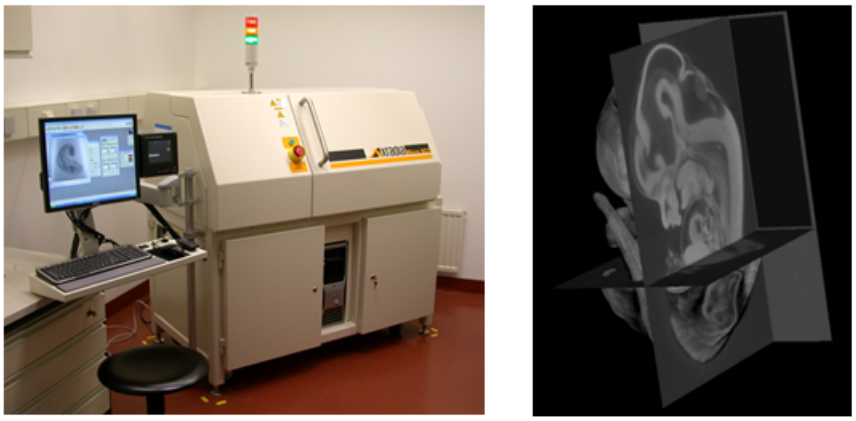
Figure 1.1: (a) An example of a radiological markup on a medical exam. (b) The corresponding text report that summarizes the radiological findings.

been achieved in solving traditional data analysis problems such as registration, segmentation and classification. Unfortunately however, only very few of these computational tools have been successful in making their way into the routine workflow. As an example let us examine the clinical routine in a typical radiology lab. The workflow involves a referring physician who requests an image exam on a patient. This exam is interpreted by a radiologist. During interpretation the radiologist may need to perform some image-based measurements or segment a region of interest. Although automated measurements and segmentation are well-addressed problems in the literature, most often than not, manual measurement and manual segmentation is what a radiologist resorts to in routine practice. The primary reason for this is the fact that these computational tools are not integrated to the medical workstations that the radiologists typically use. Often, segmentation requires radiologists to port the data into a separate segmentation capable work-

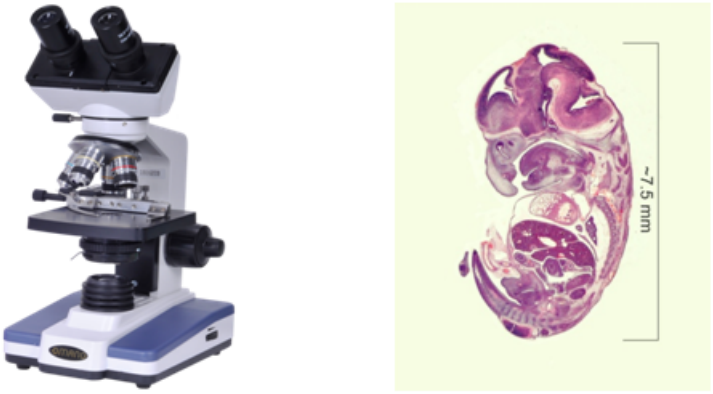
station. Further, radiologists may also need to give additional redundant input, such as seeds in case of segmentation, to initiate these tools.

After reading and interpreting the images, the radiologist prepares a text-based exam report summarizing the findings. Figure 1.1 shows an example of a radiological image markup and the corresponding text-based report. This text report is sent to the referring physician who requested the exam and his patient. The benefits of the visual markups made by radiologists on the images are well known in the medical community [Reiner and Siegel 2006] and [Fan et al. 2011]. Interestingly however, this information is lost in the workflow.

Further, once the diagnosis and treatment for this case is complete, this study is almost never used again. Today, 3D scan of a single human body produces $24,000$ slices of 512×512 pixels which is approximately 20GB of data. Most of this data, however, is discarded after the diagnosis and therapeutic success of this patient. Existing confirmed diagnosis can be utilized to aid diagnosis of new radiological cases in a content-based retrieval framework. A content-based diagnostic assistant system retrieves from a database examples and also counter-examples of confirmed cases that are similar to the case under diagnosis. Content-based retrieval systems, though known to improve diagnostic accuracy [Chi et al. 2013b], have not found their way into the routine clinical workflow. In clinical practice, diagnosis is still carried out on case-by-case basis.



(a)



(b)

Figure 1.2: (a) Each mutant mouse embryo undergoes 3D micro-CT imaging prior to sectioning. The micro-CT machine in this figure is manufactured by Xradia Inc., model MicroXCT. (b) For phenotyping, experts still rely on microscopic evaluation of the sections even though a complete 3D reconstruction of the embryo is available. The microscope in this figure is from Omano Inc., model number OM118-B4 and the mouse embryo section is available online at `http://commons.wikimedia.org/wiki/File:10dayMouseEmb.jpg` under GNU free documentation license.

Similarly, some areas of biomedical research also still rely on technologies from the pre-digital era. For example, post-completion of the human genome project, the emphasis is now on mapping each human gene to its physiological function. Genetic engineering of human is practically and ethically not possible. Hence, mouse is chosen as the model for genetic study. Large scale global efforts are underway to systematically knock out each gene in the mouse body and study the effect that it causes to the physiological makeup. In routine practice each genetically engineered mouse is first imaged using a micro-CT scanner and then sectioned into thin slices for observation under the microscope. Surprisingly, despite the availability of complete three-dimensional structure of each mouse, experts routinely rely on microscopic evaluation of the mouse sections for physiological analysis (Figure 1.2). This technique is not only antiquated but also underutilizes today's technologies and the available data.

1.2 Objectives

The objective of this thesis is two-fold. The first objective is to better utilize the existing data in routine clinical and biomedical practice. The second objective is to use modern computer vision, image processing and computer vision technologies to enhance certain archaic reporting and analysis tools used in healthcare and biomedical research. More specifically, we target three clinically important applications as described below.

Reporting still relies on age-old text-based paper reports. While interpreting a radiological exam radiologists often make visual markups on the exam images. These markups, though present in the radiological workflow, are often not used

by the physicians for diagnosis. Physicians typically rely only on the text-based reports produced by radiologists that summarize the exam findings. These reports not only lack the visual and contextual information of the pathology in the human body but are also not a convenient medium to communicate with the patients. In this thesis we aim to address this limitation in the current medical imaging and reporting workflow, in particular the out-dated reliance of physicians on text-only reports. Our goal is to use the existing radiological annotations and imaging data to design a visual reporting framework that augments radiological text reports in a format that is not only easily accessible, but is concise and visually informative to the physicians and their patients. Further, using this visual summary generation framework we want to be able to integrate computational tools, such as automated three-dimensional (3D) segmentation, into the clinical routine. This can be achieved by auto-generating seeds using existing radiological annotations. Segmentation can enhance various applications in the clinical routine some of which, including automated reporting and summary generation, are discussed in this thesis.

Diagnosis of hepatic tumors is very challenging in clinical practice and is highly experience-dependent. Research shows that diagnosis varies largely with the amount of imaging data available. Content-based access to existing tumor images has been proposed for assisting clinical decision making by re-using existing confirmed cases. The idea is to retrieve from a database of confirmed cases, instances that are similar to the one being currently diagnosed. The state-of-the art content-based retrieval systems for hepatic tumors model the tumors in two dimensions using a few slices of the 3D imaging data which is not only an incomplete representation of the tumor but also underutilizes the available data. In this thesis we quantitatively model hepatic tumors using 3D image-based spatio-temporal

features and design a fast content-based retrieval framework that outperforms existing methods. Our system is faster and incorporates a wider range of tumor pathologies than the existing systems.

Analysis of mouse images for detection of gene-induced physiological defects still largely relies on microscopic evaluation of mouse sections regardless of the availability of 3D micro-CT data. The state-of-the-art computational tools for mouse physiology analysis only offer differential volumetric analysis of various mouse organs between different transgenic mouse strains. A computational assistant for defect detection is not investigated in the literature. In this thesis we propose a generalized defect detection framework for genetically engineered prenatal mice that not only detects known defects automatically but also highlights candidate genetic defects using micro-CT images of normal and transgenic mice and computational tools like registration and deformation vectors.

1.3 Contributions

In terms of original contributions to the research community, the proposed work makes the following:

1. Proposes a visual reporting framework that enhances the text-based radiological reports by auto-generating an animated visual summary in the form of a 3D volume rendering of the exam data with the radiological markups embedded in it. The 3D volume spins to provide a comprehensive summary of the important clinical content. In a user satisfaction study conducted with physicians in Singapore it is found that the visual reports improve clarity of communication between radiologists and physicians and assists physicians

in communicating with their patients. An abstract of this work has been published in the 98th annual meeting of the Radiological Society of North America (RSNA) [Roy et al. 2011] and the full version of the work has been published in the Springer Journal of Digital Imaging (JDI) [Roy et al. 2013a]. Using the visual reporting framework we are the first to be able to integrate automatic 3D segmentation into the clinical workflow by automatically deriving segmentation seeds from radiological annotations. An abstract of this work is published in the 99th annual meeting of the RSNA [Roy et al. 2012a] and a full paper is published in the IEEE International Conference on Bioinformatics and Biomedicine Workshops (IEEE BIBMW'12) [Roy et al. 2012b].

2. Proposes 3D representation of hepatic tumors for the design of a fast content-based retrieval system for focal liver lesions. In this work we model liver tumors using 3D image-based spatio-temporal features and design a fast tumor retrieval framework that outperforms existing state-of-the-art retrieval systems which are typically based on two-dimensional (2D) features. With fast query processing and high retrieval accuracy, the proposed system has the potential to be used as an assistant to radiologists for routine hepatic tumor diagnosis. This work is submitted to the IEEE Transactions on Biomedical Engineering (IEEE TBME) and is currently undergoing review.

3. Proposes a generalized defect detection framework that automatically detects known genetic defects and highlights candidate defective areas in 3D micro-CT images of genetically engineered prenatal mice. The proposed framework greatly enhances the throughput of the traditional histology-based defect detection methodology by pruning the vast search space of novel defects and

also highlighting candidate areas that are hard to recognize by human eye due to lack of significant visual features. This work has been published in the 16th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'13) [Roy et al. 2013b].

1.4 Road Map

The rest of this thesis is organized as follows: Chapter 2 describes our visual reporting framework that uses computer graphics technologies to enhance radiological reporting. In Chapter 3 we propose 3D features for a fast content-based image retrieval system for liver tumors. Chapter 4 describes how simple image processing techniques like registration and deformation vectors can be used to automate the traditional and still state-of-the-art microscopic techniques used in mouse defect detection. Chapter 5 concludes the thesis and outlines some future research directions.

Chapter 2

Visual Interpretation with Three-Dimensional Annotations

2.1 Overview

Imaging is a vital component of modern medicine. A typical clinical workflow involves a referring physician who requests an image exam and a radiologist who interprets this exam and prepares an exam report. The exam interpretation often involves using image-based tools to markup annotations on the exam's images, for example, drawing geometric primitives to denote lengths and volumes of interest, drawing arrows with text annotations, and selecting key images from the exam that represent images of clinical significance. A final text-based report is prepared to summarize the key findings. The benefits of these radiological visual annotations are well known in the medical community [Fan et al. 2011], [Armato 2003] and [Reiner and Siegel 2006]. Unfortunately however, this information is not always easily available to the referring physician. This could be due to several reasons

starting from software incompatibilities between physician's office desktops and radiologist's workstations to tedious workflows which require the physicians to sift through exams that often contain thousands of images [Rubin 2000] to find and match radiological annotations with the findings described in the report. In addition, because the annotations are made only on a few key images, the context of the annotation within the 3D volume may not always be clear, especially to the patient.

In some cases, the radiological tool vendors employ proprietary annotation implementations. As a result, it is often not possible to view annotations generated in a radiologist's workstation on other workstations like the ones present in the referring physician's offices or on data given to the patients (e.g., CDROMS). As a result, referring physicians sometimes rely only on the text-based reports as the primary means to interpret exams and communicate diagnosis to their patients. The inherent disadvantages of text-based reports are well documented in the literature [Reiner and Siegel 2006], [Schwartz et al. 2011] and [Bosmans et al. 2011].

In this chapter we propose to enhance this archaic radiological workflow, especially the reliance on text-based reports for communication between radiologists, physicians and their patients by proposing a software framework that allows automatic generation of 3D visual reports of exam findings. Our application framework, called visual interpretation with three-dimensional annotations (VITA), extracts annotations made by radiologists and generates a visual summary in the form of an animated 3D rotating volume of the exam with the radiologist's annotations clearly highlighted. VITA summaries are intended to augment radiologists' text-based reports by placing the annotation into a better visual context in the 3D volume. This helps physicians in both understanding the radiological reports as well as

communicating diagnosis to their patients.

Further, we use the image-based 2D annotations to derive higher order 3D properties of the radiological markup by automated segmentation to produce visual reports that highlight segmented anatomy instead of the individual raw annotations. This framework is the first to attempt integration of 3D segmentation into the radiological workflow. We demonstrate the usefulness of 3D segmentation by utilizing segmentation results to enhance important clinical applications such as exam summarization, exam visualization and radiological reporting.

The rest of the chapter is organized as follows. Sections 2.2 and 2.3 give overview of a typical clinical set-up and how annotations are performed in the radiological workflow. Section 2.4 introduces the VITA framework. In Section 2.5 we describe how individual 2D annotations can be clustered to obtain 3D characteristics of the annotated anatomy. Usefulness of deriving the 3D characteristics is demonstrated by enhancing three routine clinical applications.

2.2 Radiological Reporting Workflow

A typical radiology reporting workflow employed in clinical practice is pictorially presented in Figure 2.1. Central to this framework is the Picture Archiving and Communication System (PACS) server [Choplin et al. 1992] which is used as the central repository of image-based studies and the Radiology Information System (RIS) which is typically where the text-based report of the radiology exam is stored. Medical images are stored in the PACS archive in the Digital Imaging and Communications in Medicine (DICOM) format [NEMA 2008]. It is quite common that a PACS server will be shared by a cluster of institutions via a service provider, while

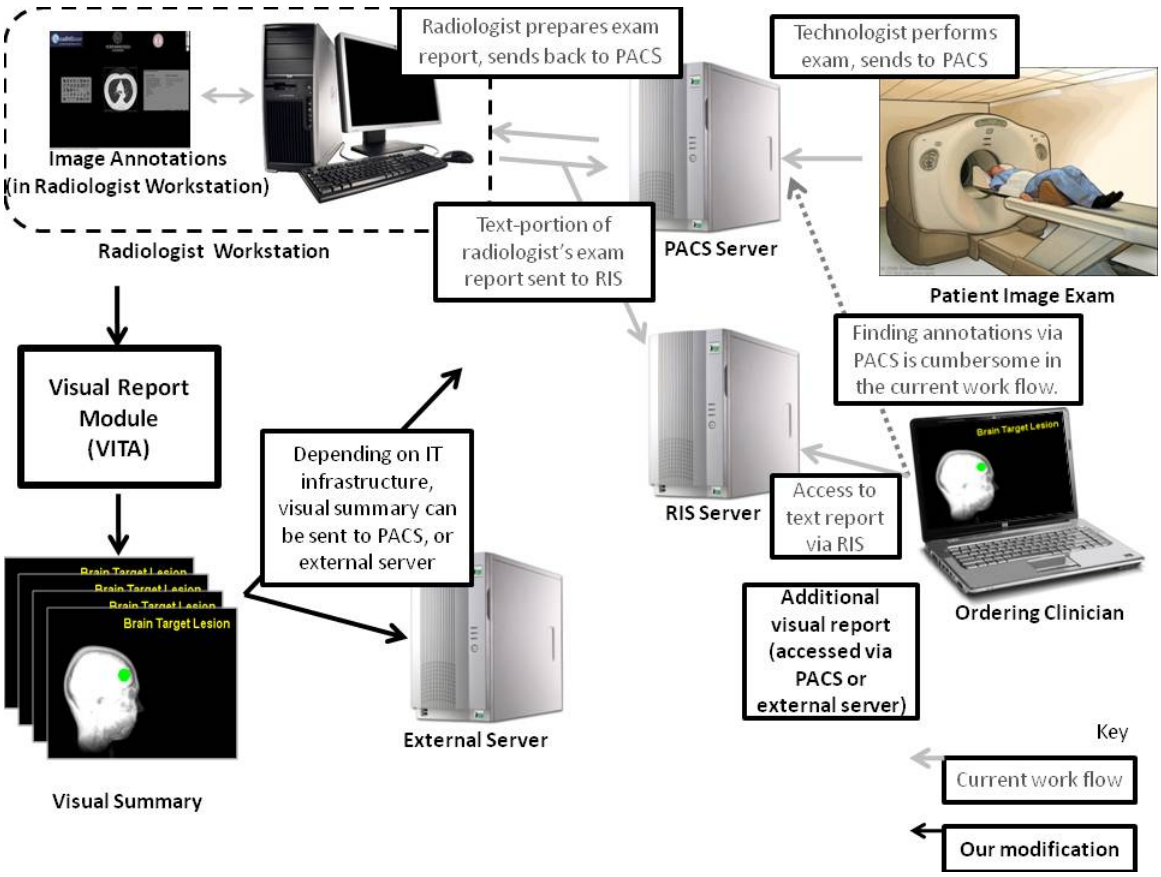


Figure 2.1: This figure gives an overview of a typical radiological reporting set-up and explains how our visual report module can be integrated into the existing workflow. Our framework can work either directly with Picture Archiving and Communication System (PACS), Radiology Information System (RIS), or even an external database that is cross-referenced via RIS.

RIS servers are operated by the institutions themselves. Under this framework, a technologist performs an image-based exam on a patient that is sent to the PACS server where a radiologist accesses the exam images and prepares a report for the referring physician. This exam report is sent back to the PACS server with the text-based portion also sent to the RIS server. The referring physician typically receives the report via RIS. The text-report in the RIS system can sometimes refer back to the original study in the PACS system.

Figure 2.1 also outlines how our visual report module, VITA, integrates with the current radiology workflow. VITA uses the exam images and annotations employed by the radiologist during exam interpretation to produce a visual summary. This visual summary is sent to the PACS server as a new DICOM series which can be downloaded by the referring physicians for diagnosis.

The key research challenge lies in exploiting annotations commonly made by radiologists to produce a structured visual report. The VITA report is generated as a series of DICOM images which is distributed back to the PACS archive. Since the annotations are now embedded in the DICOM pixel data of the visual report, many issues of software incompatibilities with regards to annotation implementation across PACS vendors are avoided. This also allows our VITA framework to seamlessly integrate within the existing workflow as no additional input is needed from the radiologists and the results are available in PACS. In addition, we have the ability to produce video versions (e.g., AVI, MOV, or MPEG) of the VITA summary for sharing with patients and for situations when access to PACS or DICOM viewers is not readily available.

2.3 Radiological Annotation Implementation

While PACS and DICOM are supported by all vendor software, the manner in which proprietary software encodes annotations and markup is often a source of incompatibility. Most software, however, use a structured format like Extendible Markup Language (XML) to implement annotations. Further, in order to unify annotations across PACS frameworks, the National Institutes of Health Cancer Biomedical Informatics Grid (NIH caBIG) has initiated the Annotation and Image

Markup (AIM) project [Channin et al. 2009], [Channin et al. 2010] and [Rubin et al. 2008a] which provides a standards-based annotation format that can be shared between different PACS. With AIM it is now easy to extract and utilize annotations generated using all PACS workstations compatible with the standard. A few PACS implementations have incorporated AIM already [ClearCanvas 2008], [Rosset et al. 2004] and [Rubin et al. 2008b] and it is being used by other academic institutions [Rubin et al. 2008b] and [Zimmerman et al. 2011] for radiological reporting. One key benefit of AIM is that it provides a well-defined and structured format using XML for radiological annotations that can be easily parsed. VITA supports the AIM initiative and is compatible with its latest version. AIM has been chosen as a representative standard to present the idea behind VITA; it is not hard to incorporate other clinically used standards like Health Level 7 (HL7) in VITA to implement the visual summaries. VITA has a built-in module to parse structured annotation files. The current version parses the AIM schema and also XML-based structured format used by ClearCanvas. The parser can be extended to read other popular formats as well.

Other works described in prior literature have examined how to use annotations to automatically generate text-based reports [Zimmerman et al. 2011], however, VITA system is the first to target 3D visual reporting. The VITA system does not require any high-level processing or understanding of the annotations; it simply uses what is already provided by the radiologist in routine practice.

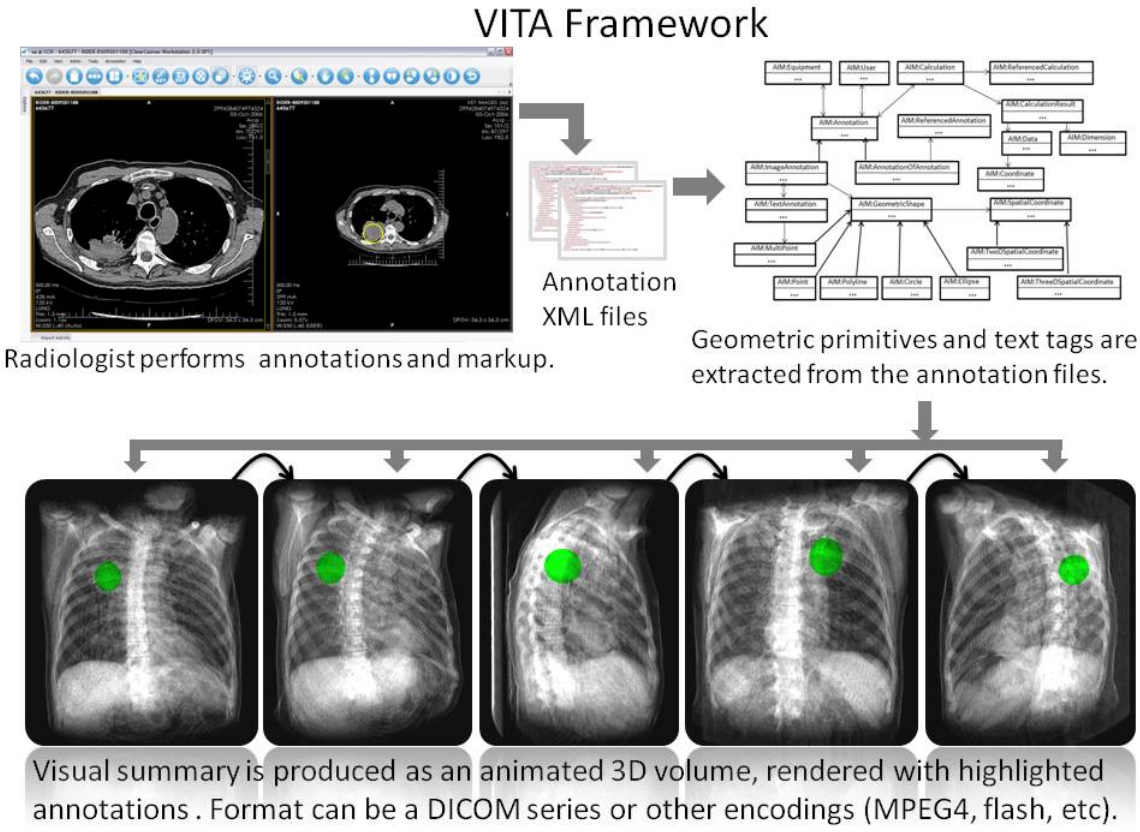


Figure 2.2: The VITA framework uses radiologist annotations prepared using a structured format (e.g., Extendible Markup Language (XML), Annotation Image Markup (AIM)). Geometric primitives are extracted from the annotation encodings and used to produce visual summary in the form of a rotating 3D volume rendering.

2.4 The VITA System

Our VITA system is developed in C++ using Nokia's Qt cross-platform application and UI framework [Nokia 2009] on an Intel core i5, 2.4 GHz processor with 3 Gigabyte Random Access Memory and NVIDIA GeForce GT 330M Graphics Card. VITA can be used on standalone computer or with ClearCanvas PACS workstation. ClearCanvas workstation is a free and open-source workstation with an active developer community [ClearCanvas 2008]. It is currently used by more than 20, 000

healthcare professionals worldwide and the clinical edition has been approved by the Food and Drug Administration (FDA). ClearCanvas workstation allows radiologists to draw geometric shapes over images, e.g., lines, ellipses etc and associate text with the drawn geometry. This markup is saved either as ClearCanvas study file in the XML format or in the unified AIM schema in both XML and Digital Imaging and Communication in Medicine Structured Reporting (DICOM-SR) format.

VITA has a built-in XML parser module that mines ClearCanvas study files and AIM XML files associated with a medical exam to extract the annotated geometric primitives, observations and text tags. The geometric primitives are sent to the rendering engine of VITA which produces a visual summary in the form of a 3D volume animation that renders the volume as it spins 360° around the spinal axis. The geometric primitives are distinctly highlighted in the volume. Figure 2.2 captures this pipeline.

