Liver Segmentation by Geometric and Texture Features using Support Vector Machine (LiGTS)

A Thesis

Presented in Partial Fulfillment of the Requirements for the Degree Master of Science in the Graduate School of The Ohio State University

By

Kovid Bhatnagar, BS, 2013

Graduate Program in Department of Electrical and Computer Engineering

The Ohio State University

2016

Master’s Examination Committee:

Dr. Arunark Kolipaka, Advisor

Dr. Rizwan Ahmad, Advisor
Abstract

Magnetic resonance elastography (MRE) is a novel, non-invasive technique to measure the stiffness of soft tissue. In MRE, mechanical vibrations are induced in the region of interest (ROI), which are then encoded in the phase of an MR image. These wave images are further processed to obtain a spatial stiffness map. The work presented in this thesis is focused on segmenting the liver to facilitate stiffness quantification for staging liver fibrosis. The current practice requires an expert radiologist to segment the liver in MRE scans so that the stiffness computation can be performed within the liver. In addition to being labor intensive, this manual intervention causes inter- and intra-observer variability in drawing the boundaries of the liver and thus impacts the computed stiffness values. Hence, there is a need to increase the consistency of the stiffness computation process, which will enhance the clinical effectiveness of the MRE-based stiffness computation. Automating the segmentation process can improve repeatability and eliminate inter- and intra-observer variability, providing the desired consistency for MRE-based stiffness quantification. In this work, an automated method for liver segmentation is presented. The method is based on supervised learning
and employs a combination of geometric and texture-based features in conjunction with support vector machine-based classification. For validation, liver MRE is performed in 14 healthy volunteers. For performance evaluation, results from automated segmentation are analyzed in terms of error rate, sensitivity, specificity, and precision, with manual segmentation providing the ground truth. The results show that the proposed automated segmentation method has an error rate of less than 4%. Also, the mean stiffness values derived from the proposed method are in good correlation ($R^2 = 0.96$) with the values obtained from a manually drawn ROI. In summary, this study demonstrates and validates a new segmentation method for liver MRE.
Acknowledgments

I would like to acknowledge my advisors, Dr. Ahmad and Dr. Kolipaka, for their support and guidance throughout my graduate studies. I would like to thank Brian Raterman for his technical support in scanning the volunteers. I would also like to extend my gratitude towards volunteers who participated in this study; the completion of this study could not be possible without their participation. I would also like to acknowledge my parents and friends for their love and support.
Vita

5 September 2008 - April 2013 ....... B.S. Electrical and Computer Engineering; University of Cincinnati
August 2013 - December 2015 ......... M.S. Electrical and Computer Engineering; The Ohio State University

Fields of Study

Major Field: Electrical and Computer Engineering
For Science.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>iv</td>
</tr>
<tr>
<td>Vita</td>
<td>v</td>
</tr>
<tr>
<td>Dedication</td>
<td>vi</td>
</tr>
<tr>
<td>List of Figures</td>
<td>ix</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xi</td>
</tr>
<tr>
<td>1. Motivation and Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2. Background</td>
<td>4</td>
</tr>
<tr>
<td>2.1 Clinical Need</td>
<td>4</td>
</tr>
<tr>
<td>2.1.1 Liver Biopsy</td>
<td>5</td>
</tr>
<tr>
<td>2.1.2 Current Non-invasive Techniques</td>
<td>6</td>
</tr>
<tr>
<td>2.2 Image Segmentation</td>
<td>12</td>
</tr>
<tr>
<td>2.3 Image Segmentation for MRI</td>
<td>14</td>
</tr>
<tr>
<td>2.4 Image Segmentation in MRE</td>
<td>16</td>
</tr>
</tbody>
</table>
3. Features and Classifiers

3.1 Gray Level Co-occurrence Matrix
3.2 Texture Features
  3.2.1 Contrast
  3.2.2 Entropy
  3.2.3 Energy
  3.2.4 Cluster Prominence
  3.2.5 Sum Variance
  3.2.6 Correlation
  3.2.7 Difference Entropy
  3.2.8 Sum Average
3.3 Geometric Feature
3.4 Classifiers
  3.4.1 Linear Discriminant Analysis (LDA)
  3.4.2 Support Vector Machine
  3.4.3 SVM-Morph Classifier

4. Materials and Methods

4.1 Experimental Setup
4.2 Image Acquisition
4.3 Data Processing
  4.3.1 Training
  4.3.2 Tuning
  4.3.3 Testing
4.4 Performance Metrics
4.5 Statistical Analysis

5. Results and Discussion

5.1 Results
5.2 Discussion
5.3 Future Work
5.4 Conclusions

Bibliography
### List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 CT imaging of a cirrhotic liver.</td>
<td>7</td>
</tr>
<tr>
<td>2.2 Ultrasound imaging for a cirrhotic liver.</td>
<td>8</td>
</tr>
<tr>
<td>2.3 MRI for liver fibrosis.</td>
<td>10</td>
</tr>
<tr>
<td>2.4 Illustration of MRE principle.</td>
<td>11</td>
</tr>
<tr>
<td>2.5 An example of liver MRE.</td>
<td>12</td>
</tr>
<tr>
<td>2.6 Histogram partitioning to segment liver.</td>
<td>16</td>
</tr>
<tr>
<td>3.1 Contrast feature.</td>
<td>21</td>
</tr>
<tr>
<td>3.2 Entropy feature.</td>
<td>22</td>
</tr>
<tr>
<td>3.3 Energy feature.</td>
<td>22</td>
</tr>
<tr>
<td>3.4 Cluster prominence feature.</td>
<td>23</td>
</tr>
<tr>
<td>3.5 Sum variance feature.</td>
<td>24</td>
</tr>
<tr>
<td>3.6 Correlation feature.</td>
<td>25</td>
</tr>
<tr>
<td>3.7 Difference entropy feature.</td>
<td>26</td>
</tr>
</tbody>
</table>
3.8 Sum average feature. ................................. 27
3.9 Run length feature. ................................. 28
3.10 Illustration of LDA. ............................... 30
3.11 Non-linear classification. ......................... 31
3.12 SVM-based non-linear classification. .......... 32
3.13 Morphological filtering. ......................... 32
4.1 Setup for liver MRE. ............................... 35
4.2 Training, tuning, and testing datasets. ........ 36
4.3 Ground truth based on manual segmentation. .. 37
4.4 MRE stiffness map. ............................... 39
4.5 Computation of mean liver stiffness. .......... 40
5.1 Nine features used for segmentation. .......... 42
5.2 Segmentation results (volunteer 1). ............ 44
5.3 Segmentation results (volunteer 2). .......... 45
5.4 Correlation plots for mean stiffness. .......... 46
5.5 Bland-Altman analysis for mean stiffness. .... 47
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 Performance comparison of LDA, SVM, and SVM-Morph.</td>
<td>43</td>
</tr>
</tbody>
</table>
Chapter 1: Motivation and Introduction

