5-2016

Adaptive kernel estimation for enhanced filtering and pattern classification of magnetic resonance imaging: novel techniques for evaluating the biomechanics and pathologic conditions of the lumbar spine

Nicholas Vincent Battaglia
University of Tennessee - Knoxville, nbattagl@vols.utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Nicholas Vincent Battaglia entitled "Adaptive kernel estimation for enhanced filtering and pattern classification of magnetic resonance imaging: novel techniques for evaluating the biomechanics and pathologic conditions of the lumbar spine." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Biomedical Engineering.

Mohamed R. Mahfouz, Major Professor

We have read this dissertation and recommend its acceptance:

Richard D. Komistek, Adrija Sharma, Aly E. Fathy

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Adaptive kernel estimation for enhanced filtering and pattern classification of magnetic resonance imaging: novel techniques for evaluating the biomechanics and pathologic conditions of the lumbar spine

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Nicholas Vincent Battaglia
May 2016

To Aimee, Jack, and Eli

Acknowledgments

First and foremost I would like to express my gratitude to my graduate advisor Dr. Mohamed R. Mahfouz, whose experience, knowledge, and keen insights have proven invaluable to the success of my graduate career. His passion to science and relentless pursuit of novel frontiers are qualities I hope to emulate as I move forward. In addition, his belief in me as a student and fellow scientist has been instrumental in motivating me to accomplish my goals.

I am also grateful to Dr. Richard D. Komistek, Dr. Adrija Sharma, and Dr. Aly Fathy for serving on my committee and providing me with valuable feedback that aided in the completion of this dissertation.

I would also like to extend my appreciation to my fellow colleagues at the Center for Musculoskeletal Research (CMR) who – no matter how inane, mundane, or insane my commentary – always greeted me with respect and professionalism. I appreciate all of the hard

work that Rebecca Robertson contributed to helping me function efficiently within the CMR. I am greatly indebted to Lyndsay Bowers for going above and beyond any obligations or expectations of her to help me whenever possible.

I am eternally grateful for the unwavering support I have received from my immediate family. To my parents, Michelle and Daniel, for the guidance and emotional support; my brother, Andrew, for instilling courage and perseverance; my sister, Kelly, for the insight and wisdom (both bestowed and taught); and brother-in-law, Joel, for lessons in curiosity.

Finally, I could not have made this journey without the support and drive of my wife Aimee and our two children, Jack and Eli. I love you.

Abstract

This dissertation investigates the contribution the lumbar spine musculature has on etiological and pathogenic characteristics of low back pain and lumbar spondylosis. This endeavor necessarily required a two-step process: 1) design of an accurate post-processing method for extracting relevant information via magnetic resonance images and 2) determine pathological trends by elucidating high-dimensional datasets through multivariate pattern classification. The lumbar musculature was initially evaluated by post-processing and segmentation of magnetic resonance (MR) images of the lumbar spine, which characteristically suffer from nonlinear corruption of the signal intensity. This so called *intensity inhomogeneity* degrades the efficacy of traditional intensity-based segmentation algorithms. Proposed in this dissertation is a solution for filtering individual MR images by extracting a map of the underlying intensity inhomogeneity to adaptively generate local estimates of the kernel's optimal bandwidth. The adaptive kernel is implemented and tested within the structure of the *non-local means* filter, but also generalized and extended to the Gaussian and anisotropic diffusion filters. Testing of the proposed filters

showed that the adaptive kernel significantly outperformed their non-adaptive counterparts. A variety of performance metrics were utilized to measure either fine feature preservation or accuracy of post-processed segmentation. Based on these metrics the adaptive filters proposed in this dissertation significantly outperformed the non-adaptive versions. Using the proposed filter, the MR data was semi-automatically segmented to delineate between adipose and lean muscle tissues. Two important findings were reached utilizing this data. First, a clear distinction between the musculature of males and females was established that provided 100% accuracy in being able to predict gender. Second, degenerative lumbar spines were accurately predicted at a rate of up to 92% accuracy. These results solidify prior assumptions made regarding sexual dimorphic anatomy and the pathogenic nature of degenerative spine disease.