VITA uses ray casting as the primary method to generate 3D volume images. The core of the ray casting algorithm is to send one ray per screen pixel and trace this ray through the volume. To exploit modern high-end Graphic Processing Units (GPU), a GPU-based ray casting engine (using NVIDIA's Cg toolkit [NVIDIA 2010]) has also been implemented which can render high quality volume images at interactive speed. The ray casting engine takes the pixel data from the exam images and the annotated geometric primitives to first compute a rotating volume with geometry highlighted. The text tags and observations, if any, are then placed over the corresponding geometry to generate the complete summary. Figure 2.3 outlines this procedure.

As the volume spins, the visual report is saved by generating image files in the

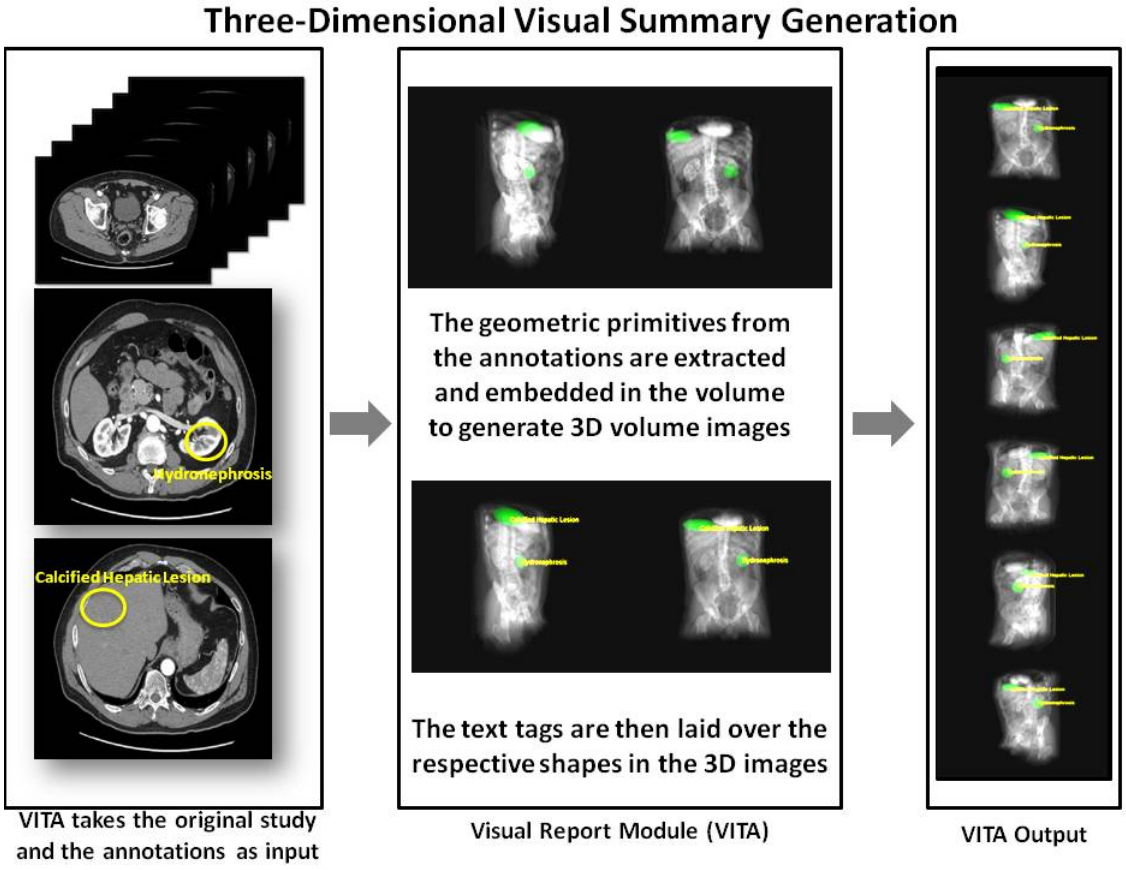


Figure 2.3: VITA needs the original stack of DICOM images and the annotation information (e.g., geometry, text tags) to generate the visual summary. The geometry is first embedded in the volume and the text tags are then overlaid to compute the final report.

DICOM format at every 10° rotation of the volume. The Insight Segmentation and Registration Toolkit (ITK) [ITK 1999] is used to generate the DICOM images. Open-source DCMTK library from Offis [DMCTK 2003] is used to insert appropriate DICOM tags that composes this new set of DICOM images into an image stack. This new image stack forms an additional series in the original exam. VITA then pushes this series to the PACS archive using DICOM message exchange functions available in DCMTK. This summary series can now be downloaded by the referring

physicians and played in cine mode in their respective DICOM viewers.

Although the VITA visual report is presently generated as a DICOM stack, it is possible to prepare it in other more compact video formats such as MPEG4 or Flash, which may be stored in other information systems such as the RIS. This non-DICOM solution might prove to be more efficient for the referring physicians and patients as the reports can now be entirely encapsulated without the need for a DICOM viewer or PACS system allowing for easier access on mobile devices such as tablets or smart-phones.

2.4.1 Results

A computer running ClearCanvas PACS workstation and the VITA application was connected to a PACS server that contained CT and MR exam images from the online cancer image archive [Armato et al. 2011]. Images were downloaded from the PACS server onto the workstation and sample annotations were created using ClearCanvas image-based measurement tools and the tools available with the AIM plug-in. VITA parsed these annotation files using its inbuilt XML parser and generated visual reports for each exam based on the respective annotations. These reports were automatically sent by VITA to the PACS server over the network as additional DICOM series to the respective exams. To communicate with the PACS server, VITA requires the server name, server application entity title and the port on which the server is running at the remote machine. The summary exams were then downloaded and checked in another workstation to ascertain that all the information was properly rendered and saved.

Figure 2.4 shows a snapshot of VITA. A ClearCanvas measurement tool was

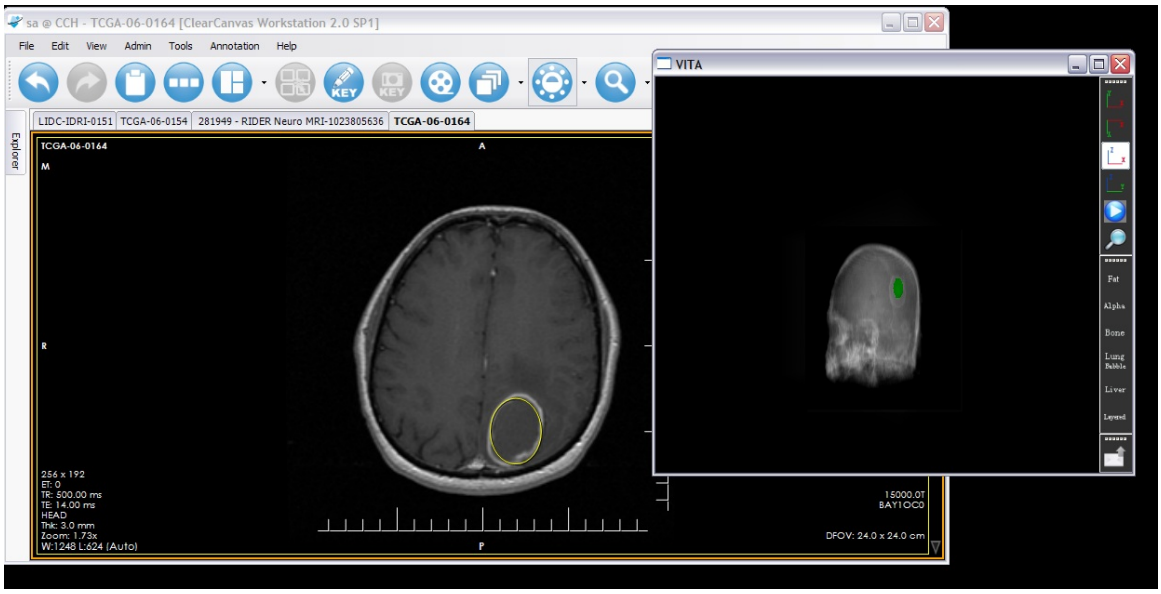


Figure 2.4: This figure shows a snapshot of VITA when used with ClearCanvas PACS workstation. VITA reads the annotations made in ClearCanvas and embeds them in the visual report.

used to mark an elliptical region of interest in an MR image of the brain. ClearCanvas workstation stores annotations generated using the measurement tools in the XML format. Figure 2.4 shows the visual report generated by VITA using the elliptical annotation. The geometric shape in the 3D visual summary shows the shape and position of the radiologist's annotation.

Figure 2.5 shows example images taken from a visual summary generated by VITA. The annotations visible in this report were generated in the standardized AIM schema using the AIM plug-in available in ClearCanvas workstation. The images used in this figure are of a healthy patient; the annotations are representative examples only.

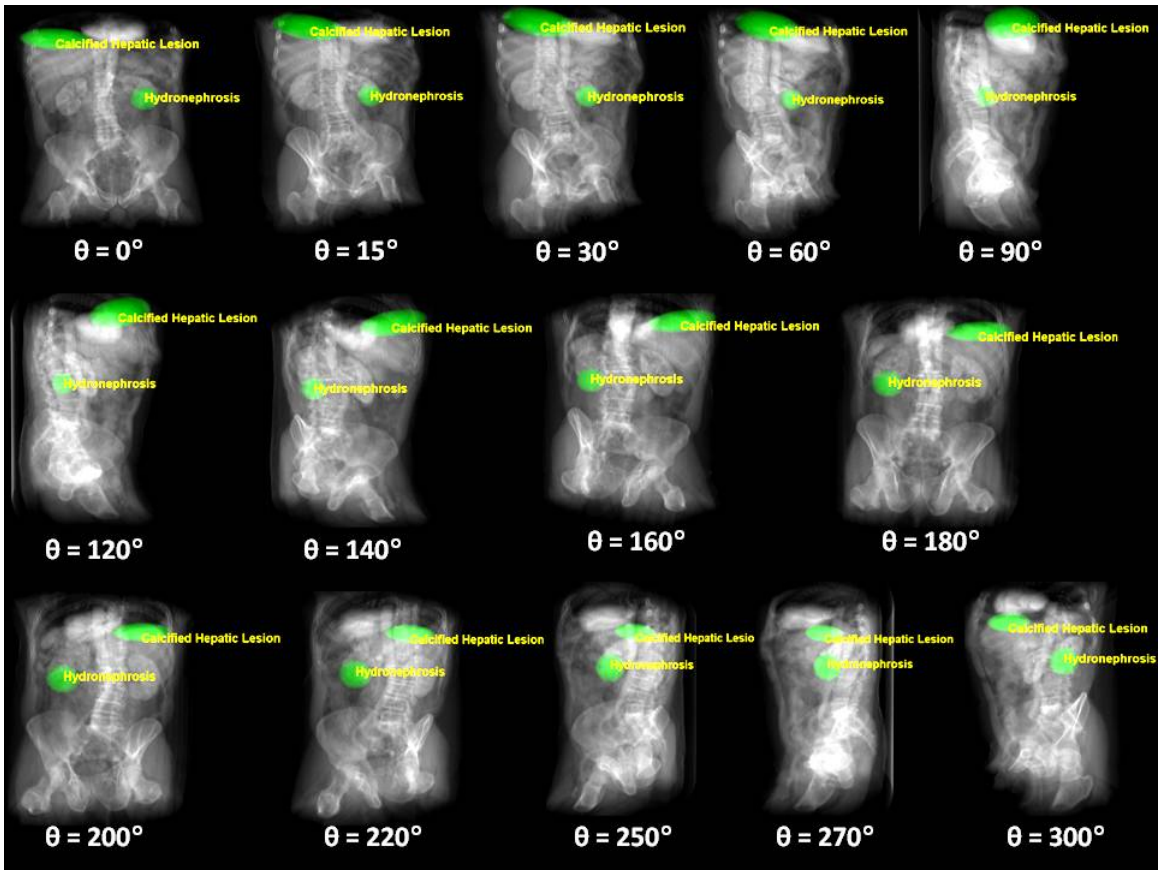


Figure 2.5: The visual summary consists of a rotating volume with annotations distinctly highlighted. The volume spins to provide a comprehensive 3D context of the important clinical observations. Θ in the figure demonstrates the angle of rotation with respect to the spinal axis.

It is possible to control the way annotation text is overlaid on the geometry. For example, VITA can place the text over the geometry and let both spin together, or have the text stationary and color-code it with the respective spinning geometry. Figure 2.6 shows examples of both the scenarios.

VITA can also selectively highlight important tissues while generating the visual report. This is achieved by applying transfer functions to the exam images such that only the relevant tissues are visible. Figure 2.7 shows images from two such

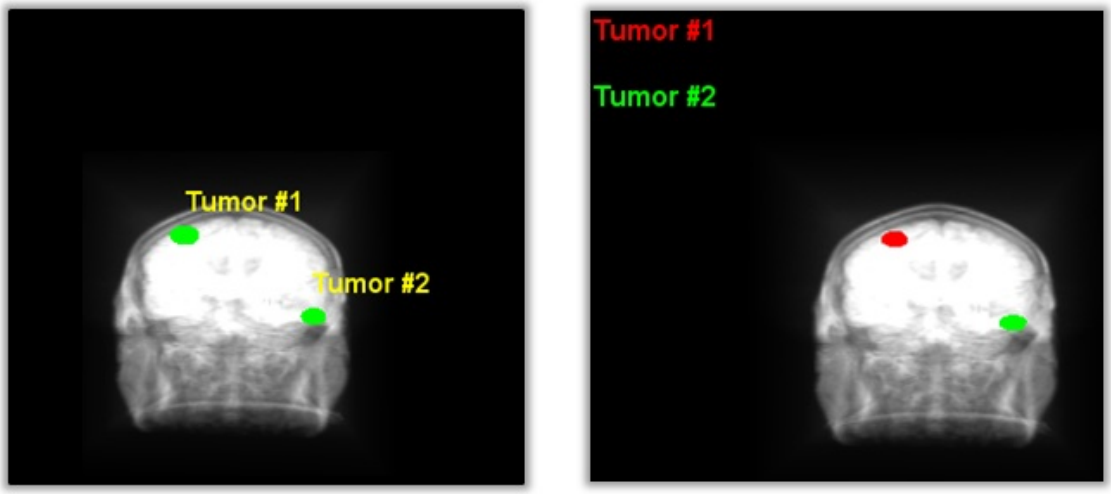


Figure 2.6: It is possible to either let the text tags move with the geometry as the volume spins or have the text stationary and color-coded with the geometry.

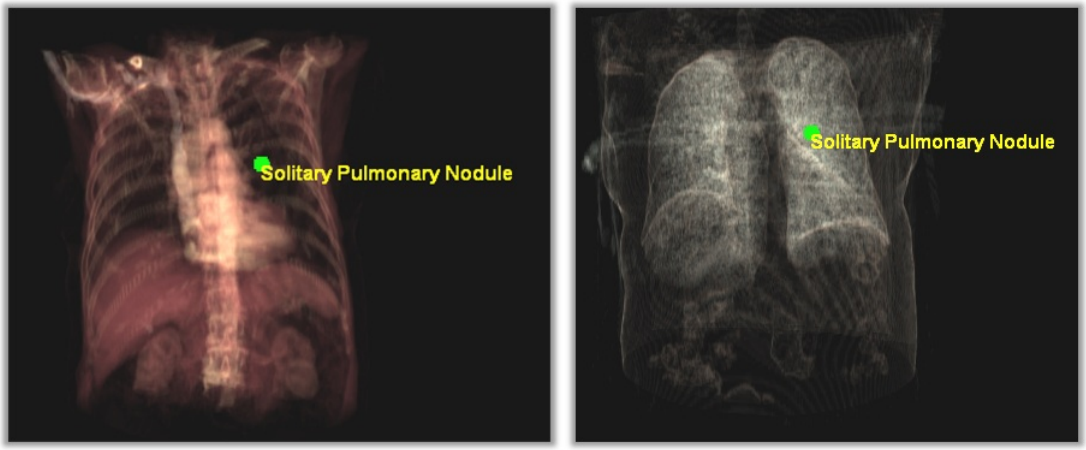


Figure 2.7: Certain body tissues can be highlighted using presets available in the volume rendering module of VITA. The left image is generated using a preset which accentuates bone tissues and the right image is generated by accentuating the lung tissues.

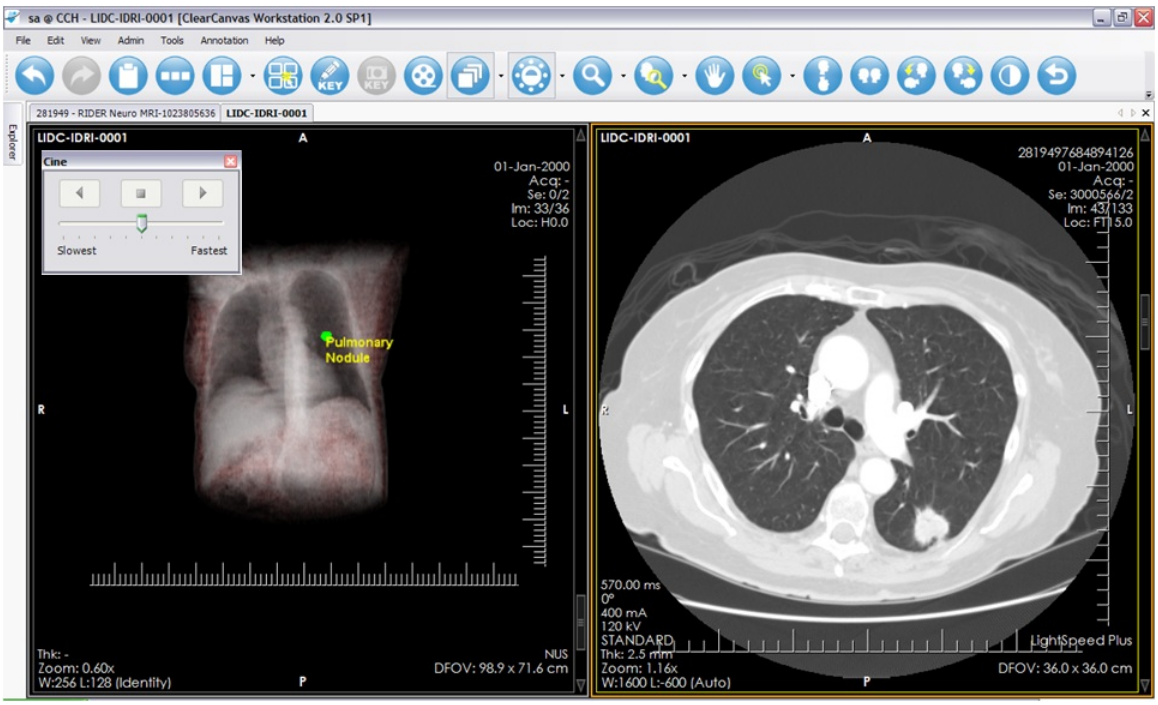


Figure 2.8: Once the visual report is placed back in the PACS archive as an additional DICOM series, it can be accessed by clinicians in their respective DICOM viewers. The animated summary can be viewed in the cine mode available in most DICOM viewers.

reports. The left image is part of a visual report which selectively highlights the bone tissues and the right image is from a report that shows no bone tissues but emphasizes the lung tissues.

After VITA placed the visual reports in the PACS archive, studies having visual reports were downloaded at another computer running a PACS workstation. Visual reports appeared along with pre-existing exam images. The animated visual summary was viewed in cine mode which played the report images at the user desired playback speed. Figure 2.8 shows a snapshot of the ClearCanvas workstation displaying the visual report.

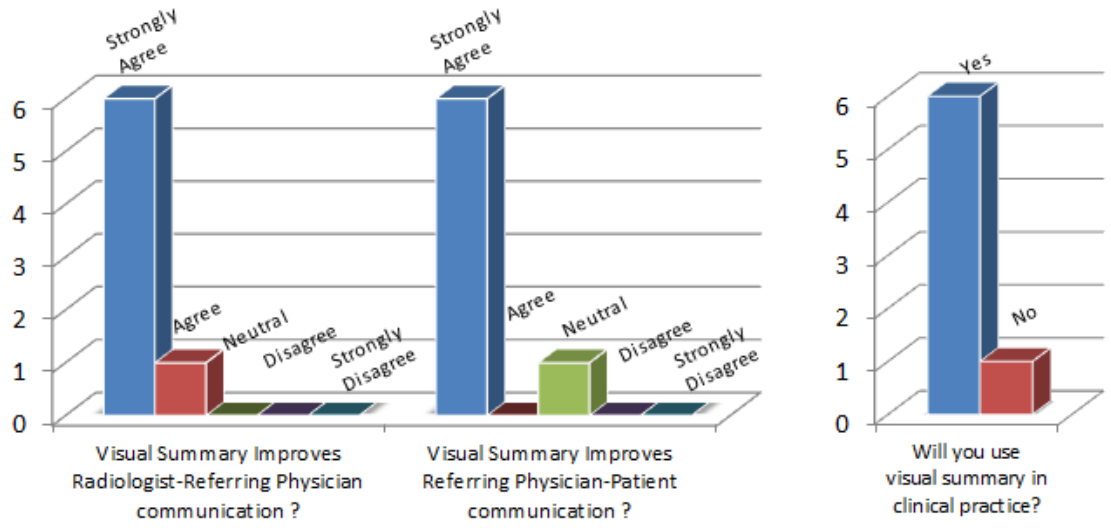


Figure 2.9: This figure shows the results of a user satisfaction study performed with seven referring physicians. Six out of seven participants strongly agreed that visual summary improves clarity of communication between radiologists and referring physicians and also agreed that visual summary aids patient communication. Six participants were willing to use this service, if provided.

2.4.2 Evaluation by User Satisfaction Survey

To test the effectiveness of VITA, a user satisfaction study was conducted with seven participating referring physicians. The participants were shown three anonymized radiological reports along with the corresponding key images that had the graphical overlays of the annotations. The first case reported a solitary pulmonary nodule in lung CT exam, second case reported tumor in brain MR images and the third case was about calcified lesion in the liver diagnosed in abdominal CT. For these three cases participants were also given the visual summaries generated by VITA. Physicians were asked three questions: 1) whether the VITA summaries improve clarity of communication between referring physicians and radiologists; 2) whether the VITA summaries would be useful in assisting physicians in communicating

diagnosis to patients and; 3) whether they would be willing to use the VITA service in their clinical routine if made available. The first two questions asked the participants to rate their answers on a 5-point Likert scale with one being strongly disagree and five being strongly agree.

Figure 2.9 shows the results of the evaluation survey. Six out of seven participating referring physicians strongly agreed that 3D visual summaries improved clarity of communication between radiologists and physicians and also strongly agreed that 3D visual summaries would aid patient communication. One physician agreed that visual report improves clarity of communication between radiologists and physicians and was neutral on whether visual summary aided patient communication. Six participants were willing to use the system in their routine clinical practice. Comments from participants were also positive, samples include "it is a new brilliant concept for patient understanding" and "this is an excellent intervention which helps in better collaboration between physician and radiologist." The only physician who was neutral on using VITA commented that "3D rendering does not add additional information for clinician. It looks nice for the layperson i.e., patient, but clinical use is very limited." Indeed for clinicians who are expert readers of 2D radiological images, VITA visual reports do not add any additional clinical content. However, for clinicians who do not have this expertise, the visual reports represent the clinical findings in a more comprehensible way.

2.4.3 Discussion

Because VITA framework requires minimal changes to the current radiological workflow, we can easily integrate it into existing systems. Currently, a typical

clinical set-up uses PACS and Electronic Medical Record (EMR) servers to share exam images and text-based reports among radiologists and doctors. The same infrastructure, along with VITA, is sufficient to share the 3D visual summaries. Radiologists can continue to use their native PACS workstations for annotations. VITA reads these annotations and places a visual summary at the PACS server for easy access to the physicians. In addition, VITA requires no user interaction, it can run in the background and automatically generate and archive summary images. If radiologists want to change the transfer function, it can be done by a single click. VITA uses the computing power of GPU available in most modern computing devices. The computation time depends upon the size of the medical exam and the hardware used. On an average, for an exam consisting of 300 images, the total time to compute visual summary is less than one minute on an Intel core i5, 2.4 GHz processor with 3 Gigabyte RAM and NVIDIA GeForce GT 330M Graphics Card.

Usability is another aspect of deploying an application into a clinical routine. We have demonstrated our VITA tool to a number of radiologists in Cornell Medical Centre, New York and at the 98^{th} annual meeting of the RSNA. VITA was appreciated by radiologists and the initial feedback was positive even though information about usability is yet unavailable.