Many pathological conditions alter the elastic properties of soft tissue. Magnetic resonance elastography (MRE) is a noninvasive, quantitative imaging technique capable of measuring stiffness of any soft tissue in the body. A distinguishing feature of MRE data acquisition, compared to other MRI techniques, is the application of external periodic motion. The resulting oscillatory motion of the tissue is captured by phase-contrast MRI (PC-MRI), performed in the presence of motion-encoding gradients (MEG) that are synchronized with the external motion. The local wavelength of the propagating waves in the phase images generated by PC-MRI is then converted to a tissue stiffness map through a process known as inversion. Organs that have been studied using MRE include liver [1], lung [2], skeletal muscle [3], breast [4], heart [5], kidney [6], brain [7], and aorta [8]. MRE-derived liver stiffness, in particular, is a useful biomarker to characterize liver fibrosis and has emerged as a noninvasive and more accurate alternative to liver biopsy.
In MRE, vibrations are sent into the liver to generate a stiffness map. Since the stiffness within the liver varies spatially, a manual ROI selection can lead to variability in mean stiffness values. Therefore, there is a need to identify a region with adequate wave penetration and signal quality within the liver to report the mean stiffness value. The manual segmentation currently employed in clinical practice is not only time consuming but also susceptible to intra- and inter-observer variability. The focus of the research described in this thesis is to develop an automated liver segmentation tool that can be used to report stiffness of the liver, without any user interaction.

Image segmentation has been an active area of research for decades, but its application to MRE is still in its infancy. Recently, Dzyubak et al. [1] presented a method for segmenting liver in MRE images. The method is based on segmenting the intensity histogram of MRE magnitude images. In this approach, Dzyubak et al. make an assumption that the liver occupies a significant fraction of the pixels, and all those pixels have similar intensity values. Therefore, the liver pixels should have the highest peak in the intensity histogram of an MRE image. Starting from this initialization, morphological operators and active contouring are employed to further refine the segmentation. Although this approach is simple to implement, intensity-based segmentation has limitations due to significant intensity variations exhibited by MRE magnitude images. Also, active contouring,
due to its propensity to ignore minute features, may fail to identify blood vessels within the liver.

The goal of this work is to develop a robust, automated liver segmentation method for MRE. The proposed method employs a support vector machine (SVM) based segmentation and utilizes a novel combination of geometric and texture-based features. The method, called Liver segmentation using Geometric and Texture features using Support vector machine (LiGTS), is tuned and trained using 32 datasets from eight volunteers and then validated using 24 datasets from six additional volunteers. The ground truth, generated from manual segmentation by an expert, is used both for SVM training and for performance evaluation. The performance of LiGTS is assessed in terms of error rate, precision, sensitivity, and specificity. The mean liver stiffness values derived from LiGTS are compared to the values obtained from a manual selection of an ROI, a practice currently used in clinic.

The ensuing chapters of this thesis are organized as follows. Chapter 2 describes the existing techniques used to study liver fibrosis and summarizes the segmentation techniques used in MRI and MRE. Chapter 3 describes the features and classifiers used in LiGTS. Chapter 4 describes the experimental setup used to validate LiGTS. Chapter 5 summarizes the results, discussion, and conclusions.
Chapter 2: Background

Liver fibrosis is a common feature of almost all chronic liver diseases [9]. Fibrosis is a dynamic process that tends to progress, leading to hepatic dysfunction. A number of clinical tools exist for staging liver fibrosis. Some of these tools, such as a biopsy, are invasive and require tissue samples to be extracted from the body for diagnosis. This invasive test is painful to the patients and has a number of adverse side effects. Non-invasive techniques include various types of imaging. In this chapter, an overview of existing clinical tools is provided for evaluating liver fibrosis as well as their limitations. A brief introduction to MRE is provided along with the segmentation techniques used in MRI and MRE.

2.1 Clinical Need

According to the National Institutes of Health (NIH), liver cirrhosis is the 12th leading cause of death [10]. Cirrhosis is the final advanced stage of liver fibrosis, which indicates scarring of liver tissue. The scarring of liver tissue causes liver dysfunction by slowing down the production of proteins that are needed by the
body, eventually leading to fatal conditions [11]. Similarly, other diseases such as non-alcoholic fatty liver disease and non-alcoholic steatohepatitis are among the most common diseases in the Western world that lead to the fibrosis of the liver [12]. It is also known that liver fibrosis/cirrhosis leads to an increase in liver stiffness [13]. Therefore, there is a need for a method for staging liver fibrosis so that treatment may be provided to patients at early stages of the disease.

2.1.1 Liver Biopsy

Liver biopsy is the gold standard procedure used for diagnosing liver fibrosis and evaluating its stage of advancement. This is an invasive procedure that can potentially cause many complications for the patient. Some of the main problems with liver biopsy reported in [14, 15] include: (i) sampling variability, that is, a biopsy samples only 1/50,000th of the organ and hence has a chance of missing the abnormalities; (ii) there are 0.03% cases of death directly related to the test; (iii) the test causes moderate pain in about 20% of the cases and severe pain requiring intravenous analgesics in about 3% of the cases; and (iv) the test has a direct cost of about $1,500 and additional indirect cost related to time off from work, hospital stay, etc. Also, the biopsy only determines the static mass of fibrosis and is unable to reveal the progression of the disease [16]. Serial follow-ups of a patient’s longitudinal condition may require multiple biopsies, compounding the above problems. It is therefore desirable to develop non-invasive alternatives to liver biopsy.
2.1.2 Current Non-invasive Techniques

The current non-invasive modalities for diagnosing liver fibrosis are computed tomography, ultrasound, and MRI. These methods are able to detect changes in the liver parenchyma when there is significant fibrosis. However, these methods are not useful for identifying patients with less advanced stages of fibrosis. The details of these techniques are as follows.