Table of Contents

Chapter 1 Background 1
Chapter 2 Literature Review 13
2.1 Noisy MRI Data 13
2.1.1 Intensity Inhomogeneity Correction (IIC) 14
2.1.2 Noise Correction (NC) 15
2.2 The Kernel 17
Chapter 3 Research Aims 19
Chapter 4 Adaptive Filtering 23
4.1 Understanding the Noise Model 24
4.1.1 Local Entropy Minimization by Bicubic Splines (LEMS) 26
4.1.2 Non-local Means (NLM) Filtering 27
4.1.3 The Coupled Algorithm: Bias Adaptive Non-Local Means (baNLM) Filtering 32
4.1.4 Exploring the Adaptive Kernel in Various Filters 36
4.2 Testing of the Adaptive Filters 38
4.2.1 Images and Simulated Corruption 39
4.2.2 Images and Simulated Corruption 41
4.2.3 Notation Convention 41
4.2.4 Performance Measures 42

4.3 Optimization and Parameterization of ***h*** *45*
4.3.1 Result of test images *47*
4.3.2 Results with Various Kernels *47*
Chapter 5 MRI of the Lumbar Spine **52**
5.1 Acquisition and Processing *53*
5.1.1 Semi-Automatic Segmentation *55*
5.1.2 Validation of baNLM on MRI Datasets *56*
5.2 Population Trends *60*
5.3 Gender Bias *66*
Chapter 6 Classification of Pathology **75**
6.1 Multivariate Analysis *82*
Chapter 7 Contributions and Future Work **90**
References **93**
Appendices **100**
Appendix A Convolution *101*
Appendix B Gaussian Smoothing *103*
Appendix C Anisotropic Diffusion *104*
Appendix D Pattern Classification *106*
D.1 Neural Networks *106*
D.2.1 Neuron Architecture *107*
D.2.2 Abbreviated Notation *110*
D.2.3 Backpropagation *111*
D.3 Recurrent Neural Networks *113*
D.4 Architecture *115*
A.2 MODULE 2: Dynamic Classification *117*
D.2 Number of Hidden Layer Neurons *118*
Appendix E Muscle Volume Data *120*
Appendix F Adipose Volume Data *126*

Vita **…….....132**

List of Tables

Table 1: Example of four frequently utilized filtering kernels and a version that includes the bias adaptive modification proposed in this dissertation. 37

Table 2: Notation for orders of operation. 42

Table 3: Average performance of the four filters on the synthetic phantom and Lena with various levels of noise. Statistically significant ($p<0.01$) performance values for a matched pairs t-test are indicated with the appropriate symbol (footnote). Bolded values indicate the filter that obtained the best average performance and that a test of significance was performed (see footnotes).....50

Table 4: Average (N=10) performances of the non-local means (NLM), anisotropic diffusion (AD), and Gaussian (G) filtering, and the bias adaptive (ba) counterparts on the Apollonian test phantom. Statistically significant ($p<0.01$) performance values for a matched pairs t-test are indicated with the appropriate symbol (footnote). Bolded values indicate the filter that obtained the best average performance and that a test of significance was performed (see footnotes).....51

Table 5: Population demographics and statistics.....53

Table 6: Classification accuracy percentages of test Images 1 and 2 (Figure 5.3). Reported are the percentages of pixels classified as: true positive (TP), true negative (TN), false positive (FP), false negative (FN), and percent accuracy 57

Table 7: Dice values for the mean differences over twenty-six subjects at each vertebral level and muscle group for baNLM against NLM 58

Table 8: Average volumetric (cm^3) data across the study population of three bilateral lumbar muscle groups shows a significantly different cross-gender comparison in all lean muscle groups for a Kruskal-Wallis test. In the adipose tissue, only the multifidus showed significant differences.....67

Table 9: Compression rations of the intervertebral spacing from the supine-to-standing positions. Bolded values indicate significantly different values for $\alpha = 0.01$. 81

Table 10: Means, standard deviations, and corresponding p-values of the intervertebral spacing for the healthy plus LBP groups versus the degenerative group. Bolded values indicate significant differences with a Kruskal-Wallis test. 81