In small-scale medical infrastructures (e.g., smaller community hospitals) it is quite common that a PACS server will be shared by a cluster of institutions via a service provider, while RIS servers will be operated by the institutions themselves. In such scenarios it may not be practical for VITA to send data to the PACS server. In such situations, the visual reports may be generated in MPEG4 or Flash formats and placed at the RIS server or any external database maintained by the hospital's information system administration. Furthermore, reports in these compact video

formats can be easily accessed in mobile computing devices that run no PACS applications.

The recent trend towards patient empowerment advocates that patients' access to their imaging reports be facilitated [Hall 2009] and [Chiaramonte 2008] and radiology reports, or the information in them should be made more patient-centric. A recent article by Berlin [Berlin 2009] concludes that radiologists have a medical-legal duty to directly communicate with patients. We believe that 3D images of the VITA visual report are well suited for this patient-radiologist communication. The 3D nature of the images in the visual report and the embedded annotations therein can greatly enhance patients' comprehension of the pathology. Patients can be given visual reports in video formats like MPEG4 or Flash so that they can view the important findings noted by the radiologist on the imaging study without having to access a PACS workstation.

For optimum results it is desirable that radiologists associate some semantic information with the image markup. Nonetheless, for generation of visual summary semantic tags are not a pre-requisite. By highlighting the geometric primitives drawn by the radiologist in the visual summary, the annotations are put into a context and the referring physicians and their patients get a better 3D comprehension of the pathology which is lacking in the text reports and the key images. Further, we believe that if radiologists know that their annotations will be used to generate 3D visual summaries that are useful to referring physicians and patients, they may be more willing to add associated text annotations along with geometric primitives. An idea for future work is to try to extract this additional text information from the text-based report itself. This is a non-trivial problem given the potential unstructured nature of the text reports.

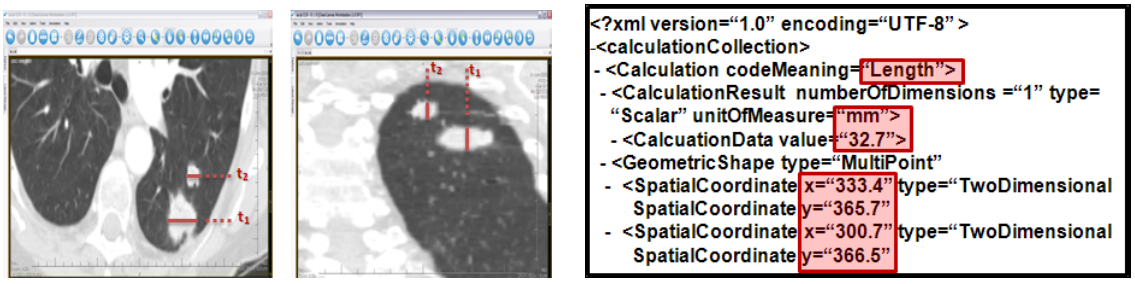
As referring physicians start to use VITA visual summaries, it is also possible to develop a simple yet powerful visual reporting language that can allow radiologists to provide directives on how the visual summary should appear in order to visually highlight salient findings to assist the referring doctor's understanding of the patient's disease processes. For example, in a case where a patient has multiple small lesions and one big lesion, the visual language can allow simple keyword annotations, e.g., <highlight>Pulmonary lesion</highlight>, as a directive to the volume rendering engine that this annotation should be displayed in a more prominent manner. By defining a small set of simple tags, radiologists could have more control over the final visual summary. This can help better guide the referring provider's focus towards key findings.

The current version of VITA tool is licensed under a BSD type license and is publicly available at http://www.comp.nus.edu.sg/~sharmili/projects/VITA/index.html along with the source code and documentation.

2.5 Extracting Volumes from 2D Annotations

While working on 3D representation of the radiological annotations we found that PACS workstations *do not* support volumetric annotations. As a result, radiologists are forced to draw two or more line segments in different planes (i.e., axial, coronal, sagittal) to denote a rough volume of an object (see Figure 2.10(a)). In addition, the PACS software does not semantically link the line segments, instead, the radiologist summarizes these measurements manually in a text-based report.

While simple 2D markup is the defacto method for radiological annotations, it has been shown that segmentation of objects of clinical significance provides better



(a) (b)

Figure 2.10: Routine radiological annotations. (a) Sample annotations (line segments) drawn over lung tumors in the axial and coronal planes. (b) Example of XML meta-data defined by AIM to store these annotations.

diagnosis for various conditions [Kostis et al. 2003] and [Jirapatnakul et al. 2011]. Furthermore, accurate volumetric measurements can have important diagnostic and therapeutic implications for certain diseases that rely on the rate of growth not easily extrapolated from 2D measurements. Segmentation is a well-studied problem in the medical image research community (see [Sharma and Aggarwal 2010] for a nice survey). Most area-based and volumetric segmentation, however, are performed outside the PACS workstation. In addition, these algorithms require user-supplied seeding or markup to initiate the segmentation procedure. This has two negative consequences. First, because volumetric segmentation is outside the conventional workflow it is not applied as often as it should be. Also, not all radiologists (e.g., smaller community hospitals) have access to workstations that provide segmentation capabilities. Second, when volumetric segmentation is applied, the radiologist is required to make redundant markup or seeding that has already been provided during the standard reporting workflow.

We propose an application framework to integrate volumetric segmentation within the standard radiological reporting workflow. Specifically, we describe

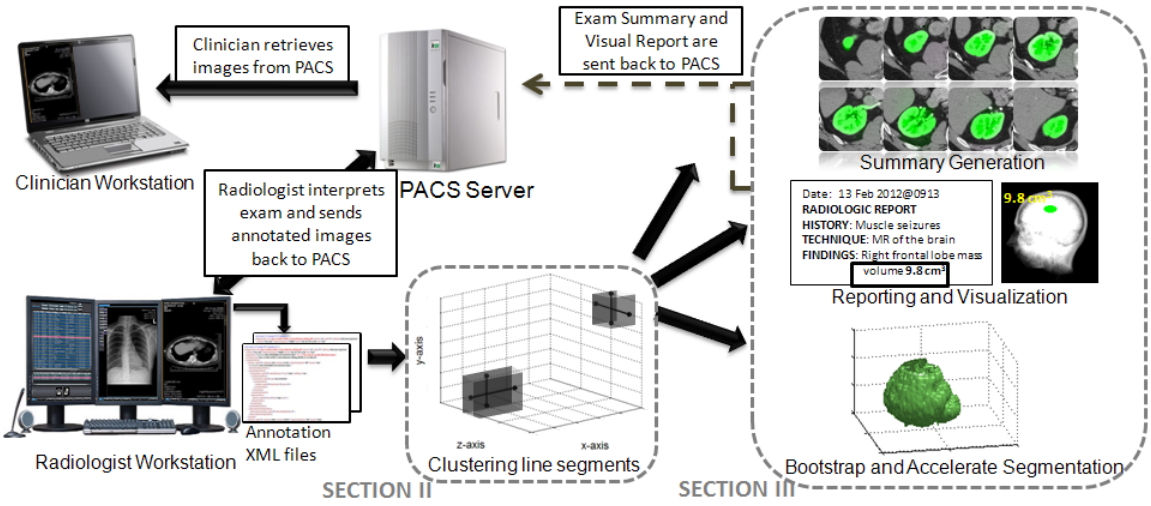


Figure 2.11: This figure gives an overview of how our application framework integrates into a typical clinical set-up. The proposed application clusters existing annotations into volumes and bootstraps 3D segmentation. The 3D information is used to enhance applications such as reporting, visualization and summary generation.

a clustering algorithm to extract volume data from unstructured line segments present in standard radiological reports. These extracted volumes are then used to bootstrap 3D segmentation. Residing on the radiologist's PACS workstation, our software parses the XML annotations to extract marked line segments. These line segments are then clustered into volumes and subsequently 3D segmentation is initiated. See Figure 2.11 for an overview of the proposed system.

2.5.1 Associating Line Segments to Volumes

As previously mentioned, the line segments are not semantically linked in the exam annotations. Our challenge is to determine which segments are intended to denote volumes (Figure 2.12) and which are not. The general practice in radiology is to use line segments in two or more orthogonal planes to denote a volume. While

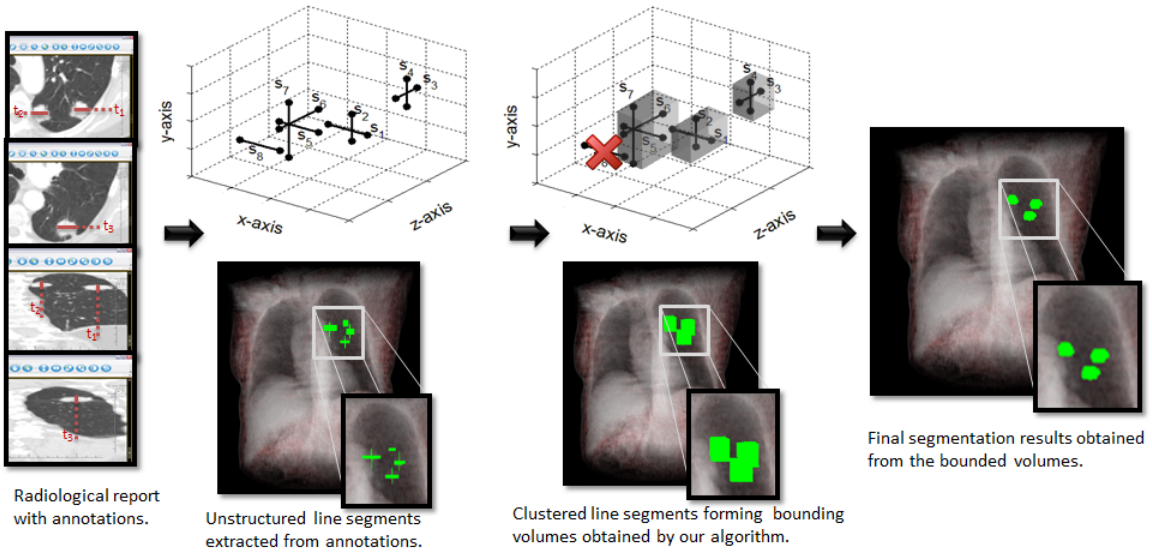


Figure 2.12: Our input includes the images from the study and the annotations reported by radiologists. We mine these annotations to get the unstructured line segments. The line segments are then clustered to determine bounding volumes. Information from the bounding volume is used to perform 3D segmentation.

ideally these line segments should intersect in the volumetric space, this may not always be the case due to variations in radiological procedures and quantization of the coordinate frames. It is however safe to assume that the segments should be close to each other and in orthogonal planes and that the segments should overlap.

To determine which segments overlap and the distance between overlapping segments, we use the segment ray distance method [Eberly 2006], described here in brief. Given two segments S_1 (between endpoints P_0 and P_1) and S_2 (between endpoints Q_0 and Q_1), the first step is to determine the closest points between the lines on which these segments lie (Figure 2.13). Segment S_1 is represented as the points on $L1 : P(s) = P_0 + s(P_1 - P_0) = P_0 + s\mathbf{u}$ where $0 <= s <= 1$ and $\mathbf{u} = (P_1 - P_0)$ is the line direction. Similarly, the segment S_2 on L_2 from Q_0 to Q_1 is given by the points $Q(t) = Q_0 + t\mathbf{v}$ with $0 <= t <= 1$ and $\mathbf{v} = (Q_1 - Q_0)$. Let $\mathbf{w} = P(s) - Q(t)$ be a

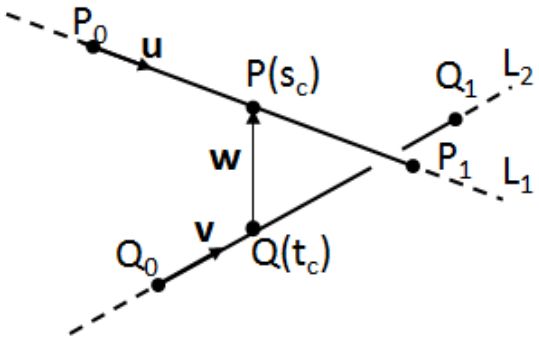


Figure 2.13: Given two segments, one between endpoints P_0 and P_1 and the other between endpoints Q_0 and Q_1, we compute the closest points ($P(s_c)$ and $Q(t_c)$) between the lines on which these segments lie. If $P(s_c)$ and $Q(t_c)$ lie within their respective line segments, then the segments overlap otherwise not.

vector between points on the two lines.

There exists unique points, $P(s_c)$ and $Q(t_c)$, at which $\mathbf{w}$ is the shortest and the two lines L_1 and L_2 are therefore closest. In addition, if L_1 and L_2 are non-parallel, then $\mathbf{w}$ is perpendicular to both L_1 and L_2 s.t. $\mathbf{w} \cdot \mathbf{v} = 0$ and $\mathbf{w} \cdot \mathbf{u} = 0$.

Substituting $\mathbf{w} = P(s_c) - Q(t_c) = \mathbf{w}_0 + s_c\mathbf{u} - t_c\mathbf{v}$, where $\mathbf{w}_0 = P_0 - Q_0$, into each of these equations, we obtain the following:

$$(\mathbf{u} \cdot \mathbf{v})s_c - (\mathbf{v} \cdot \mathbf{v})t_c = -\mathbf{w}_0 \cdot \mathbf{v},$$

$$(\mathbf{u} \cdot \mathbf{u})s_c - (\mathbf{v} \cdot \mathbf{u})t_c = -\mathbf{w}_0 \cdot \mathbf{u}.$$

Solving these equations we find s_c and t_c for L_1 and L_2. If $0 <= s_c <= 1$ and $0 <= t_c <= 1$, the segments S_1 and S_2 overlap and $|P(s_c) - Q(t_c)|$ gives the minimum distance between them. However, if $s_c < 0$ or $s_c > 1$ and similarly for t_c, then the closest points between L_1 and L_2 do not lie within the respective segments S_1 and S_2 and hence the two segments do not overlap.

Algorithm 1 Cluster line segments into volumes.

1: $\mathcal{A}$ is the set containing all line segments in axial plane, set C contains all line segments in coronal plane, set $\mathcal{S}$ contains all line segments belonging to sagittal plane
2: $\mathcal{P} = \emptyset$, $\mathcal{P}$ is the set that holds volume clusters
3: **for** i = 1:size($\mathcal{A}$) **do**
4: $S_1 = \mathcal{A}(i)$ such that $S_1 \notin \mathcal{P}$
5: find ($S_2 \in C \cup \mathcal{S}$) and ($S_2 \notin \mathcal{P}$) such that S_1 and S_2 form the closest pair
6: **if** ($S_2 \in C$) **then**
7: find ($S_3 \in \mathcal{S}$) and ($S_3 \notin \mathcal{P}$) such that S_1 and S_3 form the closest pair
8: **if** S_2 and S_3 overlap **then**
9: (S_1, S_2, S_3) form a cluster, $\mathcal{P} \leftarrow (S_1, S_2, S_3)$
10: **else**
11: (S_1, S_2) form a cluster, $\mathcal{P} \leftarrow (S_1, S_2)$
12: **end if**
13: **else**
14: find ($S_3 \in C$) and ($S_3 \notin \mathcal{P}$) such that S_1 and S_3 form the closest pair
15: **if** S_2 and S_3 overlap **then**
16: (S_1, S_2, S_3) form a cluster, $\mathcal{P} \leftarrow (S_1, S_2, S_3)$
17: **else**
18: (S_1, S_2) form a cluster, $\mathcal{P} \leftarrow (S_1, S_2)$
19: **end if**
20: **end if**
21: **end for**
22: Repeat steps 3 to 21 analogously for all segments in C and $\mathcal{S}$

Using the distance calculation method we find a pair of segments from orthogonal planes that are closest to each other. Then we search the third orthogonal plane to check if there exists any segment which overlaps with this pair. For example, if two segments, S_1 and S_2, are closest to each other and belong to axial and sagittal planes respectively, then we search the coronal plane to find a line segment, S_3, which is closest to either S_1 or S_2. If S_3 happens to overlap both S_1 and S_2 then S_1, S_2 and S_3 form a volume. Analogous logic is applied if S_1 and S_2 belong to other orthogonal planes. See Algorithm 1 for further details. The point where all

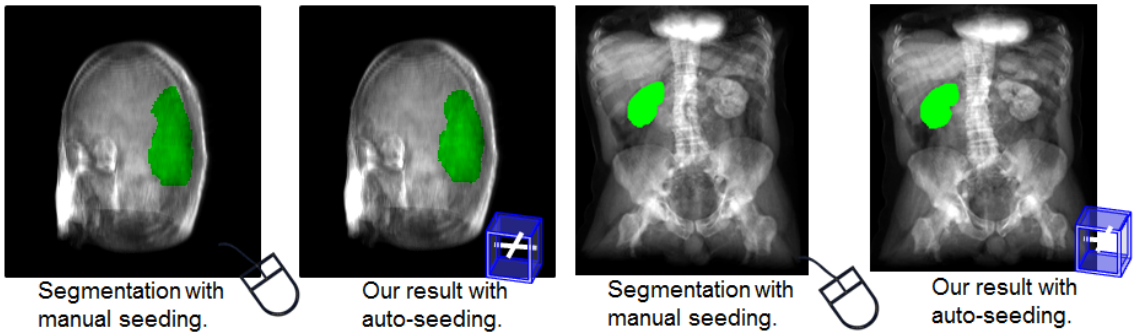


Figure 2.14: This figure shows segmentation results obtained by applying level set segmenter on brain tumor and kidney. The figure shows that our automatically seeded results are qualitatively similar to those obtained using manual seeding.

the line segments of a cluster overlap gives the location of the volume and the line segments themselves subtend a bounding volume around the anatomy marked by the radiologist. Figure 2.12 shows an example.

2.5.2 Bootstrapping and Accelerating Segmentation

The volume locations obtained by clustering the annotation data can be used to automatically seed segmentation algorithms. In our examples, we use the well-known level set approach proposed in [Li et al. 2010]; other segmentation algorithms could be used in a similar manner. Our seed is placed at the center of the bounding volume in the form of a cuboid that is 20% the size of the extracted bounding volume in height and width and is of unit depth. Figure 2.14 shows segmentation results of brain tumor and kidney obtained by automatic seed generated from the radiological annotations. The results look qualitatively similar to those obtained by manual seeding.

Table 2.1 shows the quantitative difference between the segmentation output when the segmentation seed is automatically generated versus when the seeds are

Table 2.1: This table provides quantitative difference between segmentation output of level set segmentation algorithm when seeds are automatically generated by clustering line segments versus when seeds are provided manually.

Pathology/ Anatomy	Auto Seed Volume (cm^3)	Manual Seed Volume (cm^3)	%Output Difference
Right Lung Tumor	2.96	3.05	2.9
Right Lung Tumor	5.30	5.34	0.7
Right Lung Tumor	6.94	6.76	2.6
Brain Tumor	55.66	55.28	0.7
Right Kidney	61.35	56.13	9.3

manually provided. In most cases the difference in output is less than 3% except for the case where an entire kidney is segmented where the difference is 9.3%. This happens because segmentation of kidney is inherently hard due to the organ's inhomogeneity caused by the presence of medulla.

In this work we use a level set segmentation algorithm for all images irrespective of the body part being examined. Hence, the method is most effective when the region to be segmented is homogenous and has a clear boundary. If we can design an algorithm to automatically detect the examined body part from the exam images and the annotations and subsequently initiate the most appropriate segmentation algorithm available for that specific part, the accuracy of volumetric measurements can be further improved.

The bounded volume obtained from annotations provide an approximate size of the object's structure. This can be used to accelerate the segmentation process by restricting the volume on which segmentation is applied. Table 2.2 lists execution times of the level set segmentation algorithm [Li et al. 2010] when the input is bounded based on the size of the volume obtained from annotations versus when the input is the entire data. The execution times are computed in MATLAB and for

Table 2.2: This table compares time taken by level set segmentation algorithm to segment anatomy when segmentation is bounded (t_b) or unbounded (t_{ub}). Data and anatomy sizes are given in pixels in [width,height,depth] format.

Pathology/ Anatomy	Data Size [w,h,d]	~Anatomy Size	t_b (min)	t_{ub} (min)
Right Lung Tumor	[512, 512, 133]	[26, 26, 07]	0.9	11.7
Right Lung Tumor	[512, 512, 133]	[26, 28, 07]	1.1	13.8
Right Lung Tumor	[512, 512, 133]	[30, 30, 07]	1.2	13.8
Brain Tumor	[256, 256, 052]	[60, 55, 25]	5.9	25.1

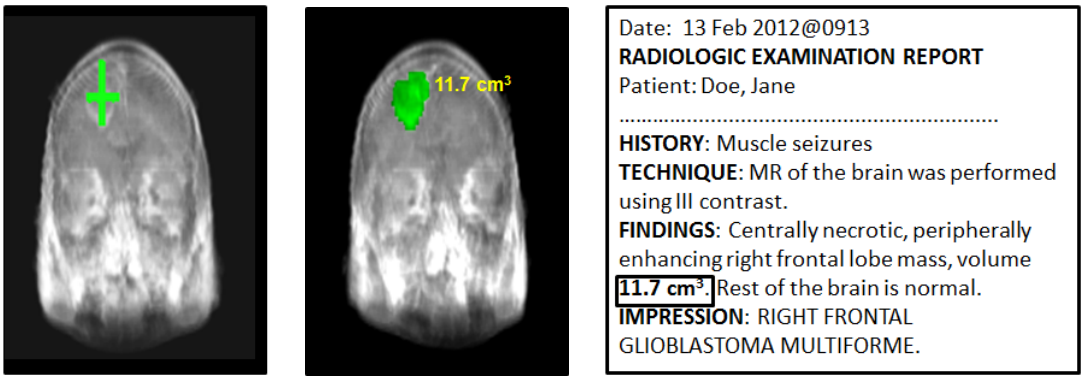
fair comparison, the seeds in both cases are kept identical. This table shows that by simply leveraging the bounding volume information in the existing radiological annotations we can achieve a significant increase in performance in addition to avoiding redundant seeding.

2.5.3 Reporting and Visualization

We discussed in the previous section that physicians seldom have access to radiologist's annotations. We have addressed this gap in communication by proposing the VITA system. A typical output generated by VITA is shown in Figure 2.15(a).

VITA produces 3D volume renderings of the annotations within the context of the whole volume. However, VITA does not attempt to derive any volumetric properties based on individual annotations. Using the clustering algorithm described in Section 2.5.1, we can now derive segmented volumes based on the 2D annotations and generate enhanced visual summary (Figure 2.15(b)).

Automatic report generation from image annotations is another interesting application [Zimmerman et al. 2011]. The existing approach in [Zimmerman et al. 2011] summarizes AIM annotations directly into the radiologist's text-based report,



(a) (b) (c)

Figure 2.15: (a) This is the output generated using VITA. (b) Our clustering algorithm can generate a segmented volume of the anatomy using 2D annotations prepared during reporting. (c) Volumetric measurements obtained from the segmented volume can be used to automatically produce value-added radiology reports.

but does not attempt to derive additional higher-level information based on the annotations. For many exams, volumetric information forms the most critical part of the report. Using our method, we can provide precise volumetric measurements from segmented data that can be included during exam reporting as shown in Figure 2.15(c).

2.5.4 Summary Generation

As opposed to communicating individual annotations to clinicians, as done in the present clinical set-up, we propose to communicate important clinical findings by automatically highlighting regions marked by the radiologist. Using the output from our volume clustering algorithm, it is possible to extract images belonging to marked volumes from the study, distinctly color the marked regions using optimization based colorization techniques [Levin et al. 2004], and send them back

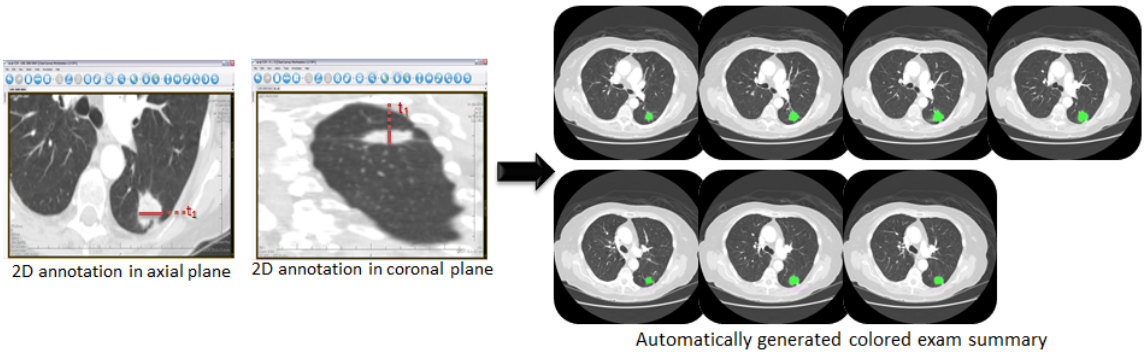


Figure 2.16: Based on the clustered volumes, key images are automatically extracted from the exam and colored to highlight the anatomy/ pathology marked by the radiologist. This summary series is pushed back to PACS as an additional series to the original exam.

to the PACS archive as an additional exam series. This has two-fold advantage. First, the software incompatibilities in handling annotations is avoided because the annotation information is embedded in the images and hence can be viewed in any PACS workstation. Second, clinicians get direct access to the annotations in the form of a concise series summarizing the study instead of having to sift through entire exams to find the marked regions. See Figure 2.16 for an example.