Computed Tomography

Computed tomography (CT) is a non-invasive imaging method for assessing the extent of liver fibrosis. Specific anatomical features such as an irregular or nodular hepatic surface, shown by arrows in the Figure 2.1 [17], a blunt liver edge, parenchymal abnormalities, and morphological changes are used to diagnose fibrosis. Some of the advantages of CT scan include high temporal and spatial resolutions, fast acquisition without the need for a breathhold, and ability to image subjects with metal implants. The main drawback of CT is the use of radiation and possible allergic reaction to the contrast dye, which might be fatal [18]. Additionally, the CT liver images are qualitatively scored with moderate sensitivity (77% – 84.3%) and specificity (52.9% – 67.6%) in staging liver fibrosis [19]. Finally, as seen in the Figure 2.1 the contrast between normal and fibrotic liver is poor.
Ultrasound imaging (USI) provides information about morphological changes in the liver, often associated with advanced liver fibrosis. Altered shape parameter of the hepatic nerve, as shown in Figure 2.2 [20], detected by USI has been shown to strongly correlate with cirrhosis [21]. Additionally, USI has lower cost, is portable, and has high temporal resolution. However, there are many challenges in using ultrasound images for diagnosis of liver fibrosis/cirrhosis. For example, the portal vein morphology can be conspicuous in the absence of ascites, and the parenchymal texture is a characteristic that is subjective and has low sensitivity for the detection of cirrhosis. Furthermore, USI is operator dependent, has low spatial resolution, and requires extensive technical expertise to diagnose early stage liver fibrosis.

Apart from the morphological or anatomical USI, ultrasound elastography (USE) has been used to determine stiffness of the liver [22] by introducing acoustic
shear waves in the liver and tracking the wave speed using ultrasound transducer. However, USE has less sensitivity and specificity [23] compared to magnetic resonance elastography (described in a later section) in staging liver fibrosis [24]. Because of significant differences in the USE techniques among different vendors, it is difficult to establish standard cutoff values to stage liver fibrosis [25]. Furthermore, USE has some technical limitations; for example, it provides only a relative measure of stiffness, and the acoustic waves introduced by ultrasound have poor penetration depth in obese patients. Additional limitations of USE are provided elsewhere [25].

Figure 2.2: Ultrasound imaging for a cirrhotic liver. Hepatic vein wall morphology, shown on left, is wavy and corresponds to cirrhosis. Liver surface nodularity seen in the image on right also corresponds to cirrhosis. Both images are adapted from [20].

**Magnetic Resonance Imaging**

Similar to CT and ultrasound imaging, MRI also uses anatomical and morphological features such as surface nodularity, heterogeneous enhancement, caudate
lobe enlargement, and other features to diagnose liver cirrhosis as shown in Figure 2.3 [18]. The sensitivity and specificity of MRI are similar to that of CT with 87% and 54%, respectively [18]. The main advantage of MRI compared to CT is that it does not require radiation. Additionally, there are other MRI techniques, such as dynamic contrast enhanced imaging and diffusion weighted imaging, which can be used for diagnosing liver fibrosis/cirrhosis. However, all the above mentioned MRI techniques require patients to hold their breath, which sometimes becomes challenging. Additionally, the spatial resolution of MRI is poor compared to CT, and this modality cannot be used to image subjects with metal implants. Contrast enhanced MRI requires use of contrast agents which have additional side effects in some subjects, including patients with renal dysfunction with glomerular filtration rate less than 30. Furthermore, diagnosis of liver fibrosis/cirrhosis using MRI is qualitative and subjective and suffers from poor sensitivity and specificity in staging lower grades of fibrosis [26]. Because of these limitations, there is a need for noninvasive imaging technique with high sensitivity and specificity to diagnose all grades of liver fibrosis.

**Magnetic Resonance Elastography**

Magnetic resonance elastography (MRE) is a noninvasive imaging technique to estimate stiffness of soft tissues [27, 28, 29]. As shown in Figure 2.4, MRE is a three-stage process [28]. In the first stage (Figure 2.4a), noninvasive vibrations are induced in the region of interest. In the second stage (Figure 2.4b), these
vibrations are tracked using a PC MRI sequence. Finally (Figure 2.4c), the wave images are mathematically converted to a stiffness map by a process known as inversion.

MRE is currently an FDA approved method to diagnose liver fibrosis and is available as a commercial tool through all major MRI vendors. An example of liver MRE is shown in Figure 2.5 [30]. As apparent in the figure, the wavelengths (red to blue distance in the wave image) of the propagating waves are longer in a fibrotic liver compared to a normal liver. These wavelengths, on a pixel-by-pixel basis, are then converted into a liver stiffness map, and the range of the
resulting stiffness values are used to stage liver fibrosis. MRE has shown to be a very promising tool with high sensitivity (73–91%) and specificity (79–85%) [31]. For each MRE scan, radiologists have to manually draw regions of interest to determine the stiffness of individual organs, e.g., liver; this introduces inter- and intra-observer variability into the stiffness computation process. Therefore, there is a need for an automated segmentation of liver in MRE scans that can take away the observer-dependent variability. This automation is the focus of the research presented here.
Figure 2.5: Liver MRE. (A) MRE of a healthy liver in top row and of a liver with cirrhosis in second row. Elastogram shows an increased stiffness in liver with cirrhosis. (B) Upper row shows MRE for an obese patient and lower row for a patient with ascites. This figure is adapted from [30]

2.2 Image Segmentation

Segmentation of a digital image is the process that breaks apart the image into smaller regions based on some characteristics of either individual pixels or groups of neighboring pixels. The purpose of segmentation is to identify different objects present in the image. Image segmentation has been used for facial recognition
[32], medical diagnosis [33], satellite image processing [34], criminal investigations [35], and many other applications. Segmentation of an image can be performed by an expert based on his or her visual understanding of the image or by a computer algorithm based on computable image properties. For the sake of consistency, accuracy, and time-efficiency, automated segmentation algorithms are very desirable, and several application-specific algorithms have been developed for image segmentation.