Table 11: Prediction accuracy of all 17-features used to classify the population. 84

Table 12: Prediction accuracy of a single feature (L1–L5 distance) used to classify the population. 86

Table 13: Maximum classification accuracy using the optimized features of the reduced (5-feature) space. ...88

List of Figures

Figure 1.1: Simplified drawings of the spinal column showing the cervical thoracic, lumbar and pelvic regions of the spine (right). Annotated are the lumbar vertebrae, spinal curves, and intervertebral discs (left). ..3

Figure 1.2: Simplified sketch of L3 highlighting relevant anatomy of the vertebrae. 4

Figure 1.3: Sample MR slice at L4 with bulk muscle boundaries highlighted. The highlighted muscles include the bilaterally occurring: psoas (red), erector spinae (green), and multifidi (blue). (Note: the image has been brightened and contrasted to enhance relevant tissue boundaries. 7

Figure 4.1: MR images of the lumbar characteristically exhibit globally inhomogeneous signals with high SNR values near the posterior midline of the lumbar (a). The signal dissipates radially from that point along a smooth gradient, the extracted bias images resembles that of vignetting distortion observed in most optical system (b). A graph of the intensity in the out-of-plane axis (z-axis) reveals the distorted nature of spatially sensitive signal intensities (c). 25

Figure 4.2: A realistic bias field (a) is corrupted with 9% Rician noise (b), and then corrected for the intensity inhomogeneity with LEMS (c). Finally, local SNR values are calculated (d). As expected, the SNR image follows a similar trend to the intensity of the extracted bias field (a). 31

Figure 4.3: Three similarity windows aligned at various coordinates along the bias image (right). The mean pixel intensity contained in each window reflects the probability density (left) for parameterizing the adaptive kernel function. 33

Figure 4.4: The phantom and Lena (a, g) corrupted with a realistic inhomogeneity bias field (b, h) and 9% Rician noise (c, i). A zoomed in section of the original images (d, j), with the added field (e, k), and 9% Rician noise (f, l). 40

Figure 4.5: Two paths utilized for the analysis of baNLM an example MRI. The method for $\Pi\beta\lambda$ (top, red) employs LEMS prior to filtering, while the method for $\Pi\lambda\beta$ (bottom, green) employs LEMS before and after filtering. 44

Figure 4.6: The mean values over 10 iterations of the optimized image performance measures, FOM, SSIM, and *i*RMSE, at the various indicated noise levels for $\Pi\lambda\eta$ and $\Pi\eta\lambda$ for the synthetic phantom and Lena (a and c, respectively). The filer strength parameter, h, that achieved the highest mean metric measurements over 10 iterations provided an indication of the optimal value at the given noise levels for the synthetic phantom and Lena (a and c, respectively). 49

Figure 5.1: Sample MR slice at L4 with bulk muscle boundaries highlighted. The bilaterally occurring muscles are the psoas (red), erector spinae (green), and multifidi (blue). Note that the image has been brightened and the contrast increased to enhance the tissue boundaries. 54

Figure 5.2: The original MR image (a) and compliment of the extracted bias field (e). Shown are the unbiased image (b) and the unbiased image with bias adaptive NLM (f). Zoomed regions of the unbiased image are shown in (c) and (d), while (g) and (h) are the same zoomed regions for the bias guided filtering. Greater filtering was applied at (g) compared to (h), while little filtering occurred at (d) compared to (h). 58

Figure 5.3: Automatic segmentation accuracy of muscle tissue for two different MR images (a, b). Each image was processed independently with baNLM and NLM, and then classified with fuzzy c-means clustering. The classification results of the images from both filters are overlayed on a single image and color coded based on correspondence of the classification results. Green depicts regions where baNLM and NLM

were mutually correct, while blue indicates where they were mutually incorrect (b, d). Red indicates baNLM was correct and NLM was incorrect; while yellow indicates NLM was correct and baNLM was incorrect (a, c). Incorrectly classified pixels were identified as both false-positives and false-negatives, while correctly identified pixels were associated with only true-positive muscle tissue. The classifications from both images show baNLM (yellow) had significantly less mutually exclusive misclassifications than NLM (red) in both images (c). 59