2.5.5 Discussion

The impetus of this work is to better utilize the information already provided in the routine radiological workflow. This allows us to avoid the need for the radiologist to use additional software outside the PACS workstation or to make redundant markup to perform segmentation. Our efforts point to an interesting chicken-and-egg issue in current radiological practice. Specifically, since volume segmentation is currently not integrated into PACS workstations, radiological protocols report volumetric data in the form of simple length measurements (i.e., line segments).

However, if volume segmentation is integrated into the PACS workstation this would influence how annotations are performed and would likely allow markup beyond basic geometric primitives. Nonetheless, we believe the work in this chapter offers a useful strategy to help leverage existing data to bootstrap segmentation. Furthermore, by demonstrating the benefits of using a standardized markup language such as AIM, we hope to further the types of value-added processing that can be exploited based on such structured information.

Chapter 3

Content-based Image Retrieval Framework for Focal Liver Lesions

3.1 Overview

As discussed in Chapter 1, clinical diagnosis today is performed on case-by-case basis. Once a pathological case is successfully diagnosed and treated, the radiological data and the diagnostic information for this case is routinely discarded. Huge amounts of such data are generated in the hospitals that can be used to design diagnostic assistants for radiologists. Content-based access to medical images has been proposed to make use of this huge data resource for supporting clinical decision making by retrieving from the image database confirmed cases that are similar to the one currently under investigation.

Figure 3.1 gives an overview of a typical content-based image retrieval (CBIR) framework. Central to a typical medical CBIR system is the image database. Regions of interest in the images are segmented and the relevant features, which could

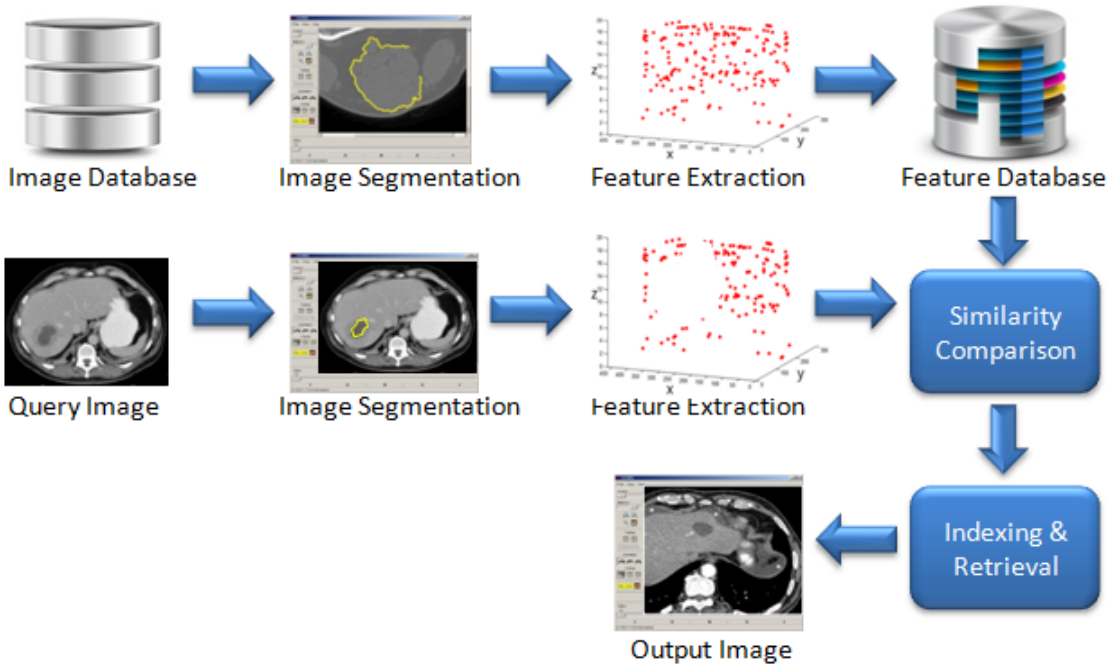


Figure 3.1: Overview of a content-based image retrieval system.

be density-based, texture-based, shape-based or other low level visual features are extracted to represent the image content. These features are then indexed and organized to form a feature database. Given a query image, in order to retrieve similar images from the database, the regions of interest from the query image are first extracted and the relevant features computed. The feature vector of the query image is then compared against the feature vectors of all the images in the database using a distance metric that measures the similarity between the query image and the images in the database. The images of the database are then ranked according to the distance of their feature vectors to that of the query image and the top matches are retrieved.

3.2 Focal Liver Lesion Characterization

Any wound, sore, ulcer, tumor or other tissue damage of an organ is termed as a "lesion." A focal lesion in the liver refers to an area of tissue damage in the liver and is identified as a region of different echogenicity, attenuation or signal intensity compared to surrounding liver parenchyma on ultrasound, CT and MR images respectively. Focal liver lesions (FLLs) can be of different pathologies. Multi-phase contrast-enhanced computed tomography is the primary imaging technique employed for the detection and characterization of FLLs [Ji et al. 2001], [Lencioni et al. 2006], [Kamel et al. 2003], [Francis et al. 2003], and [Kim et al. 2005]. It is observed that visually similar FLLs tend to belong to the same pathological category [Doi 2007] and [Muller et al. 2004]. The ability to detect and accurately characterize FLLs by qualitative visual inspection comes with years of training and experience and hence is frequently dependent on who is performing the diagnosis. CBIR systems are finding increasing use as diagnostic decision support systems. CBIR systems assist radiological diagnosis by searching and retrieving from databases of medical exams and reports confirmed cases that have image features similar to the case under investigation [Muller et al. 2004], [Akgul et al. 2011] and [Long et al. 2009].

It has been observed in clinical practice that different FLLs exhibit different visual characteristics at various time points after intravenous contrast injection. This evolution of visual features over time carry important diagnostic information and greatly influences FLL classification. Multi-phase contrast-enhanced CT procedure captures this transition by performing consecutive CT scans before and after injection of contrast. A non-contrast enhanced (NC) phase scan is usually performed before contrast injection. The patient then receives intravenous contrast injection

and three or more scans are obtained in the arterial (ART) phase (typically 25-40 seconds after start of injection), portal venous (PV) phase (60-75 seconds) and delayed (DL) phase (3-5 minutes). Diffusion of the contrast media over the different phases enhances the vessels and the lesion tissues thereby assisting in lesion type determination.

Figure 3.2 shows evolution of various lesions over different contrast phases. Liver cysts are benign fluid-filled lesions and appear as round or oval smooth edged regions with uniformly low density. Cysts do not show any enhancement after intravenous contrast injection. Hemangiomas (HEMs) are benign and typically exhibit discontinuous nodular peripheral enhancement in the ART phase with centripetal enhancement over time. Central fibrosis and calcification may sometimes be observed due to thrombosis in the vascular channels. Focal nodular hyperplasia (FNH) is a benign tumor-like mass, second only to HEM in frequency. Without any contrast the lesion is usually hypo or isodense to the liver parenchyma. FNH demonstrates bright arterial contrast enhancement except for the central scar; pronounced central arteries may be visible. In the PV phase FNH becomes isodense to liver. Metastatis (METS) on the other hand is a malignant tumor that usually spreads from other cancer affected organs. METS enhance homogeneously with contrast material administration, however, they have less well-defined margins than cysts. They typically have a band-like peripheral enhancement in ART phase and a washout in DL phase. Hepatocellular carcinoma (HCC) typically shows ART phase hyper enhancement and washout in either PV or DL phases.

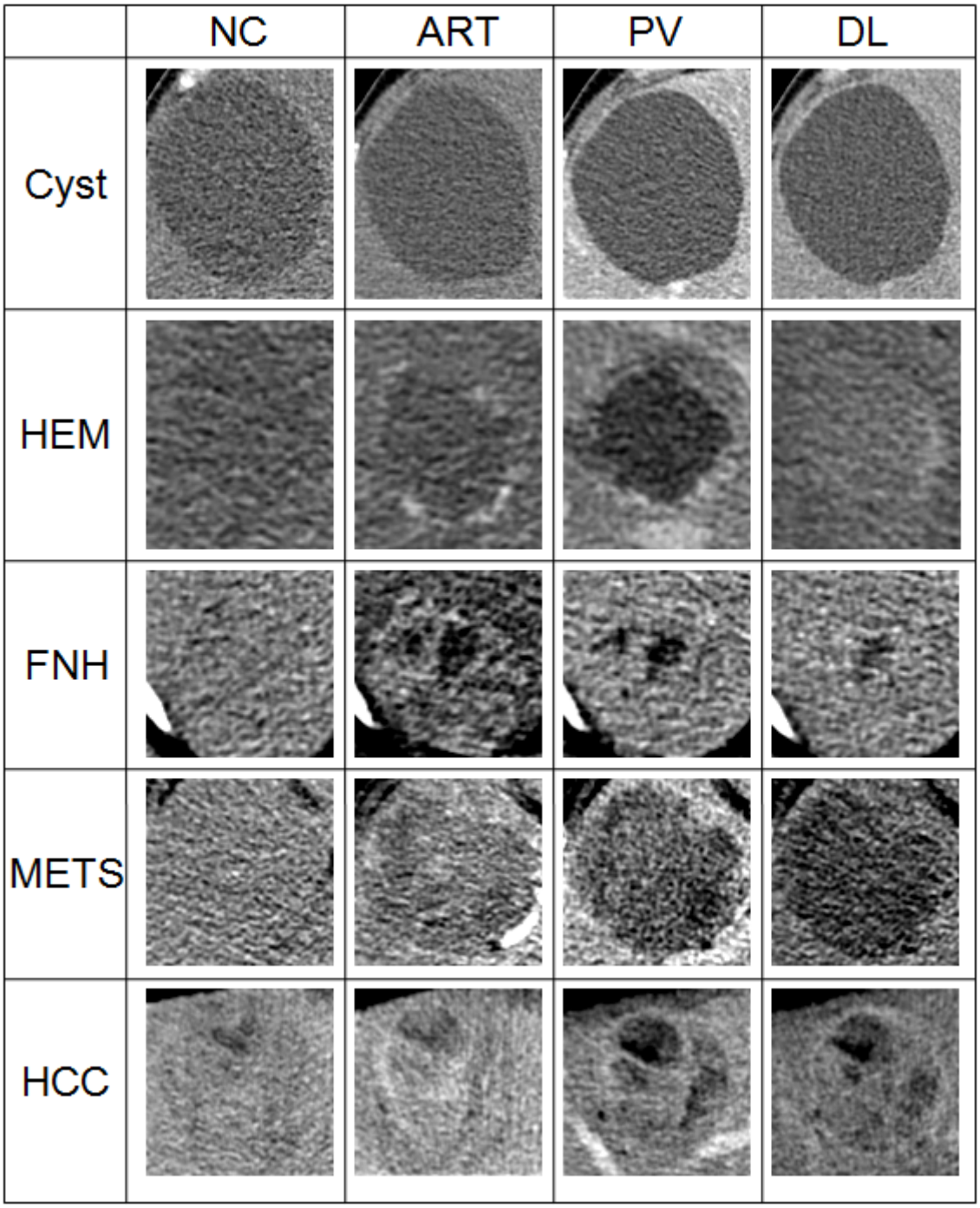


Figure 3.2: This figure shows the visual appearance of various lesions over the four phases. Images in a row are from the same lesion; cyst, hemangioma (HEM), focal nodular hyperplasia (FNH), metastatis (METS) and hepatocellular carcinoma (HCC), respectively and images in a column belong to the same contrast phase.

FLLs exhibiting similar visual appearance usually correspond to the same disease category [Doi 2007] and [Muller et al. 2004], hence radiologists can detect the lesion and recognize, to a certain extent, the pathological type of the lesion. However, accuracy in characterizing FLL usually comes with experience. Radiologists may therefore benefit from an automatic method that can provide diagnostic decision-support based on other radiologists' experience.

3.3 Related Work

Considerable research is being carried out to automate classification of liver lesions using image-based features. Some studies have reported texture-based classification of liver lesions in non-enhanced CT and ultrasonography images using techniques like neural networks [Gletsos et al. 2003] and fuzzy support vector machines [Xian 2010]. In [Mougiakakou et al. 2007], the authors provide a comprehensive performance comparison of various texture-based classifier architectures and conclude that a voting-based combination of three primary classifiers gives the best classification results. The authors in [Yu et al. 2010] developed a CBIR system to differentiate three types of hepatic lesions using global features derived from non-tensor product wavelet filter and local features based on image density and texture. However, clinical experience shows that non-enhanced CT captures limited diagnostic information. The enhancement patterns observed during various phases of contrast-enhanced images are fundamental for identifying specific focal lesions.

Some published studies have reported characterization of FLLs using multiphase features. In [Yu et al. 2012], the authors use spatially partitioned bag of

visual words (BoW) and intensity, texture and shape-based features derived from a few representative triple-phase image slices to differentiate three lesion types. The features are averaged over all phases which leads to loss of temporal enhancement information. The mean average precision of the retrieval system is reported to be 88%. In another study by the same group, the BoW-based method is improved to obtain a precision of above 90% using a different set of lesions [Yang et al. 2012]. A CBIR framework is proposed in [Chi et al. 2013b] to characterize six types of hepatic tumors using multi-phase density and texture features. The texture features are averaged over a bounding box around the tumor and tracked over multiple phases to capture their temporal evolution. A system "Bull's Eye Percentage" (BEP) score of 78% is achieved. We provide a more detailed comparison of our method with [Yu et al. 2012], [Yang et al. 2012], and [Chi et al. 2013b] in Section 3.6.

In [Costa et al. 2011], the authors use semantic features annotated by radiologists and image features derived from three orthogonal 2D planes of a single phase CT image to train a random forest classifier that distinguishes benign from malignant tumors in a retrieval framework. The framework is used to characterize sub-centimeter liver lesions. Sub-centimeter lesions are often found indistinguishable in clinical practice and hence are left unclassified though closely monitored. Further, authors in [Costa et al. 2011] neglect tumor temporal characteristics while designing their features. Napel *et al.* in [Napel et al. 2010] use high level radiological semantic features and single phase texture and boundary features to characterize three lesion types. Semantic features are unstructured subjective descriptions made by radiologists and are known to exhibit large inter-user variation. Studies show that radiologists often use different terminologies to describe the same observation in clinical routine [Sobel et al. 1996] and [Stoutjesdijk et al. 2005], hence utilization

of semantic features in image retrieval may be highly correlated to the radiological lexicon used in the particular clinical set-up.

The CBIR systems discussed thus far represent the FLLs using 2D features derived from a few representative slices of the entire exam stack. Physiologically, however, FLLs are 3D volumes. Hence, 2D features derived from a few slices is clearly an incomplete tumor representation. Further, 2D features cannot wholly capture the lesion volume, especially in cases of large and heterogenous lesions. Medical image retrieval systems based on 3D features have not been reported extensively in the literature. This is mainly due to the high computation time for 3D features, especially when retrieving high resolution datasets. In [Chi et al. 2013b] authors represent 3D liver lesions by averaging 2D texture features extracted from all the slices where the lesion is visible. However, spatial structural information interlaced within the volume is lost when considering slice-by-slice 2D features that only capture structures from the surface. Linear binary pattern (LBP) extracted from three orthogonal 2D planes have been used to approximate 3D features for fast retrieval of brain lesions in [Qian et al. 2011]. Again, by modeling the lesions using only three image slices significant part of the lesion volume is neglected. In [Burner et al. 2012], the authors use 3D LBP-based texture bags to retrieve lung lesions. Feature computation time, however, is not reported.

In this chapter we propose a fast content-based retrieval framework for FLLs based on 3D spatio-temporal features derived from 4-phase CT scans. All the features are computer generated; no radiological labels are used. The proposed retrieval framework identifies FLLs automatically and aligns the lesions in the four phases using an automated registration pipeline. Regional image-based features are computed from spatially partitioned lesion volumes and tracked over the

four phases using feature temporal derivatives. Feature similarity is then used to retrieve similar lesions from a database of confirmed cases. To the best of our knowledge this is the first study to use 3D spatio-temporal features extracted from multi-phase CT images in a CBIR framework for FLLs.

The rest of the chapter is organized as follows. In Section 3.4, we describe the evaluation database and the techniques used in the proposed retrieval framework. Section 3.5 illustrates the results and Section 3.6 provides a comparative discussion of the proposed framework with existing FLL CBIR systems. We conclude the chapter in Section 3.7.

3.4 Method

We propose a multi-phase CBIR for FLL characterization based on 3D spatio-temporal features as outlined in Figure 3.3. The proposed framework consists of two input/ output modules and three functional modules. The input to the system is the multi-phase contrast enhanced CT scan with the query FLL labeled on one phase using an automatic FLL detection method presented in [Chi et al. 2013a]. The output are the visually similar reference radiological cases. The functional modules perform multi-phase image registration, focal lesion representation and similarity assessment. Non-rigid B-spline based registration [Yushkevich et al. 2006] is employed for alignment of the FLL in the four phases of the CT scan. The FLL is quantitatively represented using 3D spatio-temporal features extracted from various regions within the FLL volume of interest (VOI). A FLL database is constructed using the resulting feature vectors and the corresponding clinical diagnosis. A L2-norm similarity measure between feature vectors of the query

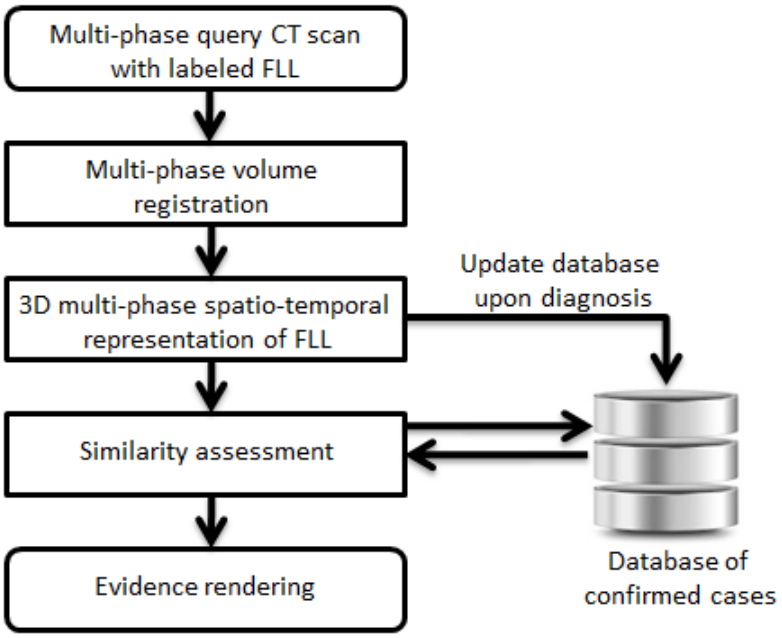


Figure 3.3: Outline of the proposed FLL content-based retrieval framework.

lesion and lesions in the database is used for retrieval. The retrieved results are ranked on the basis of similarity score and presented as evidential support to the radiologists.

3.4.1 Image Database

Institutional review board approval was obtained for retrospective analysis of 4-phase contrast-enhanced CT images of 30 de-identified patients. CT scans were acquired using a 64-detector SOMATOM sensation scanner (Siemens Medical Solutions, Forchheim, Germany) via a standard 4-phase contrast-enhanced imaging protocol with a slice collimation of $0.6\,mm$, a matrix of 512×512 pixels and an in-plane resolution of $0.59 - 0.78\,mm$. The raw data was reconstructed at an isotropic resolution of $0.6 \times 0.6 \times 0.6\,mm^3$. The evaluation database was constructed using 44

confirmed lesions identified in the 30 patients. The 44 lesions consisted of five types of lesions: cyst, HEM, FNH, METS, and HCC. There were 14 cases of cyst, 10 cases of HEM, 5 cases of FNH, 11 cases of METS and 4 cases of HCC in the 44 confirmed lesions. One representative lesion was identified in each patient for analysis. The pathology type of the lesions were confirmed based on clinical features, CT scans, data from other imaging modalities and biopsy, wherever needed.

3.4.2 Focal Liver Lesion Identification

We use a hybrid generative-discriminative method proposed in [Chi et al. 2013a] to detect FLLs in a 3D image. The method first uses a generative model to represent non-lesion components such as the healthy liver parenchyma and the enhanced liver vasculature. The candidate FLLs are then identified within the liver volume by eliminating these non-lesion areas. False positives among the identified candidate FLLs are then suppressed using a discriminative approach that uses a lesion-likelihood measure comprising of three shape-based features: spherical symmetry, compactness and size. All the detected FLLs are presented to an expert who then selects one for further processing.

3.4.3 4-phase Lesion Alignment

As mentioned in Section 3.2, temporal evolution of FLL over different phases carry rich diagnostic information. However, FLLs do not appear visually distinct in all the phases hence making their identification difficult in all phases. A FLL is typically detected in the phase in which it shows highest contrast with respect to the liver parenchyma and is localized in the other phases using a non-linear B-spline

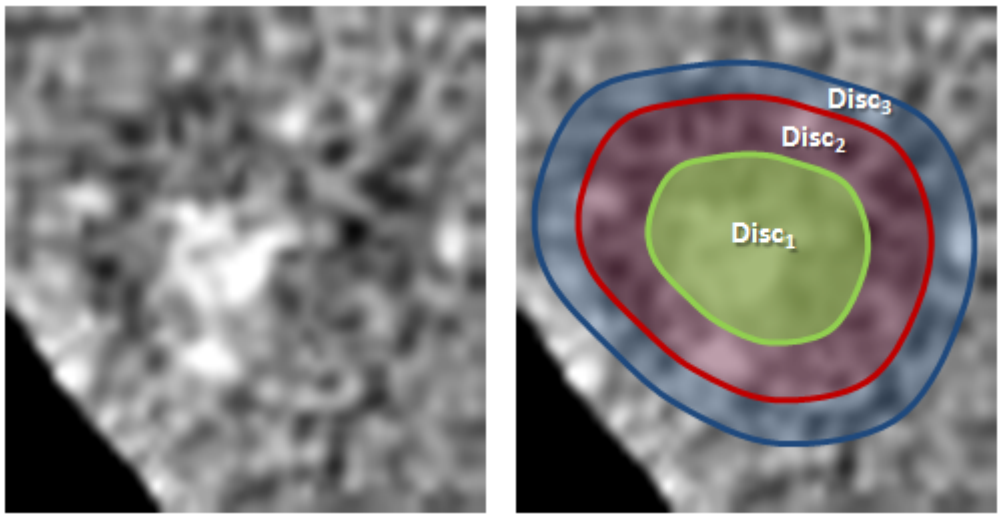


Figure 3.4: This figure shows a central calcification inside a HEM. To capture the spatial tissue characteristics we partition the lesion into three concentric discs.

registration [Yushkevich et al. 2006].

3.4.4 3D Spatio-Temporal Feature Design and Extraction

Visual patterns generated by intrusion of the contrast agent into the tumor volume over time is the primary tool used for tumor differentiation in clinical practice. Spatial visual characteristics such as ring enhancement, nodule-within-a-nodule enhancement, pseudocapsule, true and pseudo central scars, peripheral washout are fundamental to identifying specific focal lesions [Elsayes et al. 2005]. Inspired from these features used in clinical decision making, we design spatio-temporal features to model the tumor by dividing the VOI into three volumetric discs and extracting features from these discs over the four phases. An example of volumetric partitioning is shown in Figure 3.4. The innermost disc, $Disc_1$, captures central enhancement characteristics caused by structures such as the central scar, fibrosis, calcification, necrosis, if any. The intermediate disc, $Disc_2$, models the tumor tissue

characteristics and the outermost disc, $Disc_3$ is designed to represent features and the enhancement pattern of the tumor boundary.