There are many different types of characteristics that can be used to algorithmically segment images. Some well-known methods for image segmentation are based on edge detection, thresholding, region growing, active contouring, and model-based classification [36]. The edge detection based methods use intensity variation to find boundaries. These boundaries are then used to isolate an object or objects in an image. The thresholding based methods use one or more pixel characteristics in an image, with pixel intensity being the most commonly used characteristic. The pixels with intensities residing in a certain range are grouped together, resulting in separate objects containing pixels of different intensities. The region growing based methods start from an initial seed point and progressively add neighboring pixels with similar properties such as texture [37]. The active contouring based methods seek to draw object boundaries by balancing the forces exerted on a contour from its two sides; minimizing energy of the boundary
in the context of these two forces yields a segmentation of the image. The modelbased segmentation relies on matching properties, such as shape or size, of an image object to their expected values. In any particular segmentation task, more than one of the above techniques are typically combined before final segmentation is achieved. In the following sections, we briefly review some of the methods used for segmenting objects in MRI and MRE images.

2.3 Image Segmentation for MRI

There are a number of methods that have been developed for segmenting medical images [33]. Segmentation of brain images, in particular, has been an active area of research because accurate segmentation is critical for detecting tumors, edema, and other pathologies. For MRI of brain, performing segmentation can be particularly challenging due to spatially varying noise, undersampling artifacts, inhomogeneous intensity, and poor contrast-to-noise ratio. As a result, simple intensity-based segmentation methods do not yield satisfactory results [7], and more advanced, multi-layered techniques are often employed. For example, the approach proposed in [7] uses a pre-processing step of intensity inhomogeneity correction. The segmentation method itself involves intracranial mask identification via morphological operations and fuzzy connectedness followed by a maximum likelihood classification with spatial constraints. Due to the richness and complexity of brain images, the segmentation of brain MRI—despite the development
of several techniques—is still done manually by radiologists [38]. A difficult seg-
mentation challenge is also encountered in cardiac MRI images, where the most 
common objective is to evaluate the cardiac function by identifying myocardium. 
Compared to brain imaging, automated segmentation techniques have been suc-
cessfully used for cardiac application [39].

For liver segmentation in MRI, several methods have been published. Many of 
these methods, including the ones based on level set and active contours, rely on 
explicit or implicit contour extraction. Such methods, however, tend to perform 
poorly in the regions of low-gradient boundary; therefore, additional information 
is often used to assist the segmentation process in low-gradient regions. For exam-
ple, the method presented in [40] segments the liver using active contouring but 
also utilizes a liver template and manual editing in low-gradient regions. In ad-
dition, model-based segmentation has also been proposed for liver segmentation. 
For example, the method presented in [41] uses prior probability maps derived 
from liver shape and tissue properties and then employs a maximum likelihood 
classifier to segment the liver. The results show an error rate of 8.3% for normal 
livers and 11.8% for fatty livers. All the above examples show that automated 
segmentation of organs in MRI images is a difficult task and many working algo-
rithms are only semi-automatic.
Figure 2.6: Histogram showing the pixel intensity in an MRE image. The image intensity for the entire image is shown in dark blue. The peak for liver is shown in red, and the light blue and gray peaks show other tissue and adipose tissue, respectively.

2.4 Image Segmentation in MRE

In contrast to the MRI scans, the MRE measures the mechanical properties of tissues by introducing shear waves and imaging their propagation using the traditional PC MRI. In addition to the intensity image, MRE also generates a stiffness map within the field of view. For liver MRE, the goal of segmentation is to identify the liver tissue based on its distinct features in comparison to its surroundings. The segmented liver tissue is then used to quantify liver stiffness and thereby assess the stage of fibrosis in the liver. In comparison to more established MRI techniques, development of segmentation techniques for liver MRE is still in its infancy. In 2013, Dzyubak et al. presented a multi-step method for liver
The proposed method uses a combination of pixel intensity thresholding, active contouring, and morphological operations. One of the key assumptions made in this method is that the liver is the largest organ in the MRE image and the pixels within the liver have similar pixel intensities. This assumption is used to identify the highest peak in the intensity histogram of the image and select the corresponding region as the liver region. The histogram in Figure 2.6, adapted from [1], shows the complete intensity image histogram and the Gaussian fits used to extract pixels corresponding to the liver, adipose, and other tissues. In this histogram, the uppermost curve is the image histogram, and the three peaks under it represent the three Gaussian fits to the overall distribution. The highest peak in the middle is assumed to correspond to the liver, and the pixel intensities that belong to this peak are used to threshold the intensity image of the MRE scan. Starting from this initial segmentation provided by histogram partitioning, morphological operations are used to erode the fat regions near the edges of the liver. The Local Entropy Minimization based on bi-cubic Spline (LEMS) method is then used to create an active contour around the liver. Morphological operators are then used to further smoothen the contour. After segmentation, the wave propagation confidence map, which is available as a part of MRE reconstruction, is used to remove areas with poor wave penetration [1]. Finally, the resulting mask is used to compute the stiffness value.
Most MRI and MRE segmentation methods outlined above identify organs using pixel intensity values or contour extraction. Such techniques perform poorly in the presence of intensity inhomogeneities and low-gradient boundaries. It is hypothesized in [42] that texture of organs in MRI images is a better discriminator of organs and has been attempted for segmentation of kidneys. Our proposed method described in chapter 3 uses texture and geometric properties to segment liver in MRE images.
Chapter 3: Features and Classifiers

In this chapter, we describe a new liver segmentation method that is based on supervised learning. First, we present a discussion of the geometric and texture-based features that we have selected for discriminating the liver pixels from the rest of the pixels in an MRE scan. Second, we discuss three supervised learning-based classifiers that use the features to perform image segmentation: Linear discriminant analysis (LDA), support vector machine (SVM), and an extension of SVM that employs morphological operations to remove small clusters of false-positive pixels outside the liver; we refer to this last method as LiGTS.