Figure 5.4: Dice values for the mean differences over twenty-six subjects at each vertebral level and muscle group for baNLM against NLM. 60

Figure 5.5: Population average of the paraspinal muscle volumes at specific lumbar vertebral levels. 63

Figure 5.6: Gender specific relationships between percent adipose tissue and BMI. Percent adipose is a reflection of the content across the entire paraspinal column of the lumbar. Female subjects have positive correlation of adipose to BMI while males exhibited a negative correlation. 64

Figure 5.7: Observed gender differences in the muscle and adipose volume. Muscle volume was statistically less for females than males (left), while no significant difference was observed in the adipose volume (right). 65

Figure 5.8: Boxplot of the BMI separated by gender. No significant differences were observed between genders. 66

Figure 5.9: Three dimensional volumetric reconstruction of the lumbar spine and paraspinal musculature of a female (top) and male (bottom). The surface models show a significant gender differences in both the muscle girth and adipose tissue content. 68

Figure 5.10: Box plots depicting gender specific percent lean muscles values. Male individuals clearly have more lean paraspinal muscles than females. 69

Figure 5.11: Between gender percentages of adipose tissue in the paraspinal muscles of L1–L5. Female percentages were significantly greater than that of males and showed a slightly less steep regression line. 71

Figure 5.12: Combined bilateral (left and right) percent adipose tissue versus BMI with corresponding regression correlation lines for the psoas (a), multifidus (b), and erector spinae (c) muscles. 72

Figure 6.1: Distribution of lean muscle values across the three subject groups. Distributions and means are not significantly different from each other. 76

Figure 6.2: Boundaries of the bones are segmented in slices of a CT scan. Adjacent segmentations are joined together to generate 3D model of the vertebrae. 77

Figure 6.3: Overlay of 3D models on fluoroscopic imaging allowing for in vivo tracking of the vertebrae with six degree of freedom resolution during a flexion-extension activity (a). Overlay with separately colored surface models are tracked to determine the relative intervertebral kinematics during a lateral bending activity (b). 79

Figure 6.4: Hermine spline connecting the vertebral centroids by their 3D surface models. 80

Figure 6.5: The Curse of Dimensionality suggests that there exists an optimal number of features for classification; however, after exceeding this arbitrarily optimal number the performance experiences a sharp decline in accuracy. 85

Figure 6.6: Variance of individual principal components (bar) the cumulative sum of the first eight principal components (line). 87

Acronyms and Abbreviations

NIH:	National Institutes of Health
US:	United States
LBP:	Low Back Pain
FSU:	Functional Spine Unit
MRI:	Magnetic Resonance Imaging
MR:	Magnetic Resonance
OMT:	Osteopathic Manipulative Therapy
CT:	Computed Tomography
WLS:	Weighted Least Squares
SNR:	Signal-to-Noise Ratio
IIC:	Intensity Inhomogeneity Correction
RF:	Radio Frequency
LEMS:	Local Entropy Minimization with Bicubic Splines
NC:	Noise Correction
DFT:	Discrete Fourier Transform
NLM:	Non-Local Means
baNLM:	Bias Adaptive Non-Local Means

RMSE:	Root Mean Squared Error
FOM:	Figure of Merit
SSIM:	Structural Similarity Index
MMV:	Mean Metric Value
BMI:	Body Mass Index
ANOVA:	Analysis Of Variance
FCM:	Fuzz C-Means
CSA:	Cross-Sectional Area
TP:	True Positive
TN:	True Negative
FP:	False Positive
FN:	False Negative
CR:	Compression Ratio

Chapter 1

Background

According to the National Institutes of Health (NIH), approximately a quarter of the population in the US will experience at least one day of low back pain (LBP) during a 3-month span. Other estimates indicate that: 80-90% of people in the United States will experience LBP at some point in their lifetime [1-5]; up to 62% of people will report an episode within one to two years [3, 6, 7]; it is the third most common reasons for doctor visits [8]; it is estimated that chronic low back pain accounts for 33% of the cost of all workers' compensation claims [9]; and in terms of treatment costs and work days lost, the fiscal impact amounts to approximately $100 to $200 billion annually [10]. Yet, despite these burdens, the etiology and pathogenesis of LBP is largely unknown, as more than 85% of low back pain complaints are diagnosed as non-specific with unknown origins [11-13].