We use a standard distance transformation technique based on Euclidean distance to partition the tumor VOI [Rosenfeld and Pfaltz 1996]. Distance transformation converts a binary volume into a gray scale volume. The binary volume in our case is the tumor VOI where voxels inside the tumor form the foreground and the rest comprise the background. The tumor is assumed to be segmented either manually or using existing tumor segmentation methods [Park et al. 2005] and [Massoptier and Casciaro 2008]. Distance transformation of this binary volume results in a gray scale volume where each voxel of the gray scale volume represents the distance of that voxel from the closest background voxel in the binary volume. Voxels in the gray scale volume are then grouped into three discs based on these distance values. Typically a tumor is clearly visible in at least one of the four phases. Segmentation is applied on the phase in which the tumor is most clearly visible. Since all the four phases are registered, the same segmentation can be used to delineate the tumor in other phases.

An additional benefit of tumor partitioning is the computational speed up. Image features from each partition can now be computed in parallel. In effect the tumor is now partitioned into three smaller sub-volumes and the computation time is governed by the largest of these sub-volumes instead of the entire tumor. In Section 3.5.2 we discuss in detail the speed up and enhancement in retrieval performance achieved by tumor partitioning. Large computation time is the primary reason why 3D feature-based retrieval systems do not find use in the clinical routine. As a measure to accelerate tumor processing time, instead of extracting features from all voxels within the tumor partitions we perform a uniform sub-

sampling of the voxels and use only the selected samples for feature computation. In Section 3.5.3 we provide a detailed analysis of how the processing time and retrieval performance vary with various amounts of sub-sampling.

Post partitioning various features such as those based on shape, user-supplied semantics, texture, and intensity can be extracted from the partitions to model the tumor. Shape features are important in modeling liver lesions. Slowly growing benign lesions are often well circumscribed or encapsulated, while rapidly growing malignant lesions tend to have an indistinct irregular shape and are not encapsulated. Shape features are good at discriminating benign lesions form malignant lesions. In this work we aim to characterize five types of FLLs that include three types of benign lesions and two types of malignant lesions. Shape features have insufficient power to differentiate among benign lesions, or among malignant lesions [15]. In addition, FLLs are in general most distinctively observed only in one of the phases, thus, accurate shape features can only be extracted from a single-phase. They are basically single-phase features. Since we focus on features that have spatial and temporal characteristics that evolve over the four phases of the CT scan, we do not include shape features in our work. User-supplied semantic labels are subjective and often unstructured description of the tumor characteristics that have high inter-user variability. Hence, we do not want to use semantic features. Image texture is widely used in the literature to model tumor tissues. Methods that model texture are broadly categorized into statistical and structural approaches. Statistical approaches such as histogram of pixel gray levels and gray level co-occurrence have been found to work best with images that have stochastic micro-textures as against the structural approaches of textons, wavelet transforms and Gabor filters which compute weighted mean of pixel neighborhoods and hence eliminate finer

textural details [Qian et al. 2011] and [Maenpaa and Pietikainen 2005]. In this work, we use density and gray level co-occurrence-based texture features derived from the volumetric partitions and track their temporal evolution over the four phases. More specifically, we define four 3D feature vectors to model a tumor as defined below.

Density Feature

The density feature, F_1, represents the ratio of average density inside the discs to the average density of liver parenchyma. F_1 measures lesion enhancement with respect to the surrounding liver tissues and is defined as:

$$F_1 = \{D^{\mathrm{NC}}, D^{\mathrm{ART}}, D^{\mathrm{PV}}, D^{\mathrm{DL}}\}, \tag{3.1}$$

where $D^{\mathrm{NC}} = \{d^{\mathrm{NC}}_{\mathrm{Disc}_1}/d^{\mathrm{NC}}_{\mathrm{liver}}, d^{\mathrm{NC}}_{\mathrm{Disc}_2}/d^{\mathrm{NC}}_{\mathrm{liver}}, d^{\mathrm{NC}}_{\mathrm{Disc}_3}/d^{\mathrm{NC}}_{\mathrm{liver}}\}$. The variable $d^{\mathrm{NC}}_{\mathrm{Disc}_i}$ measures the average density inside Disc_i in the NC phase and the variable $d^{\mathrm{NC}}_{\mathrm{liver}}$ is the average density of the healthy liver tissue in the NC phase. Variables D^{ART}, D^{PV} and D^{DL} are defined in a similar fashion. The resulting density features obtained from all phases are arranged into a 12-dimensional density feature vector representing the FLL. The density feature aims to capture contrast enhancement and washout. For example, if a lesion has $|D^{\mathrm{ART}}| > |D^{\mathrm{NC}}|$ and $|D^{\mathrm{PV}}| < |D^{\mathrm{NC}}|$, then the lesion is enhanced in the ART phase due to contrast propagation and has a washout in the PV phase.

Temporal Density Feature

The temporal density feature, F_2, measures temporal enhancement of the lesion in ART, PV and DL phases with respect to the NC phase. It is defined as:

$$F_2 = \{TD^{\text{ART/NC}}, TD^{\text{PV/NC}}, TD^{\text{DL/NC}}\}, \tag{3.2}$$

where $TD^{\text{ART/NC}} = \{td^{\text{ART/NC}}_{\text{Disc}_1}, td^{\text{ART/NC}}_{\text{Disc}_2}, td^{\text{ART/NC}}_{\text{Disc}_3}\}$ and for $i = 1, 2, 3$

$$td^{\text{ART/NC}}_{\text{Disc}_i} = \frac{d^{\text{ART}}_{\text{Disc}_i} - d^{\text{NC}}_{\text{Disc}_i}}{d^{\text{NC}}_{\text{Disc}_i}}, \tag{3.3}$$

$$td^{\text{PV/NC}}_{\text{Disc}_i} = \frac{d^{\text{PV}}_{\text{Disc}_i} - d^{\text{NC}}_{\text{Disc}_i}}{d^{\text{NC}}_{\text{Disc}_i}}, \tag{3.4}$$

$$td^{\text{DL/NC}}_{\text{Disc}_i} = \frac{d^{\text{DL}}_{\text{Disc}_i} - d^{\text{NC}}_{\text{Disc}_i}}{d^{\text{NC}}_{\text{Disc}_i}}. \tag{3.5}$$

Similar definitions follow for $TD^{\text{PV/NC}}$ and $TD^{\text{DL/NC}}$. Temporal density features from the ART, PV and DL phases are encoded into a 9-dimensional feature vector to model tumor temporal enhancement.

Texture Feature

We use a 3D gray level co-occurrence matrix (GLCM) to quantify the gray tone distribution in the tumor sub-volumes. GLCM is an estimation of the joint probability distribution of a pair of gray level voxels. An element $G_{(\theta,d)}(i, j)$ of the GLCM matrix is the probability of the occurrence of gray levels i and j at distance of d from each other along the direction θ. The variables i and j can vary from 1 to N, where N is the number of gray levels in the volume. In 3D, θ can take 26 values

resulting from linking a voxel to each of its 26 nearest neighbors. Since directions that are 180° apart result in the same co-occurrence matrix, we only consider 13 unique directions. Given an offset d, we compute GLCM over all 13 directions and use the average to make the texture rotation invariant. We vary d and choose the value that gives the best retrieval results as described in Section 3.5.

Six texture coefficients: energy, entropy, inverse difference moment, inertia, cluster shade and correlation as defined in [Haralick et al. 1973] are derived from the rotation invariant GLCM. The texture feature, F_3, is composed as follows:

$$F_3 = \{T^{\text{ART}}, T^{\text{PV}}, T^{\text{DL}}\}, \tag{3.6}$$

where $T^{\text{ART}} = \{T^{\text{ART}}_{\text{Disc}_1}, T^{\text{ART}}_{\text{Disc}_2}, T^{\text{ART}}_{\text{Disc}_3}\}$ represents the texture features derived from the three discs in the ART phase. $T^{\text{ART}}_{\text{Disc}_i} = \{t^{\text{ART}}_{1_{\text{Disc}_i}}, \ldots, t^{\text{ART}}_{6_{\text{Disc}_i}}\}$ where $t^{\text{ART}}_{k_{\text{Disc}_i}}$'s, $k = \{1, \ldots, 6\}$, are computed as defined in Table 3.1 from the GLCM of Disc_i in the ART phase. Similar definition applies for T^{PV} and T^{DL}. The resulting texture coefficients from nine discs in ART, PV and DL phases are arranged into a 54-dimensional feature vector which encodes the tumor texture.

Temporal Texture Feature

Similar lesions are known to show similar enhancement patterns. We aim to quantify this by defining temporal texture, F_4, as the normalized difference in texture features at the three enhancement phases ART, PV and DL with respect to the NC phase. F_4 is defined as:

$$F_4 = \{TT^{\text{ART}}, TT^{\text{PV}}, TT^{\text{DL}}\}, \tag{3.7}$$

Table 3.1: This table describes the texture coefficients derived from the GLCM matrix. The term $g(i, j)$ represents the joint probability density of the gray level pair (i, j).

Texture Coefficient	Expression and Qualitative Analysis
Energy	$t_1 = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} g(i,j)^2$ Energy quantifies the repetition of gray level pairs in an image.
Entropy	$t_2 = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} g(i,j) log_2(g(i,j))$ Entropy represents the randomness in the image.
Inverse Difference Moment	$t_3 = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{1}{1+(i-j)^2} g(i,j)$ Inverse difference moment measures the local homogeneity in the image.
Inertia	$t_4 = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i-j)^2 g(i,j)$ Inertia gauges local variations in an image.
Cluster Shade	$t_5 = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i+j-\mu_i-\mu_j)^3 g(i,j)$ where, $\mu_i = \sum_{i=0}^{N-1} i \sum_{j=0}^{N-1} g(i,j)$, and $\mu_j = \sum_{i=0}^{N-1} j \sum_{j=0}^{N-1} g(i,j)$. Cluster shade quantifies perceptual uniformity and proximity.
Correlation	$t_6 = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{(i-\mu_i)(j-\mu_j)g(i,j)}{\sigma_i \sigma_j}$ where, $\sigma_i = \sum_{i=0}^{N-1} (i-\mu_i)^2 \sum_{j=0}^{N-1} g(i,j)$, and $\sigma_j = \sum_{j=0}^{N-1} (j-\mu_j)^2 \sum_{i=0}^{N-1} g(i,j)$. Correlation assesses the linearity of relationship between various gray level pixel pairs.

where, $TT^{\text{ART}} = \{TT^{\text{ART}}_{\text{Disc}_1}, TT^{\text{ART}}_{\text{Disc}_2}, TT^{\text{ART}}_{\text{Disc}_3}\}$ is the temporal texture in the ART phase formulated as derivative of the six texture coefficients in each disc in the ART phase; $TT^{\text{ART}}_{\text{Disc}_i} = \{tt^{\text{ART}}_{1_{\text{Disc}_i}}, \ldots, tt^{\text{ART}}_{6_{\text{Disc}_i}}\}$, $i = \{1, 2, 3\}$. Derivative of each texture coefficient

is defined as:

$$tt^{\mathrm{ART}}_{k_{\mathrm{Disc}_i}} = \frac{t^{\mathrm{ART}}_{k_{\mathrm{Disc}_i}} - \underset{\mathcal{P}\in\{\mathrm{ART,PV,DL}\}}{\mathrm{median}}\, t^{\mathcal{P}}_{k_{\mathrm{Disc}_i}}}{\underset{\mathcal{P}\in\{\mathrm{ART,PV,DL}\}}{\max}\, t^{\mathcal{P}}_{k_{\mathrm{Disc}_i}} - \underset{\mathcal{P}\in\{\mathrm{ART,PV,DL}\}}{\min}\, t^{\mathcal{P}}_{k_{\mathrm{Disc}_i}}} \tag{3.8}$$

for $k = 1, \ldots, 6$. Texture derivative in PV and DL phases, TT^{PV} and TT^{DL} respectively, are formulated analogously with the individual texture coefficient derivatives defined as:

$$tt^{\mathrm{PV}}_{k_{\mathrm{Disc}_i}} = \frac{t^{\mathrm{PV}}_{k_{\mathrm{Disc}_i}} - \underset{\mathcal{P}\in\{\mathrm{ART,PV,DL}\}}{\mathrm{median}}\, t^{\mathcal{P}}_{k_{\mathrm{Disc}_i}}}{\underset{\mathcal{P}\in\{\mathrm{ART,PV,DL}\}}{\max}\, t^{\mathcal{P}}_{k_{\mathrm{Disc}_i}} - \underset{\mathcal{P}\in\{\mathrm{ART,PV,DL}\}}{\min}\, t^{\mathcal{P}}_{k_{\mathrm{Disc}_i}}}, \tag{3.9}$$

$$tt^{\mathrm{DL}}_{k_{\mathrm{Disc}_i}} = \frac{t^{\mathrm{DL}}_{k_{\mathrm{Disc}_i}} - \underset{\mathcal{P}\in\{\mathrm{ART,PV,DL}\}}{\mathrm{median}}\, t^{\mathcal{P}}_{k_{\mathrm{Disc}_i}}}{\underset{\mathcal{P}\in\{\mathrm{ART,PV,DL}\}}{\max}\, t^{\mathcal{P}}_{k_{\mathrm{Disc}_i}} - \underset{\mathcal{P}\in\{\mathrm{ART,PV,DL}\}}{\min}\, t^{\mathcal{P}}_{k_{\mathrm{Disc}_i}}}. \tag{3.10}$$

Texture derivatives computed for the three enhanced phases are organized into a 54-dimensional temporal texture feature vector that represents the textural evolution of the tumor. The four feature vectors F_1, F_2, F_3 and F_4 form the FLL model.

Since we use GLCM-based texture features in this work, it turns out that we can further improve tumor processing speed by reducing the number of gray levels used while populating the GLCM matrix. Each element (i, j) of the GLCM matrix measures the probability of joint occurrence of gray level pairs i and j. Computing 3D GLCM in 13 directions using the original CT values is highly expensive both computationally and in terms of memory requirement. We quantize down the original CT values to fewer distinct gray levels in order to reduce the size and computation time of the GLCM. Optimum number of gray levels can be determined experimentally. Section 3.5 provides analysis of gain in computation time versus retrieval performance for various gray level counts.

3.4.5 Similarity Assessment and Evidence Rendering

Once the FLL feature vectors are constructed, similarity between a query FLL and the model FLLs in the database can be measured using a L^2 distance between the respective feature vectors. Distance between two lesions FLL_1 and FLL_2 in L^2 is defined as:

$$\text{D}_{L^2}(\text{FLL}_1, \text{FLL}_2) = \sum_{i=1}^{4} w_i \| F^i_{\text{FLL}_1} - F^i_{\text{FLL}_2} \|_{L^2}. \tag{3.11}$$

The term $F^i_{\text{FLL}_1}$ represents the i^{th} feature vector of FLL_1 where i iterates over density, temporal density, texture and temporal texture feature vectors and w_i is the respective weight. Weight selection is elaborated in Section 3.5.

Model FLLs in the database are sorted in increasing order of their distance to the query FLL and the closest matching FLLs are rendered to the radiologist. It is also possible to predict the pathological type of the query FLL using BEP. BEP is defined for each query as the percentage of correct retrievals with respect to the query FLL's class within the top $2C$ results where C is the size of the query FLL's class [Manjunath 2002]. The query FLL is predicted to belong to the class that has the highest BEP score as follows:

$$\text{Query} \subseteq C_i$$

$$\text{if, } \text{BEP}(C_i) = \max_{k=1,2,\ldots,5} (\text{BEP}(C_k)) \tag{3.12}$$

where C_k represents the k^{th} class of FLL pathology in the database. The term $\text{BEP}(C_i)$ represents the BEP score when query FLL is assumed to belong to class C_i. The distance of the query FLL to a class C_k can be computed using average distance to model FLLs belonging to class C_k retrieved within the top $2|C_k|$ results

as formulated below:

$$\text{Distance}\,(C_k) = \frac{1}{\mathrm{N}_{C_k}} \sum_{i=1}^{\mathrm{N}_{C_k}} \mathrm{D}_{L^2}(\mathrm{FLL}_{\mathrm{Query}}, \mathrm{FLL}_i), \tag{3.13}$$

where N_{C_k} is the number of FLLs belonging to class C_k retrieved in the top $2|C_k|$ results.

Although a CBIR system can predict the pathological type of an unknown lesion, the reason why CBIR systems find high acceptance in the clinical routine is their capability of providing evidential support in favor and also against its prediction. It is important for radiologists to not only look at examples of similar lesions from the same pathology type but also refer to visually similar lesions belonging to a different class of pathology for an informed diagnosis. For most queries, the CBIR system renders more than one type of pathology in the top retrieved results thereby making the radiologist aware of possible differential diagnosis.

3.5 Experiments and Results

The proposed CBIR framework is evaluated on a database of 44 FLLs identified in 30 patients and comprising of 5 pathological types. One representative FLL is chosen from each patient for analysis. Precision-recall curve and BEP score are used to evaluate the retrieval performance of the proposed framework. Precision is defined as the ratio of retrieved lesions that belong to the query class with respect to the total number of lesions retrieved and recall is defined as the ratio of number of retrieved lesions that belong to the query class with respect to all model lesions in the database that belong to the query class. Leave-one-out cross validation scheme

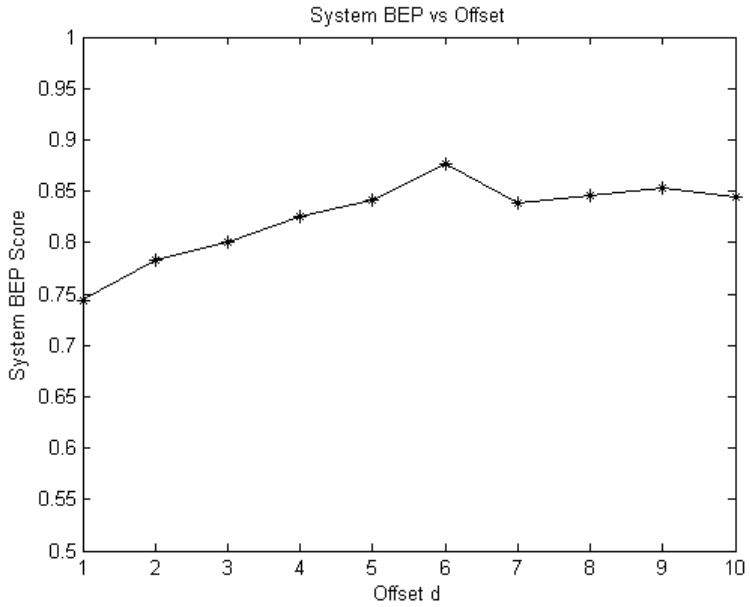


Figure 3.5: This figure plots the system BEP score for various values of offsets. The offset, d, is the distance between gray level pairs used for computing GLCM entries. The BEP score is observed to be higher for higher values of d.

is used to compute the precision-recall curves and the BEP scores.

3.5.1 Parameter Optimization

In this section we describe selection of offset (distance between gray level pairs for GLCM computation) and feature weights (for inter-lesion comparison) respectively.

Offset

We compare the retrieval performance at various values of offsets, d, by computing texture and temporal texture features from GLCM at $d = 1, 2, \ldots, 10$, and measuring the corresponding system BEP scores (Figure 3.5). Higher offsets produce better results, albeit, using a smaller subset of the dataset. Large offsets cannot be used to

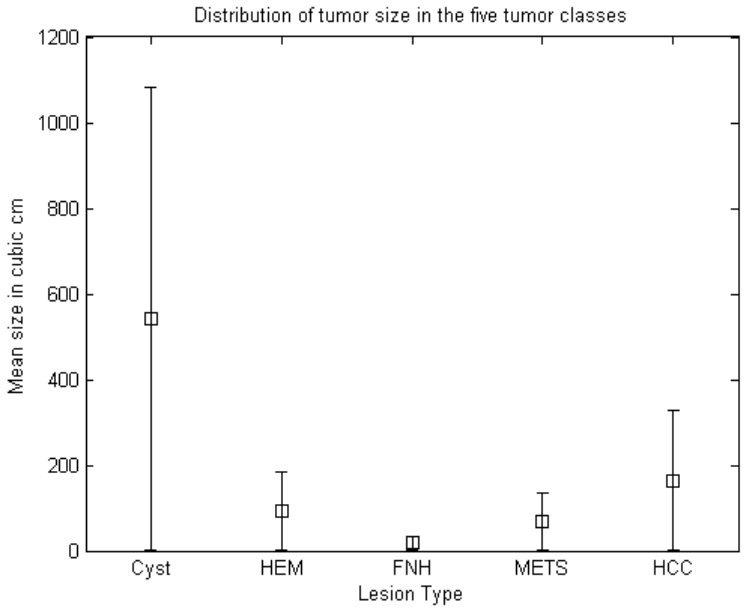


Figure 3.6: This plot shows the variation in tumor volume (in cm^3) for the five tumor classes in the database.

model small tumors. In Figure 3.6 we plot tumor size distribution for the five classes of lesions in our database. From experiments we observe that a maximum offset of four is able to model all the tumors in our database and hence for subsequent analysis we set $d = 4$.

Feature Weights

Similarity between two lesions is assessed using a weighted L^2 difference between the respective feature vectors as formulated in Equation (3.11). To compute the optimum weights, we start with identical weights for density, temporal density, texture and temporal texture features and perform a systematic greedy search for the weights that maximize the total system BEP score. It is found that a weight vector of [0.3 0.3 0.2 0.2] generates the best results. Figure 3.7 compares precision-

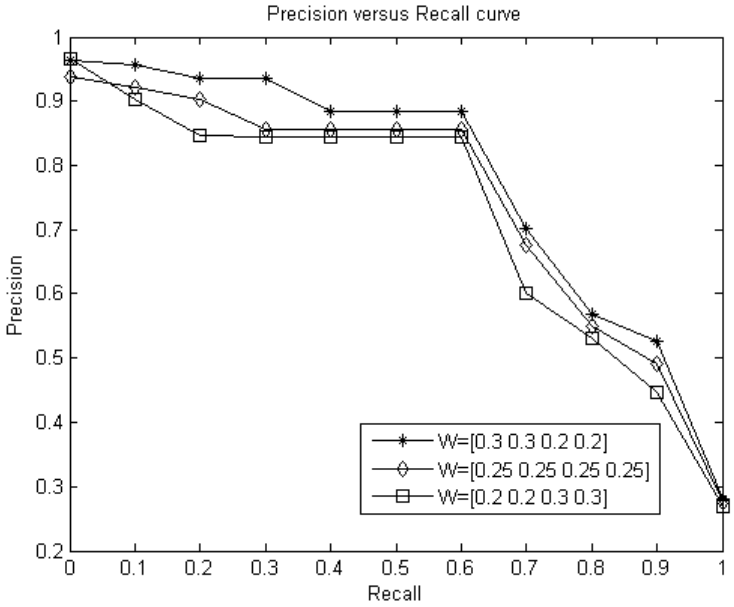


Figure 3.7: This figure plots precision versus recall curves for different feature weight vectors. Precision-recall curves for optimal and equal weight vectors are observed to be close.

recall curves for various weight vectors including the optimal and equal weight vectors.

3.5.2 Tumor Partitioning

As mentioned in Section 3.4, we partition the tumor into discs and extract features from each disc to capture the spatio-temporal characteristics of the tumor. Figure 3.8 shows the gain in retrieval performance obtained by partitioning the tumor into three discs against the case when tumors are represented by features extracted from the whole VOI. Retrieval performance post-partitioning is clearly superior to the non-partitioned case. Further, experimentally we found that an exact segmentation of the tumor is not needed. This is because the features are averaged over all the

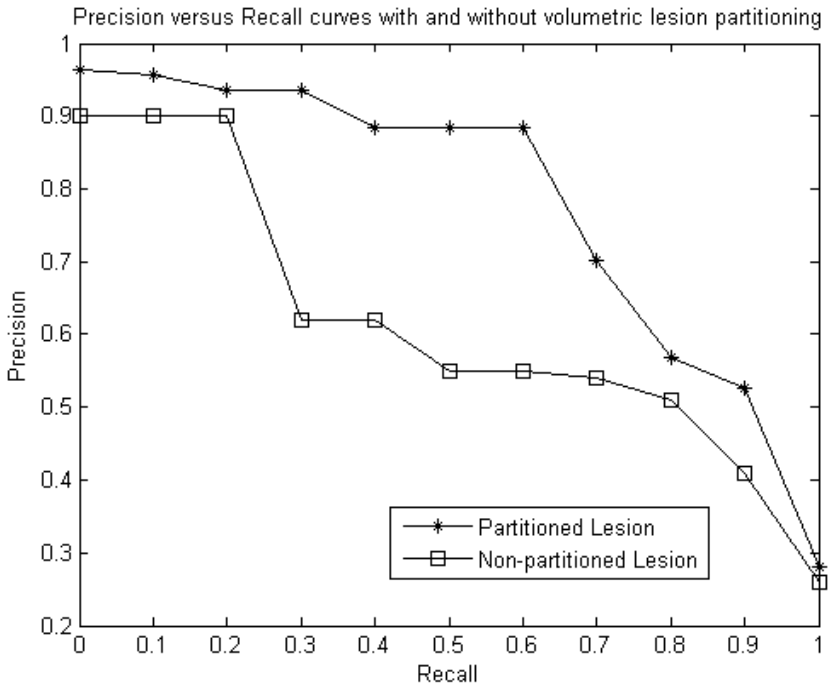


Figure 3.8: This figure compares precision-recall curves when the lesions in the database are volumetrically partitioned into three sub-volumes versus when they are not. The retrieval performance obtained by non-partitioned lesions is found to be inferior to that obtained by partitioned lesion representation.

voxels inside the respective discs.