3.1 Gray Level Co-occurrence Matrix

Texture properties of a region in an image depend on the local patterns in the pixel intensity values. These local patterns can be effectively identified by analyzing gray level co-occurrence matrix (GLCM). The GLCM is computed from the image on patch-by-patch basis. It is a square matrix of size $N \times N$, with $N$ being the number of gray levels in the image. The matrix element at location $(i, j)$
is the relative frequency with which two neighboring pixels, one with intensity $i$ and the other with intensity $j$, occur within the selected patch. In this work, we used $N = 16$ and the patch size of $5 \times 5$.

### 3.2 Texture Features

In total, eight different texture features, as described below, are computed using the GLCM [43]. These features include: Contrast, entropy, energy, cluster prominence, sum variance, correlation, difference entropy, and sum average. A brief description of these features is provided below.

#### 3.2.1 Contrast

Image contrast is a measure of the difference between the intensity values of neighboring pixels [44]. A larger difference between the intensities indicates a higher contrast value. Contrast can be computed from the GLCM using the following equation:

$$\text{Contrast}_{x,y} = \sum_{i=1}^{N} \sum_{j=1}^{N} P_{x,y}(i,j)(i - j)^2,$$

(3.1)

where $P_{x,y}(i,j)$ is the $(i,j)$ entry of the GLCM, which represents the relative frequency with which two neighboring pixels, one with intensity $i$ and the other with intensity $j$, occur for an image patch centered at $(x,y)$. As evident from Equation 3.1, $i = j$ entries represent the instances where neighboring pixels have identical intensities (zero contrast) and do not contribute to the contrast. Figure
3.1 shows an intensity image from the abdominal region and the corresponding contrast feature.

![Intensity Image and Contrast Feature](image.png)

Figure 3.1: Intensity image (left) and the corresponding contrast feature (right).

### 3.2.2 Entropy

Entropy is a measure of orderliness in the pixel intensity values in a region [44]. Entropy is computed from the GLCM using the expression given in Equation 3.2. Figure 3.2 shows an example of entropy feature computed from the GLCM.

\[
\text{Entropy}_{x,y} = \sum_{i=1}^{N} \sum_{j=1}^{N} P_{x,y}(i,j)(-\ln P_{i,j}).
\] (3.2)

### 3.2.3 Energy

Energy is a feature that measures the overall uniformity of intensity values in an image patch. Energy is computed from the GLCM using the expression given

21
in Equation 3.3. A higher energy value indicates that the neighboring pixels have very similar intensities. An example showing the energy feature is given in Figure 3.3.

\[
\text{Energy}_{x,y} = \sum_{i=1}^{N} \sum_{j=1}^{N} P_{x,y}(i,j)^2.
\]  

(3.3)
3.2.4 Cluster Prominence

Cluster prominence is a measure of asymmetry of the GLCM. A higher value implies more asymmetry about the mean value. The expression to compute cluster prominence is given in Equation 3.4. An example showing cluster prominence feature is provided in Figure 3.4.

\[
\text{Cluster Prominence}_{x,y} = \sum_{i=1}^{N} \sum_{j=1}^{N} \left\{ i + j - \mu_x - \mu_y \right\}^4 \times P_{x,y}(i, j),
\]

where

\[
\mu_x = \sum_{i=1}^{N} \sum_{j=1}^{N} i P_{x,y}(i, j), \quad \mu_y = \sum_{i=1}^{N} \sum_{j=1}^{N} j P_{x,y}(i, j).
\]

Figure 3.4: Intensity image (left) and the corresponding cluster prominence feature (right).
3.2.5 Sum Variance

The sum variance feature measures the variance of the sum of adjoining pixel values. It is computed using the expression given in Equation 3.5. An example of the sum variance feature is shown in Figure 3.5.

\[
\text{Sum Variance}_{x,y} = \sum_{k=2}^{2N} (k - \mu)^2 \times P_{x+y}(k),
\]

where

\[
P_{x+y}(k) = \sum_{i+j=k}^{i=N, j=N} P_{x,y}(i,j), \quad \mu = \sum_{k=2}^{2N} k \times P_{x+y}(k).
\]

Figure 3.5: Intensity image (left) and the corresponding sum variance feature (right).

3.2.6 Correlation

The GLCM-based correlation measures the linear dependency among the intensities of neighboring pixels. A value of 1 means the gray levels are perfectly
correlated, a value of -1 means they are anti-correlated, and a value of 0 means there is no correlation among the gray levels of neighboring pixels. This feature is computed using Equation 3.6. An example showing the correlation feature is presented in Figure 3.6.

\[
\text{Correlation}_{x,y} = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} \{i \times j\} \times P_{x,y}(i, j) - \{\mu_x \times \mu_y\}}{\sigma_x \times \sigma_y}.
\] 

(3.6)

Figure 3.6: Intensity image (left) and the corresponding correlation feature (right).

### 3.2.7 Difference Entropy

Difference entropy measures the orderliness in the difference of the gray levels of the neighboring pixels. The difference entropy feature is computed using the expression given in Equation 3.7. An example of the difference entropy feature is
provided in Figure 3.7.

\[
\text{Difference Entropy}_{x,y} = - \sum_{k=0}^{N-1} P_{|x-y|}(k) \log(P_{|x-y|}(k)),
\]

(3.7)

where \( P_{|x-y|}(k) \) is defined as:

\[
P_{|x-y|}(k) = \sum_{|i-j|=k} P_{x,y}(i,j).
\]

3.2.8 Sum Average

The sum average feature measures the local average intensity. This is computed by the expression given in Equation 3.8. An example of the sum average feature is provided in Figure 3.8.

\[
\text{Sum Average}_{x,y} = \sum_{k=2}^{2N} k \times P_{x+y}(k),
\]

(3.8)

where \( P_{x+y}(k) \) is the same as defined for the sum variance feature above.
3.3 Geometric Feature

In addition to the texture-based features described above, a geometric feature is also included to facilitate liver segmentation. The geometric feature, called run length, identifies large connected regions in the image. Two neighboring pixels are considered connected if they have the same intensity level. To compute the run length feature, first, the vertical resolution of the intensity image is reduced such that it only contains 16 shades of gray. Second, for each shade of gray, connected regions are identified. Third, all pixels within each connected region are given a value that is equal to the size (number of pixels) of that region. This feature is helpful in identifying pixels that are a part of large connected regions. Since liver is a large organ, this geometric feature helps identify the pixels that belong to the liver. An example of run length feature is shown in Figure 3.9.