The lumbar spine is situated in the lower torso, arranged between the pelvic and thoracic regions

of the trunk. Along the spinal column, the lumbar region is comprised of five vertebrae—L1 through L5—which provide structural stability and mobility to the trunk (Figure 1.1). It is an exceedingly complex system of soft and boney tissues that function synergistically to propel the torso about the lower extremity.

Located between adjoining vertebrae are intervertebral discs that provide cushioning and support for the spine. The discs form a fibrocartilaginous joint that: (1) allow for modest six degree-of-freedom motion and (2) hold the vertebral bodies together with tensile forces that function similar to a ligament. In addition, the lumbar vertebrae are also joined by two sets of bilateral synovial joints at the inferior and superior facets (Figure 1.2). The facet joints provide structural support and limit the range-of-motion between adjacent vertebrae. This "articular triad" of two facet joints and the intervertebral disc form a functional spine unit (FSU). Although motion at a single FSU is relatively limited, articulation across multiple FSUs stacked on top of each other allow for considerable global motion.

The sacrum and coccyx are located at the base of the vertebral column (Figure 1.1), and are comprised of fused, rudimentary vertebrae. The sacrum contains five vertebrae (S1 through S5) that are generally fused by early adulthood. Inferior to the sacrum, the coccyx is comprised of between three to five vertebral analogs. Gray indicates that the bones of the coccyx are likewise joined by adulthood [14]; however, other more recent studies suggest that the coccyx is more likely to be constructed of two to three separate segments [15, 16]. Nonetheless, neither the sacrum nor the coccyx contributes motion to the trunk or lumbar.