3.5.3 Retrieval Performance and Processing Speed

Figure 3.8 plots the retrieval performance of the proposed retrieval framework in terms of precision and recall. The system's precision remains above 0.85 till a recall of 0.6. The BEP score for the five lesion pathologies is tabulated in Table 3.2. The global mean score of 82.6% demonstrates good discriminatory properties of the 3D spatio-temporal features. A more detailed examination of the results shows excellent BEP scores, between 87% and 100%, for cyst, METS and HCC. This can be

Table 3.2: This table enlists the Bull's Eye Percentage for various lesion classes.

Lesion Class	Bull's Eye Percentage Score
Cyst	0.87
HEM	0.62
FNH	0.70
METS	0.94
HCC	1.00

contributed to the markedly different temporal enhancement of these three lesion types. HEM and FNH, however, report lower BEP scores. $Disc_3$, which captures the peripheral enhancement, tends to show similar enhancement in ART phase for both HEM and FNH. Further, $Disc_1$ also shows similar temporal washout in the DL phase for both HEM and FNH due to the occasional presence of a central scar in FNH. This may explain why lower scores are obtained for HEM and FNH. FNH is difficult to detect and it is well known clinically that they are called "stealth lesions" if the ART phase enhancement is not well demonstrated.

Figure 3.9 shows the top retrievals for five query lesions, one from each lesion class. Experiments show that for 98% query lesions, at least one lesion of the same pathological type as the query lesion is rendered in the top two retrieval results.

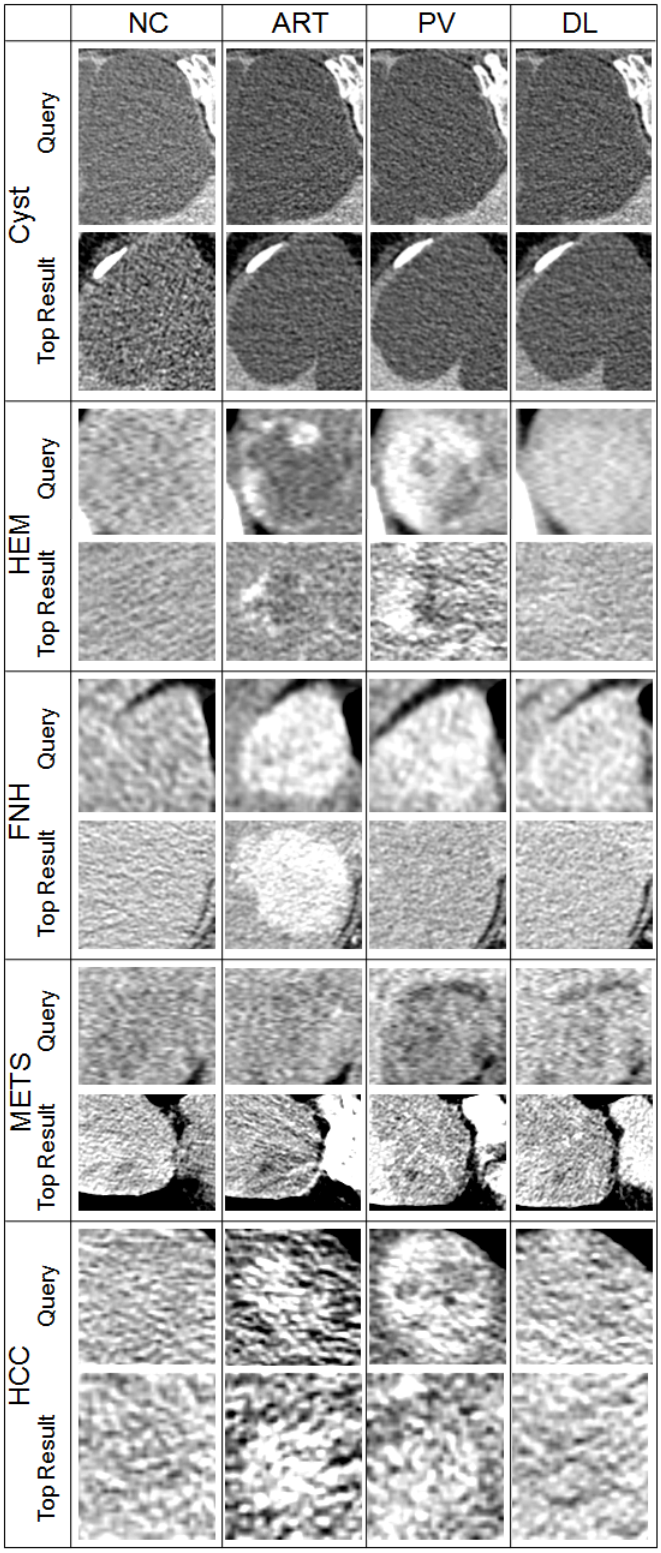


Figure 3.9: This figure shows the top retrieval results for five query lesions, one from each of the five lesion classes.

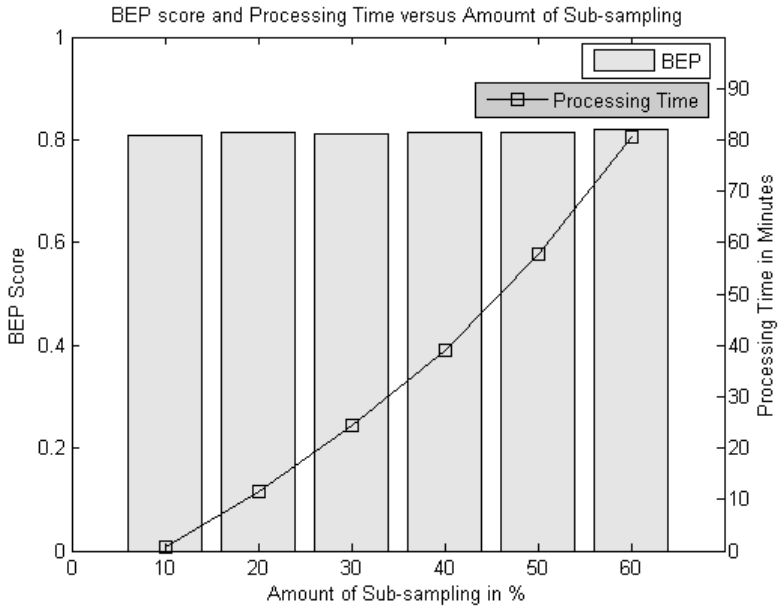


Figure 3.10: This figure plots the BEP scores and the processing times for various amounts of volumetric sub-sampling.

Low query processing time is critical to make a system clinically viable in addition to high retrieval accuracy. Various characteristics of our feature extraction framework contribute to accelerating query processing time. Tumor partitioning is the first contributor to computation acceleration. By partitioning the tumor into smaller sub-volumes we can process all the sub-volumes concurrently in a multi-core computing framework. This leads to a reduction in the processing time by one third compared to the case when the tumor is processed as a whole.

Further, we perform sub-sampling of the sub-volumes instead of using each voxel for feature computation. Figure 3.10 shows how the retrieval performance in terms of system BEP score and the total feature computation time for *all* lesions in the database vary with varying amounts of sub-sampling. The computation time is measured using MATLAB R2011b without any GPU acceleration in an Intel Xeon

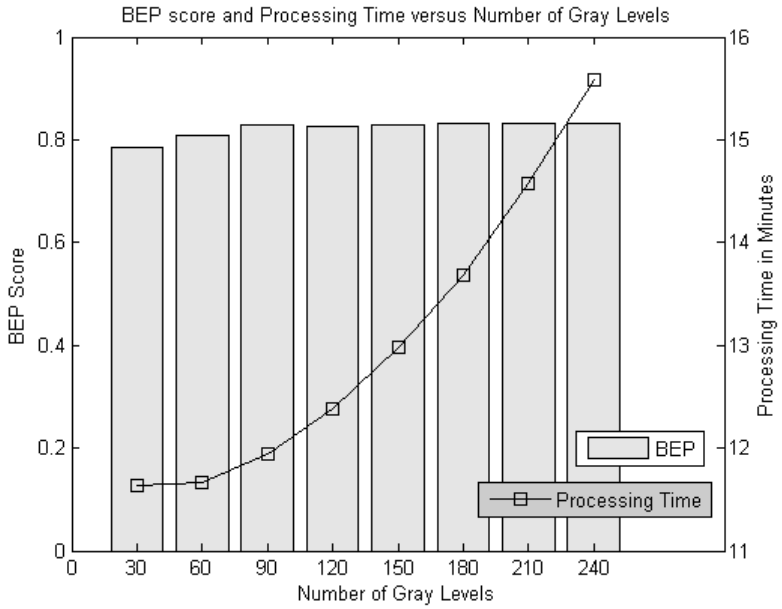


Figure 3.11: This figure plots the BEP scores and the processing times for various counts of distinct gray levels.

2.4 GHz 4 core processor with 6 GB RAM. We observe that there is no considerable increase in BEP score with additional sampling, however, gain in speed up is substantial when lower number of voxels are sampled. We use 25% sub-sampling to compute features.

Use of GLCM for texture computation gives us another parameter to gain additional speed up, namely, the number of distinct gray levels used for GLCM computation. Larger the number of distinct gray levels, bigger is the GLCM matrix and hence slower is the computation. We quantize down the original CT gray levels to a lower number of distinct values and study its effect on the computation time and the retrieval accuracy. Figure 3.11 plots the system BEP score versus number of gray levels used for feature computation. Total time taken to compute 3D features for all lesions in the database is also plotted against the number of gray levels. As

Table 3.3: This table compares the processing times for some FLLs when tumor partitioning, volumetric sub-sampling and gray level quantization are used to accelerate feature computation versus when no acceleration is used. For acceleration we use 25% sub-sampling and 60 gray levels.

FLL Size (in cm^3)	Processing time with acceleration (in min.)	Processing time without without acceleration (in min.)
69.6	0.08	2.41
133.6	0.19	5.34
184.1	0.26	7.66
286.4	0.33	8.95
328.1	0.46	12.84

expected, higher number of gray levels increases the computation time, however, the gain in performance saturates after certain gray levels. We use 60 gray levels for feature extraction.

Combining acceleration due to tumor partitioning, volumetric sub-sampling and gray level quantization, Table 3.3 compares the total processing time for some FLLs with and without computation acceleration. On average the computation time with acceleration is more than 28 times faster than when no acceleration is used.

3.6 Discussion

3.6.1 System Comparison

In this chapter we propose a retrieval framework for FLL characterization using 3D image-based spatio-temporal features. No FLL CBIR systems based on 3D features have been reported in the literature. The closest related works have studied lesion

retrieval based on 2D features derived from representative slices of single or multi-phase CT images [Yu et al. 2012], [Yang et al. 2012], [Chi et al. 2013b], [Costa et al. 2011],and [Napel et al. 2010]. In the following paragraphs we compare the proposed system with these prior studies in detail.

Yu *et al.* in [Yu et al. 2012] propose retrieval of three types of FLLs–cyst, HEM, and HCC–using BoW and 2D image-based features. The lesions are spatially partitioned and BoW histograms are computed for each partition. The visual vocabulary for BoW histogram is constructed using image patches of the training lesions without normalization. Additionally 93 image-based global features are constructed from the un-partitioned tumor region of interest based on intensity, GLCM, Gabor filter and tumor shape. The lesion is represented by averaging spatial BoW and global image-based features across multiple phases. A mean average retrieval precision of 88% is reported. In an extension to this work, the authors in [Yang et al. 2012] eliminate lesion partitioning and use distance metric learning methods to compute similarity between global BoW histograms and report an average precision of above 90% when evaluated on a database of cyst, HEM and hepatomas. Processing time, however, is not reported. In [Yu et al. 2012] lesion spatial-partitioning is used, though only to construct the BoW histograms. For other image-based features no spatial information is preserved. Further, averaging BoW and image features over multiple phases leads to loss of temporal information. In [Yang et al. 2012] both spatial and temporal information is neglected for BoW and image features. Pathologically, different lesions may appear visually similar in some phases. By combining features from various phases, a good correspondence between lesions in sequential phases is not guaranteed. Similar sequential evolution of two lesions is essential for them to be categorized to the same patho-

logical type. Further, from experiments we observe that computing GLCM texture features over the whole tumor without acceleration is slow. Research shows that computing texture using Gabor filter is even slower than GLCM [Qian et al. 2011]. We propose a much simpler modeling of tumors using density and only six GLCM-based texture coefficients that preserve both spatial and temporal characteristics of the tumor and are also faster to compute as against the elaborate modeling of tumors proposed in [Yu et al. 2012] and [Yang et al. 2012] using 93 features and BoW learning. When evaluated using cyst, HEM and HCC, our system achieves a higher mean average precision of 92.4%.

In [Chi et al. 2013b], the authors design a FLL retrieval framework using GLCM-based 2D temporal features derived from multi-phase CT images. The features are, however, derived by averaging density and texture over a tumor bounding box. The system is tested on a database of 69 FLLs comprising of six pathological types. A BEP score of 78% is obtained. It is reported that for 90% query lesions, the processing time is more than 10 minutes. 2D GLCM captures the joint probability distribution of gray level pairs in only four directions: 0°,45°,90°, and 135°. However, a 3D GLCM represents gray level distribution in 13 directions along the 13 neighbors of a voxel. Averaging 2D features over multiple slices does not accurately approximate the 3D texture. Further, by averaging features over the whole bounding box, the authors dismiss the spatial enhancement characteristics of the tumor. We use 3D regionally-partitioned temporal features in our system and obtain a superior precision-recall curve and a higher system BEP score than in [Chi et al. 2013b] with more than 20 times faster processing speed.

A CBIR system is proposed in [Costa et al. 2011] to differentiate cyst from METS using radiological semantic labels and computer-generated features based

on density histogram and its moments. The density histogram and its moments are obtained from three orthogonal 2D cuts of a single-phase scan volume. A random forest classifier is used to learn a discriminant distance between various FLL attributes. The classification performance is measured using a receiver operating characteristic curve. The framework proposed in [Costa et al. 2011] uses only global density-based features derived from the lesion area and the whole liver in a single contrast phase. Moments are well-known quantitative measures of the global shape of a set of points. FLL shapes, however, are rarely used to differentiate different lesions in the clinical routine. This may explain why inferior results are obtained using moments as the discriminating features of the FLLs.

In [Napel et al. 2010], the authors propose retrieval of cyst, METS and HEM using only a single image in the PV phase on a database of 30 images. Computer-generated image-based features and higher level radiological semantic labels are used to represent a FLL. Visual similarity between each pair of lesions is adjudged by two senior radiologists based on texture, boundary shape and boundary sharpness. The similarity measure between two FLLs is defined as 3/2/1 for very similar, somewhat similar and not similar pairs respectively. The system is evaluated in terms of precision and recall on how well the system retrieves visually similar lesions in comparison to radiology experts. A mean precision greater than 90% is achieved. The retrieval framework proposed in [Napel et al. 2010] is optimized and characterized for retrieving visually similar lesions as perceived by expert radiologists as opposed to retrieving lesions belonging to the same lesion class. Retrieval performance in terms of FLL characterization is not reported which makes a formal performance comparison with our system difficult. Further, only one slice in the PV phase, selected manually, is used for feature computation. Higher level

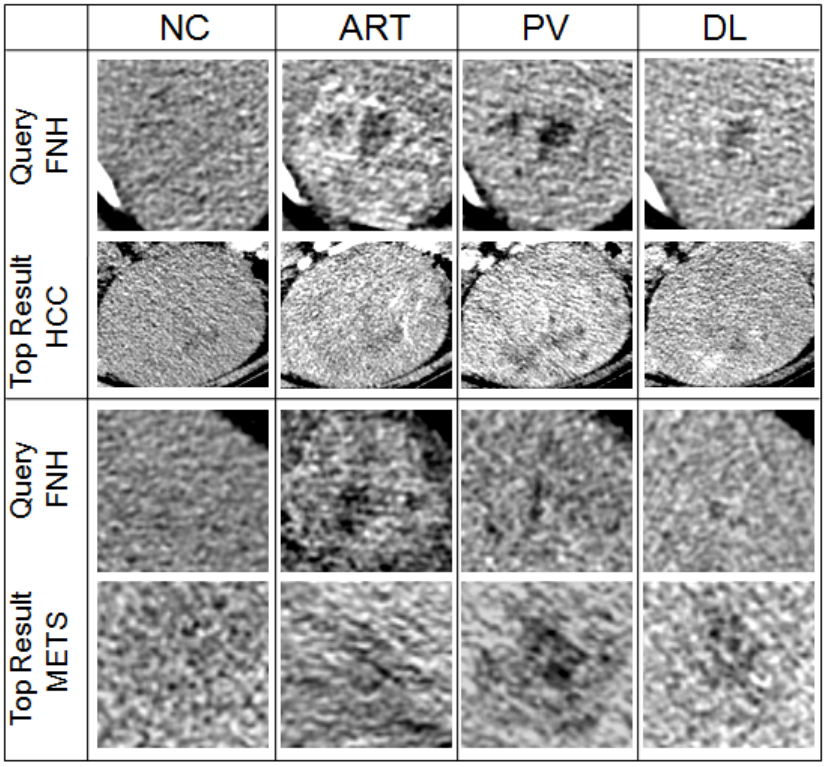


Figure 3.12: This figure shows some cases where the top retrieved lesion does not belong to the query lesion class.

radiological annotations that are inherently known to be subjective and widely user-dependent are used to bridge the performance gap.

3.6.2 System Performance

For most cases the proposed CBIR system ranks lesions belonging to the same pathological type as the query lesion higher than lesions from other pathological groups. However, in certain cases lesions from a different lesion class may be ranked higher as shown in Figure 3.12. This is due to variation in visual appearance among lesions belonging to the same pathological group. In clinical practice other higher level semantic information and clinical history are used to distinguish such

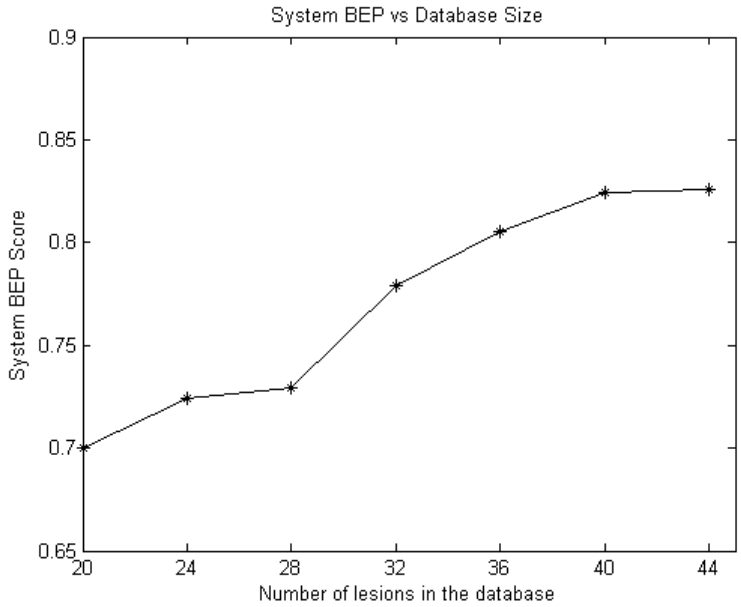


Figure 3.13: This plot shows the variation in system BEP score with respect to the database size.

cases. In this work we do not use any semantic information. However, in future we would like to explore other features that are more efficient in distinguishing visually similar lesions from different pathological classes.

3.6.3 Sensitivity to Database Size

In order to test the sensitivity of the proposed CBIR framework with respect to the database size, we test the system performance in terms of the BEP score versus the number of lesions in the database. Since our database is small, we want to check whether the system has passed the knee point and the improvement in performance curves have slowed down with respect to the database size. Figure 3.13 shows the variation in the BEP score with respect to the number of lesions in the database. From the figure we observe that there is no considerable improvement in the

retrieval performance if we have a database size of larger than 40 lesions.

3.7 Conclusion

In this chapter we propose a retrieval framework for FLLs using a simple tumor model based on 3D spatio-temporal features derived from 4-phase contrast-enhanced CT images. Acceleration techniques are employed to speed up the computationally heavy 3D feature extraction process known to be the primary bottleneck in integration of 3D feature-based retrieval systems into the clinical routine. The proposed system is evaluated in terms of precision-recall and system BEP score on a database of 44 lesions comprising of five pathological categories. The proposed system performs better and faster than existing 2D feature-based FLL CBIR systems.

In future work, we would like to conduct a clinical validation of the proposed system and evaluate the system's performance on a larger database that includes more FLL pathologies. We acknowledge that the database used in this work, though at par with some of the existing studies [Chi et al. 2013b] and [Napel et al. 2010], is small. CBIR systems are known to improve radiological diagnostic accuracy [Chi et al. 2013b], however, high processing times have rendered their integration into the clinical routine impractical. By keeping the query processing time low and including more pathological cases we hope to be able to integrate the proposed system as a diagnostic assistant into the routine radiological practices.

Chapter 4

Phenotype Detection in Mutant Mice

In addition to *reporting* and *diagnosis*, the third application we look at for outdated methodologies is *analysis*. Image analysis in certain biomedical research applications still rely on microscopic evaluation, an example of which is gene-to-physiology mapping. Completion of the human genome project brought comprehension of location and sequence of each human gene. Scientists now want to map these individual genes to their corresponding physiological functionalities. Mouse is chosen as the principal study model for this gene mapping due to its 99% genetic similarity with human [Collins et al. 2007] and [Consortium 2002].

4.1 Overview

Gene targeting technology is being actively employed by many international organizations like Knockout Mouse Project (KOMP), the EUropean Conditional Mouse Mutagenesis Program (EUCOMM), North American Conditional Mouse Mutagenesis Project (NorCOMM), and the Collaborative Cross to generate transgenic mouse

lines by knocking out each of the approximately 25,000 mouse genes (i.e., systematically removing each gene one by one and growing the mouse). High-throughput phenotypic assessment systems are necessary to systematically analyze and interpret the genetic information generated by these large-scale mutagenesis programs. A significant proportion of the generated mice strains are embryonic lethal resulting in the shift towards embryo-centric phenotyping.

Current embryo phenomic analysis largely relies on microscopic histological examinations which is not only labor intensive, highly time consuming and prone to distortion induced during sectioning but also has limited anatomical coverage. 3D non-destructive volumetric imaging has been proposed as the next technology for embryonic morphological analysis [Nieman et al. 2011]. Given that each individual 3D image can be very large, comparison of an ensemble of images for subtle differences is not practical by traditional microscopic observation. Further, if more than one mutant at multiple time points and under various experimental conditions is to be evaluated, then the task of systematic anatomical analysis can quickly become prohibitive. An automated protocol for the initial evaluation of mouse phenotypes would clearly be beneficial in identification of anatomical differences consistent over a mutant population.

4.2 Related Work

Worldwide efforts are underway to perform systematic anatomical analysis of the 3D mouse data. Coarse phenotypes where organs are absent or grossly malformed are easily detected without computer-aided mechanisms [Schneider and Bhattacharya 2004]. The research community mainly focusses on semi-automatic

differential volumetric analysis of morphological structures between normal and gene knockout strains using average images representing the normal (also called wild-type) and transgenic mice. In [Zamyadi et al. 2010], the authors use high resolution magnetic resonance imaging, non-linear image alignment and statistical analysis techniques to perform volumetric comparison of various organs including heart, lung, brain and liver of two transgenic mice strains. In a more exhaustive volumetric study the authors analyze volumetric variation of 48 anatomical structures between two genetically engineered mice strains using micro Computed Tomography (μCT) imaging, 3D mouse atlas and manual segmentation [Wong et al. 2012]. [Cleary et al. 2011] investigates volumetric difference of various brain structures in two mice strains using magnetic resonance and 3D atlas.

Although evaluation based on average image may be beneficial for an initial examination, phenotypes that are randomized in position and texture such as the intestines and developing trabeculae of the heart [Nieman et al. 2011] or subtle structural organ failures without large volume changes cannot be characterized using this technique. Another body of work focusses on automatic/semi-automatic detection of known abnormalities via segmentation [Hadjidemetriou et al. 2009] and [Norris et al. 2013] or by better data visualization resulting from enhanced tissue contrast [Degenhardt et al. 2010] and [Cleary et al. 2009]. Segmentation techniques fail if the defect characteristics are unknown or if the anatomy is hard to label such as bone joints. Enhanced visualization, while useful, still requires long expert hours to identify phenotypes, possibly multiple, in each of the thousands of mutant populations.