Figure 3.8: Intensity image (left) and the corresponding sum average feature (right).
3.4 Classifiers

Utilizing the above mentioned nine features, we employ both linear and non-linear classifiers to segment the liver. For linear classification, LDA is used. Because linear classifiers, such as LDA, may perform poorly when the data points in the feature space are not well separated by a hyperplane, we also propose a non-linear segmentation method based on SVM. To further improve the performance of SVM, we also present an extension, called SVM-Morph, where the output of SVM is subjected to morphological filtering. Before application, all the classifiers are trained using training datasets where the correct segmentation (ground truth) is available.

3.4.1 Linear Discriminant Analysis (LDA)

LDA is a technique which finds an optimal line, plane, or hyperplane in the feature space to separate the training instances of two classes. The optimality
is described in terms of minimizing the intra-class variance and maximizing the
inter-class separation [45]. An example in Figure 3.10 shows a 2-dimensional
feature space occupied by two classes: the class 1 (blue) and class 2 (red). The
line, $T$ on the right shows the optimal linear surface on which the projections of
all data points are examined. This optimal direction is such that the variance
of projected points within each class is minimized and the distance between the
projection means is maximized. The line, $W_{LDA}$, normal to $T$ is the resulting
linear classifier. While computing the $W_{LDA}$ vector, an assumption is made that
the two classes have Gaussian distributions after projecting onto the line $T$. Once
the $W_{LDA}$ is determined using training datasets, segmentation can be performed
based on which side of $W_{LDA}$ the test sample resides in.

3.4.2 Support Vector Machine

Being a linear classifier, LDA has its limitations. For example, the hyperplane
learned by LDA or any other linear classifier will fail to separate the data points
shown in Figure 3.11. In contrast, nonlinear classifiers, such as SVM, are not
restricted to finding a hyperplane in the feature space. Nonlinear SVM employs
the kernel trick and projects the feature space data to a higher dimensional trans-
formed space; by performing a linear segmentation in the transformed space, SVM
leads to a nonliner segmentation in the original feature space. Also, unlike LDA,
it does not depend on the distributions of the data points within the individual
classes.
The illustration in Figure 3.12, adapted from [46] shows an example of a 2-class data in a 2-dimensional feature space that is not linearly separable. The nonlinear transformation to a 3-dimensional space makes it easy to find a hyperplane separating the two classes. The SVM finds an optimal linear surface in this higher dimensional space.

### 3.4.3 SVM-Morph Classifier

SVM-Morph is an extension of the SVM classifier described in the previous section. SVM-Morph performs additional morphological filtering on the SVM segmentation results. The goal is to retain only the largest connected object in
Figure 3.11: A depiction of nonlinear classification for two classes that are not separable linearly.

the binary segmentation provided by SVM and remove small, isolated fragments that most likely do not belong to the liver. SVM-Morph is obtained by following these three steps. Step-1: Dilate SVM segmentation to merge nearby objects in the binary map. Step-2: Discard all regions except for the largest connected region. Step-3: Perform a logical AND operation between the SVM output and the output from Step-2. Figure 3.13 shows the impact of applying the above mentioned three-step procedure.
Figure 3.12: Mapping data from a 2D feature space (left) to a 3D transformed space (right) allows a hyperplane to separate the classes in the transformed space, leading to a nonlinear classification in the original feature space.

Figure 3.13: Output of SVM (left) and SVM-Morph (right). The small isolated regions that were misclassified by SVM as liver have been successfully removed in SVM-Morph.
Chapter 4: Materials and Methods

This chapter describes materials and methods used for data acquisition and image analysis. In addition, it describes the statistical analysis used to evaluate the performance of LiGTS.

4.1 Experimental Setup

Fourteen normal (healthy) volunteers recruited for this study (10 male, 4 female; age: 20-30 years) were imaged after obtaining approval from the Institutional Review Board and a written informed consent from each subject.

4.2 Image Acquisition

All imaging was performed in a 3T (TimTrio, Siemens, Erlangen, Germany) MRI scanner. For each examination, the subject was positioned supine with head first into the scanner. For MRE acquisition, mechanical waves were introduced into the subjects liver by a commercial pneumatic driver system (Resoundant Inc., Mayo Clinic Foundation, Rochester, MN). As shown in Figure 4.1, the passive driver was placed anterior on the patients body close to the liver region (centered
at the level of the xiphoid process) and secured with an elastic Velcro belt. The passive driver was connected via a plastic tube to the active driver which was placed outside the scan room to induce 60 Hz vibrations into the liver. A rapid gradient recalled echo (GRE) MRE sequence (MREr) [47] was used to acquire four axial slices covering the major portion of the liver. Four sets of wave images spaced equally over a period of the wave motion were obtained by changing the temporal relationship between the motion encoding gradients (MEGs) and the external vibrations. MREr sequence imaging parameters included: TE=21.4 ms; TR=25 ms; FOV=40 cm; $\alpha$ (Ernst Angle)=16°; slice thickness=5 mm; matrix=128x64; Generalized Autocalibrating Partially Parallel Acquisitions (GRAPPA) acceleration factor of 2; vibration frequency=60 Hz; and MEG of 16.67 ms duration (60 Hz) was applied to measure the through-plane tissue motion. Additionally, a balanced steady state free precession sequence (bSSFP) was performed to obtain four axial slices with same resolution and slice location to that of MRE acquisition.

4.3 Data Processing

The segmentation was performed on the intensity images acquired using bSSFP sequence. As depicted in Figure 4.2, all the data were separated into three groups: training data, tuning data, and testing data. The training data consisted 24 datasets from 6 subjects, i.e., four slices per subject. The tuning data consisted of 8 datasets from 2 subjects. The testing data consisted of 24 datasets from 6 subjects. To establish ground truth for the training and tuning phases and for
evaluating performance for the testing phase, an expert manually segmented the liver for all the 56 datasets. As depicted in Figure 4.3, the manual segmentation deliberately excluded larger blood vessels in the liver.