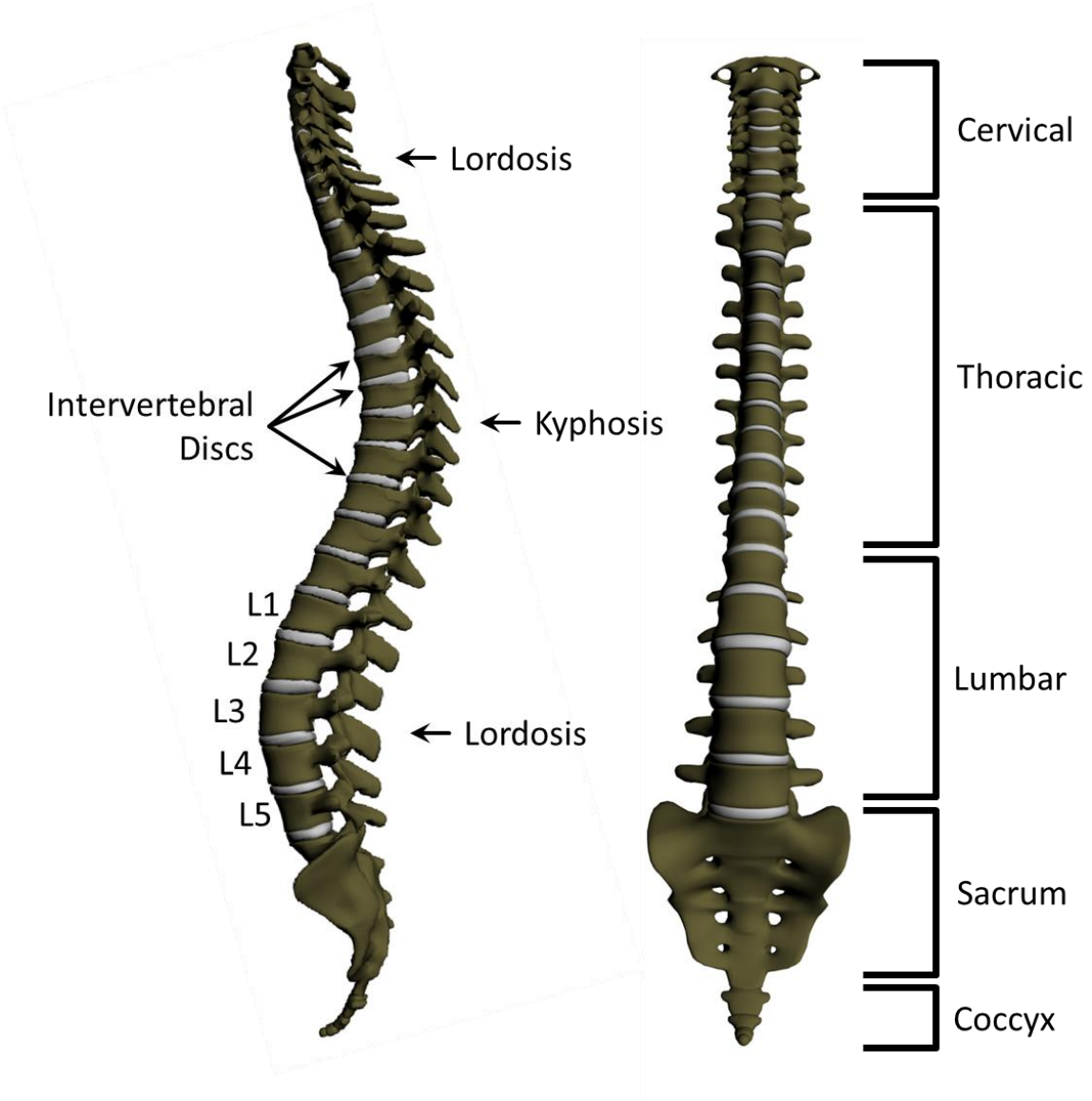

Figure 1.1: Simplified drawings of the spinal column showing the cervical thoracic, lumbar and pelvic regions of the spine (right). Annotated are the lumbar vertebrae, spinal curves, and intervertebral discs (left).