In this chapter, we present a generalized defect detection framework that automatically computes candidate phenotypic areas without using atlas, segmentation

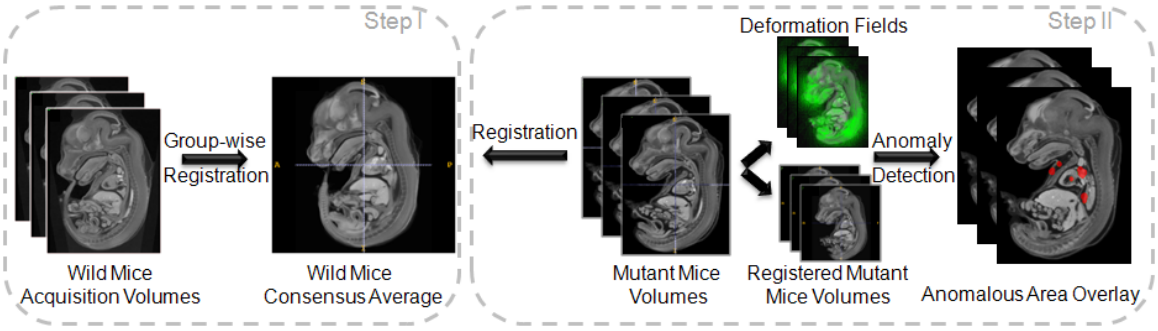


Figure 4.1: Defect detection consists of two steps. A mean of the normal mouse group is computed in the first step. In the second step, mutant group is registered to the normal mean and the resulting deformations are analyzed to detect defects.

or any defect-specific features. Instead, our approach uses deformation fields that are widely used to study anatomical variations [Zamyadi et al. 2010], [Wong et al. 2012], [Xie et al. 2010], and [Nieman et al. 2006]. We extract various features from deformation fields obtained by registering mutant mice to a normal mean and combine them to detect coarse, subtle as well as randomized defects (Figure 4.1). Statistical characteristics of deformation fields have been previously studied to detect gross defects in mice brain using multi-modality images [Nieman et al. 2006]. Our approach, however, targets a single imaging modality and successfully handles both subtle as well as significantly differing anatomy.

The rest of the chapter is organized as follows. In Section 4.3 we describe in detail the techniques used for automated defect identification. Section 4.4 shows some results obtained by our defect detection framework. Section 4.5 discusses the merits and de-merits of the proposed approach and concludes the chapter.

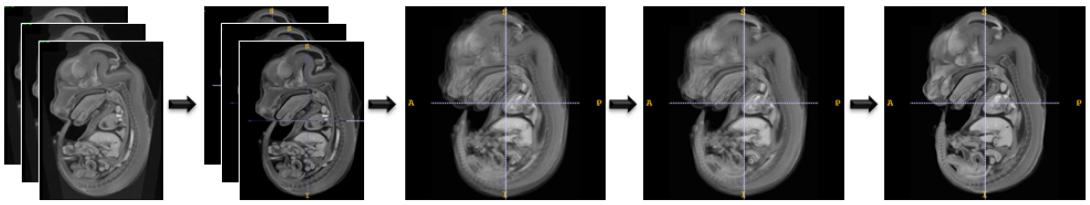
4.3 Methods

4.3.1 Sample Preparation

This study is performed using C57BL/10 mutant mice generated at the National Institute of Genetics, Japan. Post organogenesis, growth and development in the embryo starts at ~14.0 days post-coitum (dpc). Image registration cannot be applied at stages earlier than this due to unformed or absent organs. Further, at a relatively mature stage such as 15.5 dpc accurate registration of abdomen is difficult to achieve due to variation in intestinal position and crowding within the abdominal cavity [Wong et al. 2012]. Therefore, embryo samples at 14.5 dpc were collected for this study. A total of 14 embryos were used out of which 3 were normal and 11 had chromosomal aberrations. Out of the 11 mutant embryos, 3 were homozyogote generated by inbreeding C57BL/10 mice and 8 were heterozygote obtained by cross breeding C57BL/10 and normal littermates. The samples were washed in phosphate buffered saline and fixed in 4% paraformaldehyde until imaging. Before scanning, embryos were soaked in 1 : 3 mixture of lugol solution and double distilled water. Scan was carried out with the embryos fitted in 1.5 milliliters eppendorf tube fixed using wet paper.

4.3.2 Imaging Protocol

Short scan times and excellent soft tissue contrast obtained by tissue staining [Degenhardt et al. 2010] has made μCT a popular imaging technique for phenotyping studies. Embryo samples used in this study were imaged on a Scanxmate-E090S 3D μCT system (Comscantecno, Japan). Keeping the X-ray source at 60kVp and



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

Figure 4.2: This figure illustrates the steps in the computation of normal mean image. (a) Acquisition volume, (b) extracted normalized embryo images, (c)-(e) consensus average images at rigid, affine and B-Spline registration stages respectively.

130mA, each specimen was rotated 360° in steps of 0.36° generating 1000 projections of 640×480 pixels. The 3D μCT data was reconstructed at an isotropic resolution of $9.5 \times 9.5 \times 9.5$ μm^3.

4.3.3 Normal Mouse Consensus Average Image

The proposed defect detection algorithm can be broadly broken down to two major steps. The first step comprises of constructing a consensus average for the normal mouse population. As the second step, mutant mouse image is non-linearly registered to the normal mouse average and the resulting deformation field is analyzed to detect the defective areas (Figure 4.1). This section details the first step.

For computing the normal consensus average, embryo pixels are extracted from the acquisition volumes in an automated fashion using simple operations like Gaussian mixture modeling, thresholding and mathematical morphology such as selection of connected components and erosion/ dilation. We model the voxels of the acquisition volume using three Gaussian distributions. The three distributions

represent the voxels belonging to gel/air, minimum CT value inside the mouse area and the maximum CT value inside the mouse area. The Gaussian distribution that has its peak at the minimum value correspond to the gel/air. By thresholding the acquisition volume using the mean of this Gaussian distribution, we extract out the mouse area. Figure 4.2(a) and (b) show examples of an acquisition volume and the extracted embryo respectively. The intensity ranges of the embryos are then normalized prior to registration. One embryo is selected and taken as the reference, with all further processing performed in its image space. The others are first spatially normalized using a rigid registration algorithm consisting of six global degrees of freedom (DOF): three translation and three rotation parameters, in order to correct the differing orientations of each embryo in the normalized images. All the resulting images are then averaged into a new blurry reference image (Figure 4.2(c)). This initial registration step, due to the rigid transformation, does not affect the geometry of the subjects [Cleary et al. 2011]. Thus the average image is not biased towards the selected embryo's geometry. The embryos are then aligned using affine registration (12 global DOF: three translations, three rotations, three scales and three shears) to the previously created average image. Rigid and affine registration is performed using a block matching technique. Figure 4.2(d) shows the resulting average after affine registration. As a final step B-Spline-based non-linear registration is applied to locally align the affine registered embryos to the reference. The non-linear registration is formulated with a similarity energy function comprising of mutual information [Mattes et al. 2001] and a rigidity penalty [Staring et al. 2007]. We use 32 histogram bins to compute mutual information. Ten iterations of this registration are applied in a multi-resolution fashion where the control point spacing gradually reduces to eight voxels. We use four

resolutions of the images and an adaptive stochastic gradient descent optimizer. The reference image is updated after each iteration leading to the final consensus average (Figure 4.2(e)). Elastix toolbox is used to implement this registration scheme [Klein et al. 2010].

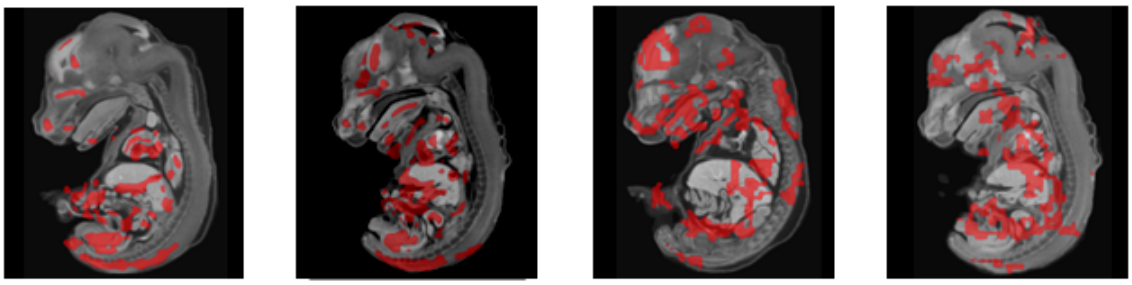
4.3.4 Deformation Features and Masks for Defect Detection

To detect defects in mutant mice, they are registered to the normal average image using the same three-stage registration pipeline as above except that in each stage the reference is always kept fixed to the normal average. Registration of each mouse results in the corresponding deformation field. These deformation fields encode detailed description of the anatomical differences between each mutant embryo and the normal group average, including the displacements, pose, and relative scale of every anatomical feature.

We use the deformation fields to compute 3D Jacobian maps using determinant of local Jacobian matrix at each voxel. F represents the deformation vector at voxel (x, y, z). The Jacobian $J(x, y, z)$ at voxel (x, y, z) is defined as:

$$J(x,y,z) = \begin{vmatrix} 1+\frac{\partial F_x}{\partial x} & \frac{\partial F_x}{\partial y} & \frac{\partial F_x}{\partial z} \\ \frac{\partial F_y}{\partial x} & 1+\frac{\partial F_y}{\partial y} & \frac{\partial F_y}{\partial z} \\ \frac{\partial F_z}{\partial x} & \frac{\partial F_z}{\partial y} & 1+\frac{\partial F_z}{\partial z} \end{vmatrix} \tag{4.1}$$

The Jacobian determinant is a local measurement of volumetric difference of each mutant subject relative to the normal group average. Jacobian determinant greater than one represents voxel expansion and less than one represents voxel compression. Jacobian of deformation is a popular tool to study inter-group structural



(a) (b)

Figure 4.3: (a) Example of Jacobian masks, I_J, overlaid on mutant images. (b) Example of stress masks, I_S, overlaid on mutant images.

Table 4.1: This table compares VSD and polydactyly detection accuracy (in %) of various features. VSD is assumed detected if the ventricular area is highlighted.

	I_J	I_S	$(I_J \cap I_S)$ F_1	$(I_{IV} \cap I_J)$ F_2	$(I_{IV} \cap I_S)$ F_3	$(F_2 \cup F_3)$ F_4	$(F_1 \cup F_4)$ F_5
VSD Sens.	88.8	100.0	88.8	77.7	77.7	77.7	100.0
VSD Spec.	50.0	100.0	100.0	100.0	100.0	100.0	100.0
Polydactyly Sens.	76.9	84.6	61.5	46.1	76.9	76.9	92.3
Polydactyly Spec.	48.4	80.6	87.1	90.3	87.1	87.1	87.1

differences [Xie et al. 2010], and [Nieman et al. 2006]. We apply Jacobian determinant in phenotyping by computing a Jacobian mask, I_J, one for each mutant mouse, that selects voxels at which Jacobian determinant is δ units away from one. I_J is defined as:

$$I_J = \begin{cases} 1, & \text{if } (1-\delta) \leq J \leq (1+\delta); \\ 0, & \text{otherwise.} \end{cases} \tag{4.2}$$

Figure 4.3(a) shows two Jacobian masks overlaid on a mutant embryo images. δ is kept 0.5 for experiments. To better characterize the local morphological differences, the Jacobian map only included non-linear deformation. Global transformations such as scaling of the whole embryo was not taken into account. We realized,

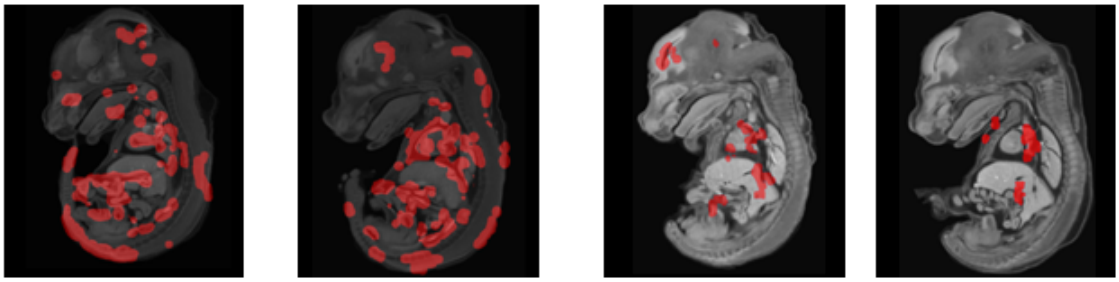
however, that Jacobian determinant fails for defects where volume changes are minimal. Further, we find that Jacobian introduces numerous false positives by highlighting areas that are found to be non-defective by the phenotyping experts, thus resulting in low precision. Performance of I_J in detecting known defects such as, Ventricular Septum Defect (VSD) and polydactyly is summarized in Table 4.1. Detection specificity for both the defects is very low with I_J.

To capture subtle defects with low volumetric changes, we compute another deformation feature that we call deformation stress. Deformation stress (D_s) is computed by dividing the volume into small blocks and measuring the entropy of deformation direction inside each block.

$$D_S(v) = -\sum_{u \in B(v)} p(\theta(u)) \log(p(\theta(u)); \tag{4.3}$$

$B(v)$ in Equation (4.3) represents the block in which voxel v lies and $\theta(u)$ is the displacement direction at voxel u. For experiments the volume was divided into cuboids of size eight voxels. Using D_S we compute a mask, I_S, that chooses voxel blocks that have high entropy of deformation direction and hence are undergoing incoherent deformation. For experiments I_S selected the top 50% blocks that exhibited highest deformation entropy. Figure 4.3(b) shows examples of I_S and Table 4.1 enlists its performance in detecting known phenotypes. Some false positives are introduced due to inclusion of sources and sinks in the deformation field.

Since I_J and I_S individually fail to detect all the defects and both introduce false positives, a simple combination of the two does not give satisfactory results as shown in Figure 4.4 and Table 4.1. In practice multiple mice from a mouse line are imaged before phenomic analysis is performed. We introduce this group



(a) (b)

Figure 4.4: (a) This figure shows detection results obtained by $I_J \cup I_S$. Simple union does not work because it introduces the false positives of both the individual components. (b) This figure shows detection results obtained using $I_J \cap I_S$. Many true positives detected by the individual masks are left out when the two masks are intersected.

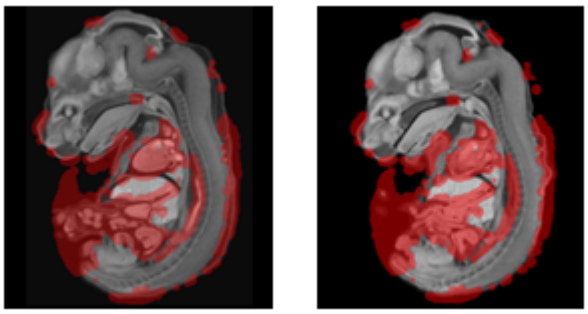


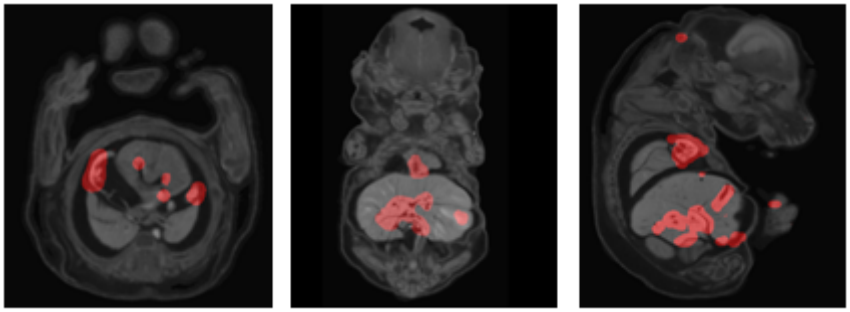
Figure 4.5: Example of intensity variance masks, I_{IV}, overlaid on mutant images.

information in defect detection by calculating voxel-wise intensity variance (V_{IV}) across the group of mutant mice that are registered to the normal mean. V_{IV} is defined as:

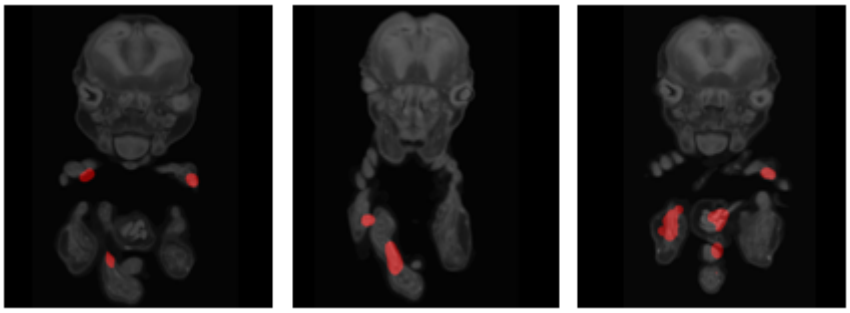
$$V_{IV}(v) = \frac{1}{N_M - 1} \sum_{i=1}^{N_M} (M_i(v) - N_{Avg}(v))^2. \tag{4.4}$$

The variables N_M, M_i and N_{Avg} in Equation (4.4) are mutant mouse population size, i^{th} registered mutant mouse image and the normal consensus average respectively.

From V_{IV} we compute a mask I_{IV} by selecting the top 50% voxels that have high intensity variance and hence low registration accuracy (Figure 4.5).



(a)



(b)

Figure 4.6: (a) Defects detected by ($I_{IV} \cap I_J$). (b) defects detected by ($I_{IV} \cap I_S$).

We find that many false positives introduced by I_J and I_S are pruned when these features are combined with I_{IV}. Table 4.1 lists the accuracies when the detection criterion is ($I_{IV} \cap I_J$), ($I_{IV} \cap I_S$) or both combined. From experiments on C57BL/10 mutant mice we find that ($I_{IV} \cap I_J$) mainly captures VSD and unusually wide liver lobe junctions. Figure 4.6(a) shows some detection results obtained by this factor. The factor ($I_{IV} \cap I_S$) captures polydactyly and unnatural position and deformations of tail and limbs as shown in Figure 4.6(b).

It is possible that some inaccurately registered morphological structures are not captured by the intensity variance mask. Uniform body cavities (dark regions)

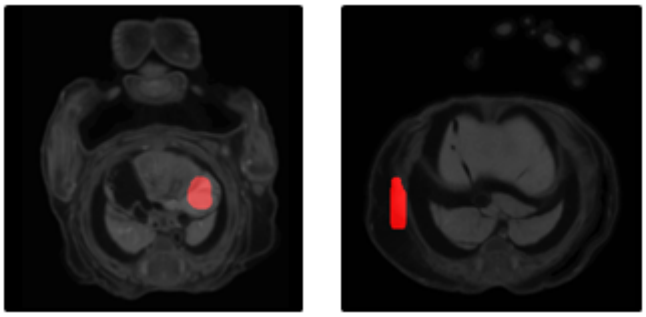


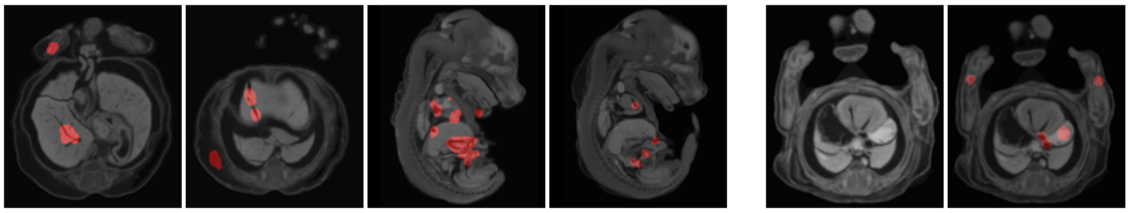
Figure 4.7: This figure illustrates regions detected by $(I_J \cap I_S)$.

or muscular organs (like liver lobes and heart atria) are some examples where we find that intensity variance is low due to spatially uniform intensity values even though there is an abnormality. The factor $(I_J \cap I_S)$ addresses regions where both Jacobian and stress are high irrespective of the intensity variance. By adding this term in the detection rule we are able to detect spatially uniform defective regions. Some secondary phenotypes like enlarged heart atrium due to high blood pressure induced by VSD are captured by this term as shown in Figure 4.7.

Combining the three terms we propose the defect detection rule as:

$$I_{Defect} = (I_{IV} \cap I_J) \cup (I_{IV} \cap I_S) \cup (I_J \cap I_S). \tag{4.5}$$

Table 4.1 enlists the performance of this detection rule. Number of detected regions can be readily increased or decreased by relaxing or tightening the thresholds while generating the Jacobian, deformation stress and intensity variance masks. Simple morphological operations like dilation and erosion are applied as noise reduction measures to clean up the detection results.



(a) (b)

Figure 4.8: (a) Defect detection results obtained using the complete detection rule (Equation (4.5)) in the liver lobe junctions, heart and intestine of C57BL/10 mice. (b) The left and right images depict a healthy heart and the misjudged defect respectively.

4.4 Results

Complete phenomic analysis of a mouse strain is a very tedious and slow process. C57BL/10 strain is still under investigation and hence full phenotypic characteristic consisting of all phenotypic defects is yet unknown. Therefore, even though we can evaluate the detection rule (Equation (4.5)) in terms of precision, a formal evaluation of recall is not possible. VSD and polydactyly are two established genetic defects in C57BL/10 mice. We compute sensitivity and specificity of the detection algorithm with respect to these two defects.

When evaluated over the mutant database of three homozygote and eight heterozygote embryos, the algorithm detected all cases of VSD without generating any false positives. Out of the 13 cases of polydactyly 12 were successfully detected and one was missed. The missed case belonged to the only mouse in the database that had its umbilical chord removed resulting in registration errors at the nearby areas. Four false positive polydactyl cases were reported in situations where due to high proximity toes of both the feet seemed fused in the 3D renderings.

To evaluate the rest of the detected areas, a user study was conducted with

phenotyping experts having long experience in mouse imaging and phenotyping. The expert comments were very encouraging and they noted that all the regions detected by the algorithm had biological significance. Some of the areas were due to genetic defects, some due to non-genetic defects and some due to organogenesis or procedural interventions. After careful histological examination, it was found that majority (57%) of the total detected regions belonged to genetic defects. Apart from known phenotypes, the algorithm detected areas in the liver lobe junctions that are candidate areas for potential phenotype of this mouse strain and are under active phenotypic evaluation.

Malformed body cavities constituted 14% output regions. Though these regions do not represent defects due to genetic makeup, they still signify biological malformations. Another 8% regions were noted to be due to genesis and extensive developmental remodeling of gonads at this gestational stage. The rest of the output was attributed to blood clots, randomized umbilical chord regions and pancreatic genesis.

When the detection rule was applied to wild-type mice, some areas were reported. These areas represent blood clots, umbilical chord, malformed body cavities and pancreatic and gonadic organogenesis. One false positive was generated for heart septal defect in a case where low spatial intensity variance makes the judgement hard even for an unexperienced human eye (Figure 4.8(b)). With further image processing it is possible to improve the detection accuracy by neglecting the high intensity blood clots and masking out umbilical chord regions.

4.5 Discussion and Conclusion

We propose a generic deformation-based defect detection framework for 3D μCT images of mutant mice. Our system has the potential to greatly enhance phenotyping throughput by automatically detecting all known phenotypes. Unlike other algorithms designed to detect specific known defects, our system also highlights candidate novel defects that may not be readily recognized by human experts due to absence of significant visual features. Owing to voxel-by-voxel analysis, defects are localized to sub-structures and those affecting multiple structures are visualized collectively. Though our evaluation database is small, the results clearly establish the potential of the proposed system in patterning defects. Our framework can be easily adapted to examine other 3D scan images amenable to registration. We acknowledge that the registration method may effect the detection results, however, the registration scheme used in this work is widely employed in mouse phenotyping [Zamyadi et al. 2010], [Wong et al. 2012], and [Cleary et al. 2011]. Deformation field resulting from only the non-linear registration step is used for defect detection. The detection performance is found to be fairly robust to parameter variation.

Since the proposed framework is independent of the defect features, classification of defects into those that are genetically induced and those that are not is out of scope for the current system. Currently we provide frequency of occurrence as an indicator of whether or not a defect is genetic. As an example, since VSD and polydactyly are detected in all homozygote embryos, the probability of these defects being genetic is reported to be 100%. Similar probabilities are assigned to all detected regions. As a future work it is possible to use advanced image processing and statistical techniques to device classifiers that can differentiate between genetic

and non-genetic defects.