4.3.1 Training

For each bSSFP slice, all the nine features described in the previous chapter were computed. These nine features for all 24 datasets along with the manually drawn ground truth were used to train the LDA and SVM classifiers.

4.3.2 Tuning

The LDA classifier does not possess any free parameter and therefore does not require any tuning. However, the SVM classifier, when using the Radial Basis Function (RBF) [48], possesses two free parameters that require tuning. These
parameters, called $\nu$ and $\sigma^2$, control the nature of the learned classifier [48]. The parameter $\sigma^2$ provides a trade off between the simplicity of the decision boundary, i.e., the surface that separates the two classes, and extent of errors in classification resulting from this boundary. A high value of $\sigma^2$ prefers the decision surface to be very regular and smooth at the cost of potential errors in classification. A low value of $\sigma^2$ prefers a decision boundary that minimizes the classification errors but is more complex. This added complexity of the decision boundary makes it prone to over-fitting, resulting in potentially poor generalization. The parameter $\nu$
influences the number of support vectors that contribute to the decision boundary. To tune $\nu$ and $\sigma^2$, the SVM classifiers trained on different combinations of $\nu$ and $\sigma^2$ were used to perform segmentation on the tuning datasets; the combination that yielded the minimum error rate was selected for further testing.

### 4.3.3 Testing

After training and tuning the classifiers, testing was performed on 6 subjects i.e. 24 datasets. As described earlier, nine features were computed from each dataset and fed to the trained classifiers. For the testing phase, the ground truth was not provided to the classifiers. For SVM-Morph, the segmentation results from the SVM classifier were subjected to morphological filtering as described in the previous chapter.
4.4 Performance Metrics

The performances of LDA, SVM, and SVM-Morph classifiers were evaluated based on the error rate, precision, sensitivity and specificity. With FP, FN, TP, and TN representing false positive, false negative, true positive, and true negative, respectively, the above mentioned performance metrics are defined as follows.

1. Error Rate = \[ \frac{FP + FN}{FP + FN + TP + TN} \]
2. Precision = \[ \frac{TP}{TP + FP} \]
3. Sensitivity = \[ \frac{TP}{TP + FN} \]
4. Specificity = \[ \frac{TN}{TN + FP} \]

These metrics were computed based on a direct comparison of the pixel labels output by the classifiers with the labels in the ground truth created by manual segmentation. In addition to the metrics that explicitly characterize the performance of the classifiers, mean liver stiffness values from the LDA, SVM, and SVM-Morph segmentations were compared against the values obtained using a clinical method, where an ROI is directly drawn on the stiffness map by an expert user. Figure 4.4 shows intensity image, four phase offsets, and the corresponding stiffness map. A 95% confidence map is overlaid on the stiffness map, with the hatched regions indicating the pixels with unreliable stiffness values due to poor signal to noise ratio of the propagating waves. The process of mean stiffness
computation is depicted in Figure 4.5. The axial through-plane wave images for each volunteer were processed automatically on-line using a standard inversion algorithm to generate stiffness maps [31]. For liver stiffness estimation, the stiffness map was masked by both the binary output of the classifier and the 95% confidence map; the stiffness values of the remaining pixels were averaged to yield an estimate of mean liver stiffness. To establish reference values for mean liver stiffness, an experienced user was asked to manually draw an ROI on the stiffness map, avoiding the hatched region.

Figure 4.4: MRE phase images and the corresponding stiffness map. (a) Intensity image from bSSFP sequence; (b-e) MRE-based phase images at four different phase offsets, describing the wave propagation; (f) Stiffness map computed from the phase offset images.

4.5 Statistical Analysis

The mean stiffness values obtained using the three classifiers (LDA, SVM, and SVM-Morph) were compared against the values obtained by manually drawing ROIs on the stiffness maps using a paired students t-test to determine the significance of difference ($P < 0.05$). In addition, a least squares linear regression
Figure 4.5: The process of determining the mean liver stiffness. (a) Intensity image from bSSFP sequence; (b) Segmentation output of a classifier; (c) Stiffness map with pixels outside the 95% confidence interval marked using the hatched regions; (d) The stiffness map in (c) masked by the binary map in (b); (e) the map in (d) after discarding the pixels belonging to the hatched regions. The mean stiffness is computed from non-zero pixels in (e) and compared against the values obtained from manual ROIs directly drawn on (c) by an expert.

was performed to evaluate correlation between the stiffness values obtained using classifiers and the values based on manual ROI selection. Furthermore, Bland-Altman analysis was performed to determine the agreement in mean stiffness values obtained using classifiers and from the manual ROI selection.
Chapter 5: Results and Discussion

In this chapter, results of segmentation using the three classifiers previously described are presented. For error rate, precision, sensitivity, and specificity, manual segmentation of the liver is used as the ground truth. For mean stiffness of liver, the ROIs drawn on the stiffness maps by an expert are used to create the reference values. Furthermore, this chapter provides conclusions and future directions.

5.1 Results

Figure 5.1 provides an example of all nine features, normalized between 0 and 1, used in LDA and SVM classifiers. It can be observed from the figure that each feature highlights a different property of the image. In particular, the run length feature that identifies large connect objects provides additional information, which is distinct from the information provided by the texture based features. Our preliminary results (data not shown) strongly suggest that the addition of run length feature significantly improves the performance of all the classifiers.
Figure 5.1: Nine features—eight texture based and one geometric—computed from one of the bSSFP datasets.

Table 5.1 shows the error rate, precision, sensitivity, and specificity values of each classifier. It can be observed that SVM-Morph outperforms other methods in terms of error rate, precision and specificity (bold), whereas LDA is superior in terms of sensitivity (bold). These values were obtained by collectively analyzing all 24 datasets used for testing. Figures 5.2 and 5.3 show representative results.
from two of the six test subjects. The figures show all four slices collected for the
two test subjects. To facilitate visual assessment, the binary segmentations are
overlaid on the corresponding intensity images.

<table>
<thead>
<tr>
<th></th>
<th>LDA</th>
<th>SVM</th>
<th>SVM-Morph</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision</td>
<td>48.00%</td>
<td>71.95%</td>
<td>78.80%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>76.07%</td>
<td>69.09%</td>
<td>68.36%</td>
</tr>
<tr>
<td>Specificity</td>
<td>92.89%</td>
<td>97.48%</td>
<td>98.29%</td>
</tr>
<tr>
<td>Error Rate</td>
<td>8.45%</td>
<td>4.61%</td>
<td>3.88%</td>
</tr>
</tbody>
</table>

Table 5.1: Performance metrics averaged across 24 datasets obtained from six
test subjects. The best performance is highlighted by bold font.

SVM-Morph showed an excellent correlation in stiffness values to the values
obtained from the manual ROI selection when compared to other classifiers. Fig-
ure 5.4 shows correlation plots for all three classifiers for the 24 test datasets.
LDA, SVM, and SVM-Morph exhibited $R^2$ values of 0.79, 0.81, and 0.96, re-
spectively. Also, for SVM-Morph, the bias between the estimated stiffness values
and the values obtained from manually drawing ROIs was minimal, as indicated
by a slope of nearly one and a small intercept of 0.037. A significant difference
($P = 0.00051$) in mean stiffness values when compared to manually drawn ROI
was observed for LDA, but no significant difference for SVM ($P = 0.3019$) and
SVM-Morph ($P = 0.0924$) was observed.
Figure 5.2: Classification results for one of the test subjects. All four slices (left to right) are shown. The results from the three classifiers are overlaid on the intensity image.

Figure 5.5 shows the Bland-Altman analysis for the mean stiffness values. Compared to the other classifiers, SVM-Morph exhibited the smallest limits of agreement (-0.10–0.14 kPa). Also, for SVM and SVM-Morph, negligible bias was observed in the estimation of stiffness, while a slight negative bias (-0.1 kPa) was observed for LDA. However, LDA exhibited a significantly larger limits of agreement (-0.36–0.18 kPa). Collectively, Table 5.1, Figure 5.4, and Figure 5.5 indicate that the SVM-Morph outperforms the other two classifiers in identifying the liver tissue and in facilitating the computation of liver stiffness.
5.2 Discussion

The results demonstrated that SVM-Morph (LiGTS) was superior in terms of error rate, precision, and specificity compared to the other two classifiers. Additionally, SVM-Morph exhibited high correlation and narrow limits of agreement with low mean difference compared to the values obtained from the manual ROI selection. These preliminary findings suggest that LiGTS is a promising tool to perform automated liver segmentation in clinical setup.
Compared to the LDA method, the SVM-based methods exhibited superior performance. In particular, SVM-based methods are more selective and do not incorrectly identify the pixels from the spleen as liver, which may have texture properties similar to that of liver. The added selectivity of SVM-based methods can be attributed to the effective utilization of the features, particularly the run length feature. For the run length feature, the values assigned to the liver pixels are large but not necessarily the largest; see Figure 5.1. LDA, being a linear classifier, is not equipped to handle the situation where feature values for one class are surrounded by the feature values for the other class. Being a nonlinear classifier, SVM can handle this situation by employing more complex decision boundaries.

SVM-Morph performs better compared to the SVM classifier. SVM-Morph and SVM classifiers are identical except that SVM-Morph employs additional
Figure 5.5: Bland-Altman analysis for the mean stiffness results from the three classifiers. x-axis indicates the pair-wise mean of the stiffness values from the classifier and the values obtained from the manual ROI selection, and y-axis indicates the pair-wise difference in the stiffness values from the classifier and the values obtained from the manual ROI selection.

morphological filters to exclude smaller regions. As observed from Figures 5.2 and 5.3, SVM-Morph is effective in eliminating isolated regions that have been incorrectly recognized by SVM as liver.

A recent study [1] used MRE images to segment the liver and report the mean stiffness values. In that study, Dzyubak et al. assumed that the liver is the largest connected region in the abdomen and used histogram partitioning and an active contouring technique to segment the liver. However, this method may not work well in the presence of low-gradient boundaries and intensity inhomogeneities. Also, this method may not be able to identify the vessels and other subtle features from liver parenchyma. As observed in the MRE stiffness maps, the vessels have lower stiffness values when compared to the liver parenchyma.
Therefore, inclusion of vessels may lower the mean stiffness value of the liver, potentially resulting in misclassification of early/bridging stage fibrosis. In contrast, our method can identify vessels based on their texture features, which are distinct from the liver parenchyma.

There are a few limitations in our study. First, only few training datasets \((n=24)\) were used in this study. More training data may be needed for a performance that is robust across different acquisition setups. Second, the testing of the classifiers was also performed on a small sample size \((n=24)\). Third, subjects used in our study are normal volunteers \((20-30\text{ years old})\) and may not provide a complete distribution of the population with varying body habitus, which might alter the texture features. Finally, additional studies need to be performed to understand the performance of the classifiers in patient populations with different liver diseases. Despite these limitations, we have demonstrated an initial feasibility of the proposed technique.

### 5.3 Future Work

Future work will involve: 1) increasing the sample size with varying body habitus for a more robust training of the classifiers; 2) including additional texture-based or geometric features; 3) applying the algorithm to segment the liver in different patient populations and report mean stiffness values; and 4) extending the application of this technique from MRE to other MRI applications, including T1 and T2 quantification and fat fraction assessment.
5.4 Conclusions

In conclusion, we have implemented and validated a technique, LiGTS, to automatically segment the liver and report mean stiffness values of the liver using MRE-derived stiffness maps. The preliminary results indicate that LiGTS exhibits low error rate (less than 4%) and the resulting mean stiffness values demonstrate an excellent correlation with the values obtained from a manual ROI selection. Therefore, LiGTS can potentially be used to automatically segment the liver and report different intrinsic parameters, such as T1, T2, stiffness, and fat fraction values.
Bibliography


[26] A. E. Bohte, J. R. van Werven, S. Bipat, and J. Stoker, “The diagnostic accuracy of us, CT, MRI and 1H-MRS for the evaluation of hepatic steatosis
compared with liver biopsy: a meta-analysis,” *European radiology*, vol. 21, no. 1, pp. 87–97, 2011.