Chapter 5

Conclusion

We set out to achieve two objectives in this thesis. The first objective was to better utilize the existing data in routine clinical and biomedical practice and the second objective was to use modern computer vision, image processing and computer graphics technologies to enhance the archaic reporting, diagnosis and analysis tools used in healthcare and biomedical research. In this chapter we summarize how these objectives were realized by the systems proposed in this thesis.

5.1 The VITA System

In Chapter 2, we propose a visual reporting framework, called VITA, to enhance traditional paper-based radiological reporting. VITA uses existing, otherwise discarded, radiological image markups to generate visual reports that embed radiological annotations in a 3D animated visualization of the exam data. Mining relevant information from the radiological annotations was the main technical challenge in this project. Placing the radiological findings in a 3D volumetric context within

the body part being analyzed makes it easier for physicians and their patients to comprehend the pathology. We validate that the visual reports generated by VITA improve the clarity of communication between radiologists, physicians and their patients via a user satisfaction study conducted with physicians in Singapore. By better utilizing discarded data and using computer graphic technologies, the VITA system aims to enhance the age-old paper-based radiological reporting routine.

As an extension to the VITA system, we cluster together individual annotations that denote a volume and automatically perform 3D segmentation of the anatomy marked by the radiologist. 3D segmentation, though rigorously studied in the literature, is not used in the clinical routine. Traditionally, radiologists still use manual segmentation or bounding box representation of 3D volumes. This is mainly because most PACS vendors do not integrate 3D segmentation into their systems. Further, radiologists are often required to provide redundant seeds to initiate these segmentation routines. By utilizing the radiological annotations to auto-generate segmentation seeds and integrating 3D segmentation into the VITA system, we are able to integrate 3D segmentation into the clinical routine. Again, we show that using discarded annotations we are able to provide radiologists automated segmentation tools within their PACS systems thereby relieving them from the tedious manual segmentation protocols.

5.2 Content-based Retrieval of Focal Liver Lesions

Chapter 3 proposes a fast content-based image retrieval framework for focal liver lesions. Using existing confirmed cases of liver tumors, the retrieval framework aids diagnosis of new tumors by retrieving from a database cases that appear

visually similar to the one under investigation. A novel 3D spatio-temporal feature extraction framework is proposed for an objective modeling of the tumor. Retrieval systems have been proved to improve diagnostic accuracy. However, high query processing times have rendered their use impractical in routine practices. We propose acceleration techniques to make query processing real-time. With high retrieval accuracy and fast processing the proposed system has the potential to be used as a diagnostic assistant in routine radiological practices. Content-based retrieval is yet another example of how existing data can be utilized to build clinically useful diagnostic tools using simple image processing techniques.

5.3 Phenotyping of Mutant Mice

In Chapter 4 we examined mouse phenotyping, a key biomedical research application that aims to map each mouse gene to its corresponding physiological function. The ultimate goal of this research is to be able to map each human gene to its corresponding physiological functionality in the human body. Routine mouse phenotyping still relies on the laborious and extremely low throughput microscopic examination of mouse sections even though prior to sectioning each mouse undergoes a 3D scan. This 3D data, though available, is not used in practice. Using this discarded 3D scan data we propose a generalized defect identification framework that automatically highlights physiological defects in micro-CT images of mutant mice using image processing techniques like non-linear image registration and deformation vector analysis. We have developed a novel deformation feature that is used to formulate the defect detection scheme. Since the proposed defect detection framework does not use any defect-specific features, it can detect both

known and novel defects. By pruning the vast search space of novel defects and also highlighting candidate defective areas that are hard to recognize by human eye due to lack of distinct visual features, the proposed defect detection framework greatly enhances the extremely low throughput of traditional microscopic mouse phenotyping.

5.4 Lessons Learned

While working with radiologists and phenotyping experts, we realized that inertia to changes in the routine workflow is the main reason why computational tools are not easily integrated into the routine practices. Hence, our efforts were continuously directed towards making our application frameworks least demanding in terms of user input and alterations to daily operations. Nonetheless, we believe that if our solutions do not make their way into the routine workflow, resistance to workflow changes will be the main malefactor.

5.5 Future Directions

There are several directions in which this thesis work could be extended, some of which are summarized in the following:

The current version of the VITA system does not provide much control to radiologists on how the visual reports will appear except for the choice of transfer functions. The transfer functions are used only to highlight or suppress certain types of tissues. As physicians start using VITA visual summaries, it is possible to develop a simple yet powerful visual reporting language that can allow radiolo-

gists to provide directives on how the visual summary should appear in order to visually highlight salient findings to assist the doctor's understanding of the patient's pathology. For example, in a case where a patient has multiple small lesions and one big lesion, the visual language can allow simple keyword annotations, e.g., <highlight>Pulmonary lesion</highlight>, as a directive to the volume rendering engine that this annotation should be displayed in a more prominent manner. By defining a small set of simple tags, radiologists could have more control over the final visual summary. This can help better guide the physician's focus towards key radiological findings.

Currently VITA simply embeds the available textual and geometric annotations in the 3D volume. If the radiologist does not associate any text tag with the geometry, there is no textual information in the volume rendering. An idea for future work is to extract text information from the radiological text report if no tag is associated with a geometrical annotation. This, however, is a non-trivial problem given the potential unstructured nature of the text reports.

The content-based retrieval system proposed in Chapter 3 is evaluated in a relatively small database. Hence, database indexing did not play a critical role in the system design. For larger databases indexing strategies make significant contribution to the retrieval speed. To keep retrieval systems real-time for larger clinical databases, it is important to investigate design of efficient indexing structures for fast data search and retrieval.

Chapter 4 proposes an automated defect detection framework for phenotyping of transgenic mice. The defect detection framework is based on registration of mutant mice to a normal mouse average and subsequent analysis of the resulting deformation vectors. The registration scheme used in the proposed system is

the standard registration pipeline used for mouse phenotyping in the literature. However, it is important to study the robustness of detection output when other registration schemes are employed. Further, the current version of the defect detection framework does not differentiate between genetic and non-genetic defects. All possible defects are highlighted. As a future work it is possible to use advanced image processing and statistical techniques to device classifiers that can differentiate between genetic and non-genetic defects so that defects can be categorized and selectively highlighted.

References

Akgul, C. B., Rubin, D. L., Napel, S., Beaulieu, C. F., Greenspan, H., and Acar, B. 2011. Content-based image retrieval in radiology: Current status and future directions. *J. Digit. Imaging 274*, 2 (April), 208–222. 42

Armato, S. G., et al. 2011. The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans. *Med. Phys. 38*, 2011, 915–931. 19

Armato, S. G. 2003. Computerized lung nodule detection: Effect of image annotaiton schemes for conveying results to radiologists. *Proc. of SPIE Medical Imaging 5032* (May), 854–859. 10

Berlin, L. 2009. Communicating results of all outpatient radiologic examinations directly to patients: The time has come. *Am. J. Roentgenol. 192*, 3 (March), 571–573. 27

Bosmans, J. M., Weyler, J. J., Schepper, A. M. D., and Parizel, P. M. 2011. The radiology report as seen by radiologists and referring clinicians: Results of the COVER and ROVER surveys. *Radiology 259*, 1 (April), 184–195. 11

Burner, A., Donner, R., Mayerhoefer, M., Holzer, M., Kainberger, F., and Langs, G. 2012. Texture bags: Anamoly retrieval in medical images based on local 3D-texture similarity. *Proc. of International Workshop on Medical Content-Based Retrieval for Clinical Decision Support*, 116–127. 47

Channin, D. S., Mongkolwat, P., Kelper, V., and Rubin, D. L. 2009. The annotation and image mark-up project. *Radiology 253*, 3 (December), 590–592. 15

Channin, D. S., Mongkolwat, P., Kelper, V., Sepukar, K., and Rubin, D. L. 2010. The cabig^{TM}annotation and image markup project. *J. Digit. Imaging 23*, 2 (April), 217–225. 15

Chi, Y., Zhou, J., Venkatesh, S. K., Huang, S., Tian, Q., and Liu, J. 2013. Computer-aided focal liver lesion detection. *Int. J. Compt. Assist. Radiol. Surg. 8*, 4, 511–525. 48, 50

Chi, Y., Zhou, J., Venkatesh, S. K., Tian, Q., and Liu, J. 2013. Content-based image retrieval of multiphase CT images for focal liver lesion characterization. *Med. Phy. 40*, 10. 3, 46, 47, 70, 71, 75

Chiaramonte, D. 2008. Who's afraid of the empowered patient? *JAMA 300*, 12 (September), 1393–1394. 27

Choplin, R. H., II, J. M. B., and Maynard, C. C. 1992. Picture archiving and communication systems: An overview. *RadioGraphics 12*, 1, 127–129. 12

ClearCanvas. 2008. RIS/PACS and DICOM viewing software solution. http://www.clearcanvas.ca/, Accessed January 25th, 2011. 15, 16

Cleary, J. O., Price, A. N., Thomas, D. L., Scambler, P. J., Kyriakopoulou, V., McCue, K., Schneider, J. E., Ordidge, R. J., and Lythgoe, M. F. 2009. Cardiac phenotyping in ex vivo murine embryos using μMRI. *NMR Biomed. 22*, 8 (October), 857–866. 78

Cleary, J. O., Modat, M., Norris, F. C., Price, A. N., Jayakody, S. A., Martinez-Barbera, J. P., Greene, N. D. E., Hawkes, D. J., Ordidge, R. J., Scambler, P. J., Ourselin, S., and Lythgoe, M. F. 2011. Magnetic resonance virtual histology for embryos: 3D atlases for automated high-throughput phenotyping. *Neuroimage 54*, 2 (January), 769–778. 78, 82, 91

Collins, F. S., Rossant, J., and Wurst, W. 2007. A mouse for all reasons. *Cell 128*, 1 (January), 9–13. 76

Consortium, M. G. S. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature 420*, 1 (December), 520–562. 76

Costa, M., Tsymbal, A., Hammon, M., Cavallaro, A., Suhling, M., Seifert, S., and Comaniciu, D. 2011. A discriminative distance learning-based CBIR framework for characterization of indeterminate liver lesions. *Proc. of International Workshop on Medical Content-Based Retrieval for Clinical Decision Support*, 92–104. 46, 70, 71, 72

Degenhardt, K., Wright, A. C., Horng, D., Padmanabhan, A., and Epstein, J. A. 2010. Rapid 3D phenotyping of cardiovascular development in mouse embryos by micro-CT with iodine staining. *Circ. Cardiovascular Imaging 3*, 3 (May), 314–322. 78, 80

DMCTK. 2003. Dicom offis toolkit (DCMTK). http://dicom.offis.de/, Accessed February 18th, 2011. 18

Doi, K. 2007. Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. *Comput. Med. Imaging Graph. 31*, 4-5 (March), 198–211. 42, 45

Eberly, D. H. 2006. *Game engine design, second edition: A practical approach to real-time computer graphics*. Morgan Kaufmann. 31

Elsayes, K. M., Narra, V. R., Yin, Y., Mukundan, G., Lammle, M., and Brown, J. J. 2005. Focal hepatic lesions: Diagnostic value of enhancement pattern approach with contrast-enhanced 3D gradient-echo MR imaging. *Radiographics 25*, 5, 1299–1320. 51

Fan, S. F., Xu, Z., He, H. Q., Ding, J. R., and Teng, G. J. 2011. The role of key image notes in CT imaging study interpretation. *J. Digit. Imaging 24*, 2 (April), 366–372. 3, 10

Francis, I. R., Cohan, R. H., McNulty, N. J., Platt, J. F., Korobkin, M., Gebremariam, A., and Ragupathi, K. 2003. Multidetector CT of the liver and hepatic neoplasms: Effect of multiphasic imaging on tumor conspicuity and vascular enhancement. *Am. J. Roentgenol. 180*, 5, 1217–1224. 42

Gletsos, M., Mougiakakou, S. G., Matsopoulos, G. K., Nikita, K. S., Nikita, A. S., and Kelekis, D. 2003. A computer-aided diagnostic system to characterize CT focal liver lesions: Design and optimization of a neural network classifier. *IEEE Trans. Inf. Technol. Biomed. 7*, 3 (September), 153–162. 45

Hadjidemetriou, S., Reichardt, W., Buechert, M., Hennig, J., and Elverfeldt, D. V. 2009. Analysis of MR images of mice in preclinical treatment monitoring of polycystic kidney disease. *Medical Image Computing and Computer Assisted Intervention-MICCAI 5762*, Pt 2, 665–672. 78

Hall, F. M. 2009. The radiology report of the future. *Radiology 251*, 1 (May), 313–316. 27

Haralick, R. M., Shanmugam, K., and Dinstein, I. 1973. Textural features for image classification. *IEEE Trans. Syst. Man Cybern. 3*, 6, 610–621. 56

ITK. 1999. Insight segmentation and registration toolkit (ITK). http://www.itk.org/, Accessed August 25th, 2010. 18

Ji, H., McTavish, J. D., Mortele, K. J., Wiesner, W., and Ros, P. R. 2001. Hepatic imaging with multidetector CT. *Radiographics 21*, Suppl 1, S71–S80, 2001. 42

Jirapatnakul, A. C., Mulman, Y. D., Reeves, A. P., Yankelevitz, D. F., and Henschke, C. I. 2011. Segmentation of juxtapleural pulmonary nodules using a robust surface estimate. *Int. J. Biomed. Imaging 2011*. 29

Kamel, I. R., Choti, M. A., Horton, K. M., Braga, H. J. V., Birnbaum, B. A., Fishman, E. K., Thompson, R. E., and Bluemke, D. A. 2003. Surgically staged focal liver lesions: Accuracy and reproducibility of dual-phase helical CT for detection and characterization. *Radiology 227*, 3, 752–757. 42

Kim, H. J., Kim, A. Y., Kim, T. K., Byun, J. H., Won, H. J., Kim, K. W., Shin, Y. M., Kim, P. N., Ha, H. K., and Lee, M. G. 2005. Transient hepatic attenuation differences in focal hepatic lesions: Dynamic CT features. *Am. J. Roentgenol. 184*, 1, 83–90. 42

Klein, S., Staring, M., Murphy, K., Viergever, M. A., and Pluim, J. P. W. 2010. elastix: A toolbox for intensity-based medical image registration. *IEEE Trans. Med. Imag. 29*, 1 (January), 196–205. 83

Kostis, W. J., Reeves, A. P., Yankelevitz, D. F., and Henschke, C. I. 2003. Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images. *IEEE Trans. Med. Imag. 22*, 10 (October), 1259–1274. 29

Lencioni, R., Cioni, D., and Bartolozzi, C. 2006. *Focal liver lesions: Detection, characterization, ablation*. Springer. 42

Levin, A., Lischinski, D., and Weiss, Y. 2004. Colorization using optimization. *ACM T. Graphic. 23*, 3 (August), 689–694. 37

Li, C., Xu, C., Gui, C., and Fox, M. D. 2010. Distance regularized level set evolution and its application to image segmentation. *IEEE Trans. Image Process. 19*, 12 (December), 3243–3254. 34, 35

Long, L. R., Antani, S., Deserno, R. M., and Thoma, G. R. 2009. Content-based image retrieval in medicine: Retrospective assessment, state of the art, and future directions. *Int. J. Healthc. Inf. Syst. Inform. 4*, 1 (January), 1–16. 42

Maenpaa, T., and Pietikainen, M. 2005. Texture analysis with local binary patterns. *Handbook of Pattern Recognition and Computer Vision 3*, 197–216. 54

Manjunath, B. S. 2002. *Introduction to MPEG-7: Multimedia content description interface*. John Wiley & Sons, Inc. 59

Massoptier, L., and Casciaro, S. 2008. A new fully automatic and robust algorithm for fast segmentation of liver tissue and tumors from CT scans. *Eur. Radiol. 18*, 8, 1658–1665. 52

Mattes, D., Haynor, D. R., Vesselle, H., Lewellyn, T. K., and Eubank, W. 2001. Nonrigid multimodality image registration. *Proc. of SPIE Medical Imaging 4322*, 1 (July), 1609–1619. 82

Mougiakakou, S. G., Valavanis, I. K., Nikita, A., and Nikita, K. S. 2007. Differential diagnosis of CT focal liver lesions using texture features, feature selection and ensemble driven classifiers. *Artificial Intelligence in Medicine 41*, 1 (September), 25–37. 45

Muller, H., Michoux, N., Bandon, D., and Geissbuhler, A. 2004. A review of content-based image retrieval systems in medical applications - clinical benefits and future directions. *Int. J. Med. Inform. 73*, 1 (February), 1–23. 42, 45

Napel, S. A., Beaulieu, C. F., Rodriguez, C., Cui, J., Xu, J., Korenblum, D., Greenspan, H., Ma, Y., and Rubin, D. L. 2010. Automated retrieval of CT images of liver lesions on the basis of image similarity: Method and preliminary results. *Radiology 256*, 1 (July), 243–252. 46, 70, 72, 75

NEMA. 2008. Digital imaging and communications in medicine (DICOM) standard. http://medical.nema.org/, Accessed January 20th, 2011. 12

Nieman, B. J., Flenniken, A. M., Adamson, S. L., Henkelman, R. M., and Sled, J. G. 2006. Anatomical phenotyping in the brain and skull of a mutant mouse by magnetic resonance imaging and computed tomography. *Physiol. Genomics 24*, 2 (January), 154–162. 79, 84

Nieman, B. J., Wong, M. D., and Henkelman, R. M. 2011. Genes into geometry: Imaging for mouse development in 3D. *Curr. Opin. Genet. Dev. 21*, 5 (October), 638–646. 77, 78

Nokia. 2009. Qt application and UI framework. http://qt.nokia.com/, Accessed August 7th, 2010. 16

Norris, F. C., Modat, M., Cleary, J. O., Price, A. N., McCue, K., Scambler, P. J., Ourselin, S., and Lythgoe, M. F. 2013. Segmentation propagation using a 3D embryo atlas for high-throughput MRI phenotyping: Comparison and validation with manual segmentation. *Magn. Reson. Med. 69*, 3 (March), 877–883. 78

NVIDIA. 2010. Cg toolkit. http://developer.nvidia.com/cg-toolkit/, Accessed October 10th, 2010. 17

Park, S. J., Seo, K. S., and Park, J. A. 2005. Automatic hepatic tumor segmentation using statistical optimal threshold. *Computational Science - ICCS 2005 3514*, 934–940. 52

Qian, Y., Gao, X., Loomes, M., Comley, R., Barn, B., Hui, R., and Tian, Z. 2011. Content-based retrieval of 3D medical images. *Proc. of the Third International Conference on eHealth, Telemedicine, and Social Medicine*, 7–12. 47, 54, 71

Reiner, B., and Siegel, E. 2006. Radiology reporting: Returning to our image-centric roots. *Am. J. Roentgenol. 187*, 5 (November), 1151–1155. 3, 10, 11

Rosenfeld, A., and Pfaltz, J. L. 1996. Sequential operations in digital picture processing. *J. ACM 13*, 4, 471–494. 52

Rosset, A., Spadola, L., and Ratib, O. 2004. Osirix: An open-source software for navigating in multidimensional DICOM images. *J. Digit. Imaging 17* (September), 205–216. 15

Roy, S., Yao, M., Brown, M. S., and Shih, G. 2011. Visual interpretation with three-dimensional annotation (VITA): Open source automated 3D visual summary application using AIM (annotation imaging markup) enabled PACS based on radiologist annotations. *Radiological Society of North America*. 8

Roy, S., Brown, M. S., and Shih, G. 2012. Automatic 3D volume extraction from 2D annotations. *Radiological Society of North America*. 8

Roy, S., Brown, M. S., and Shih, G. 2012. Extracting volumetric information from standard two-dimensional radiological annotations within the clinical workflow. *Proc. of the IEEE International Conference on Bioinformatics and Biomedicine Workshops* (October), 725–731. 8

Roy, S., Brown, M. S., and Shih, G. 2013. Visual interpretation with three-dimensional annotations: Three-dimensional image interpretation tool for radiological reporting. *J. Digit. Imaging*. 8

Roy, S., Liang, X., Kitamoto, A., Tamura, M., Shiroishi, T., and Brown, M. S. 2013. Phenotype detection in morphological mutant mice using deformation features. *Medical Image Computing and Computer Assisted Intervention-MICCAI 8151*, Pt 3 (September), 437–444. 9

Rubin, D. L., Mongkolwat, P., Kelper, V., Supekar, K., and Channin, D. S. 2008. Medical imaging on the semantic web: Annotation and image markup. *AAAI Spring Symposium: Semantic Scientific Knowledge Integration*, 93–98. 15

Rubin, D. L., Rodriguez, C., and Shah, P. 2008. ipad: Semantic annotation and markup of radiological images. *AMIA Annual Symposium Proceedings 2008*, 626–630. 15

Rubin, G. D. 2000. Data explosion: The challenge of multidetector-row CT. *Eur. J. of Radiol. 36*, 2 (November), 74–80. 11

Schneider, J. E., and Bhattacharya, S. 2004. Making the mouse embryo transparent: Identifying developmental malformations using magnetic resonance imaging. *Birth Defects Res., Part C, Embryo Today 72*, 3 (September), 241–249. 77

Schwartz, L. H., Panicek, D. M., Berk, A. R., Li, Y., and Hricak, H. 2011. Improving communication of diagnostic radiology findings through structured reporting. *Radiology 260*, 1 (July), 174–181. 11

Sharma, N., and Aggarwal, L. M. 2010. Automated medical image segmentation techniques. *Journal of Medical Physics/ Association of Medical Physicists of India 35*, 1 (January), 3–14. 29

Sobel, J. L., Pearson, M. L., Gross, K., et al. 1996. Information content and clarity of radiologists' reports for chest radiography. *Acad. Radiol. 3*, 9 (September), 709–717. 46

Staring, M., Klein, S., and Pluim, J. P. W. 2007. A rigidty penalty term for nonrigid registration. *Med. Phys. 34*, 11 (November), 4098–4108. 82

Stoutjesdijk, M. J., Futterer, J. J., Boetes, C., Die, L. E. V., Jager, G., and Barentsz, J. O. 2005. Variability in the description of morphologic and contrast enhance-

ment characteristics of breast lesions on magnetic resonance imaging. *Invest. Radiol. 40*, 6 (June), 355–362. 46

Wong, M. D., Dorr, A. E., Walls, J. R., Lerch, J. P., and Henkelman, R. M. 2012. A novel 3D mouse embryo atlas based on micro-CT. *Development 139*, 17 (September), 3248–3256. 78, 79, 80, 91

Xian, G. M. 2010. An identification method of malignant and benign liver tumors from ultrasonography based on GLCM texture features and fuzzy SVM. *Expert Systems with Applications 37*, 10 (October), 6737–6741. 45

Xie, Z., Yang, D., Stephenson, D., Morton, D., Hicks, D., Brown, T., and Bocan, T. 2010. Characterizing the regional structural difference of the brain between tau transgenic (rTg4510) and wild-type mice using MRI. *Medical Image Computing and Computer Assisted Intervention-MICCAI 6361*, Pt 1, 308–315. 79, 84

Yang, W., Lu, Z., Yu, M., Huang, M., Feng, Q., and Chen, W. 2012. Content-based retrieval of focal liver lesions using bag-of-visual-words representations of single- and multiphase contrast-enhanced CT images. *J. Digit. Imaging 25*, 708–719. 46, 70, 71

Yu, M., Lu, Z., Feng, Q., and Chen, W. 2010. Liver CT image retrieval based on non-tensor product wavelet. *Proc. of International Conference on Medical Image Analysis and Clinical Applications*, 67–70. 45

Yu, M., Lu, Z., Feng, Q., and Chen, W. 2012. Extraction of lesion-partitioned features and retrieval of contrast-enhanced liver images. *Computational and Mathematical Methods in Medicine 2012* (July). 45, 46, 70, 71

Yushkevich, P. A., Piven, J., an R. G. Smith, H. C. H., Ho, S., Gee, J. C., and Gerig, G. 2006. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage 31*, 3 (March), 1116–1128. 48, 51

Zamyadi, M., Baghdadi, L., Lerch, J. P., Bhattacharya, S., Schneider, J. E., and Henkelman, R. M. 2010. Mouse embryonic phenotyping by morphometric analysis of MR images. *Physiol. Genomics 42A*, 2 (October), 89–95. 78, 79, 91

Zimmerman, S. L., Kim, W., and Boonn, W. W. 2011. Informatics in radiology: Automated structured reporting of imaging findings using the AIM standard and XML. *RadioGraphics 31*, 3, 881–887. 15, 36