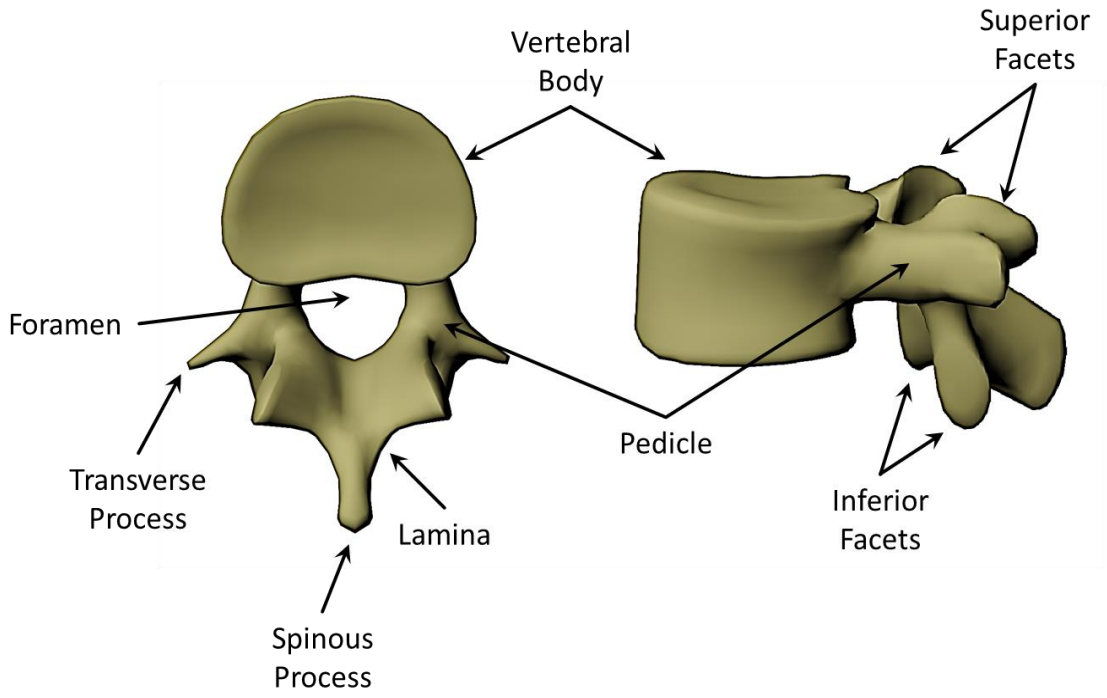

Figure 1.2: Simplified sketch of L3 highlighting relevant anatomy of the vertebrae.

This leaves the lumbosacral joint (L5-S1) as the biomechanical axis between the sacrum and trunk. Thus, it is common practice in kinematic analysis to represent motion of the lumbar vertebrae with a six degrees of freedom (three translational and three rotational) relative to the sacrum.

Another feature that is significant to the biomechanical function of the lumbar spine is the lordotic curve of the spine (Figure 1.1). As opposed to the kyphotic curve observed in the thoracic spine (Figure 1.1), lordosis is inwardly anterior curvature of the spine and is a biomechanically unfavorable configuration due to increased asymmetrical loading in the FSU. That is to say that the downward gravitational force is displaced onto the facet joints, while

forces are offloaded from the dampening properties of the intervertebral disc. This notion is made further apparent by epidemiological evidence showing significantly higher incidence of back pain at the lumbar and cervical levels compared to the thoracic [17]. These naturally occurring anatomical conditions leave the lumbar (and cervical) spine inherently more susceptible to injury.

In addition to the boney anatomy, the soft tissue, chiefly the musculature, that surrounds the lumbar vertebrae play a crucial role in the biomechanic function of the lumbar spine. The lumbar back muscles, often referred to as the *paraspinal muscles*, add structural stability to the spine and also act directly on the vertebrae themselves to assist in propelling the trunk. From the perspective of gross anatomy, muscles of the lumbar can be classified into three distinct groups: 1) psoas major, 2) quadratus lumborum and lateral intertransversarii, and 3) interspinales, intertransversaii medials, multifidi, erector spinae (longissimus and iliocostalis) [18]. However, strict anatomical classification of the lumbar paraspinal muscles becomes problematic when intermuscular fascial boundaries are not clearly delineated through medical imaging. This is often the case in MRI studies where poor superoinferior resolution distorts between slice interpolation of the shorter and less clearly defined muscles such as the medial and lateral intertransversarii, and the boundary between the longissimus and iliocostalis. To mitigate misclassifications, more encompassing bulk boundary definitions have been devised that group muscles not only by function, but also by proximity as well.

Strategies for evaluating bulk lumbar musculature via MRI centers around devising boundary definitions that can be clearly and consistently segmented across any given population set (Figure 1.3). Bulk analysis of lumbar MRIs reveals that three distinct bilateral muscle groups emerge which are easily identifiable across a large population base [14], these are the:

(1) Psoas Major (Figure 1.3, red): long spanning muscles that originate at the transverse processes of T12 through L5 and insert at the lesser trochanter of the femur. It lies anterior to the vertebral bodies and slightly lateral of midline. Its primary actions are to flex the hip joint and to bend the trunk forward (when acting bilaterally from the point of insertion) [19]. Any secondary action(s) of the psoas major is not well understood and has incited significantly more controversy [20-34]. Hansen *et al* [18] provide a concise elucidation of the numerous hypotheses that have been proposed, which have more or less reached a consensus "that the psoas major probably functions as a stabilizer of the lumbar lordosis in upright posture." As for imaging, axial cross-sections of the psoas major show no major intersection between it and other prominent anatomical features and, thus, are generally segregated from nearby tissues with little ambiguity and high accuracy.

(2) Multifidi (Figure 1.3, blue): part of larger system of muscles known as the transversospinales. The muscles that are contained within the lumbar region of the transversospinal system include the: multifidi, rotatores, interspinales, and intertransversarii. The multifidus is the most prominent and longest spanning of the

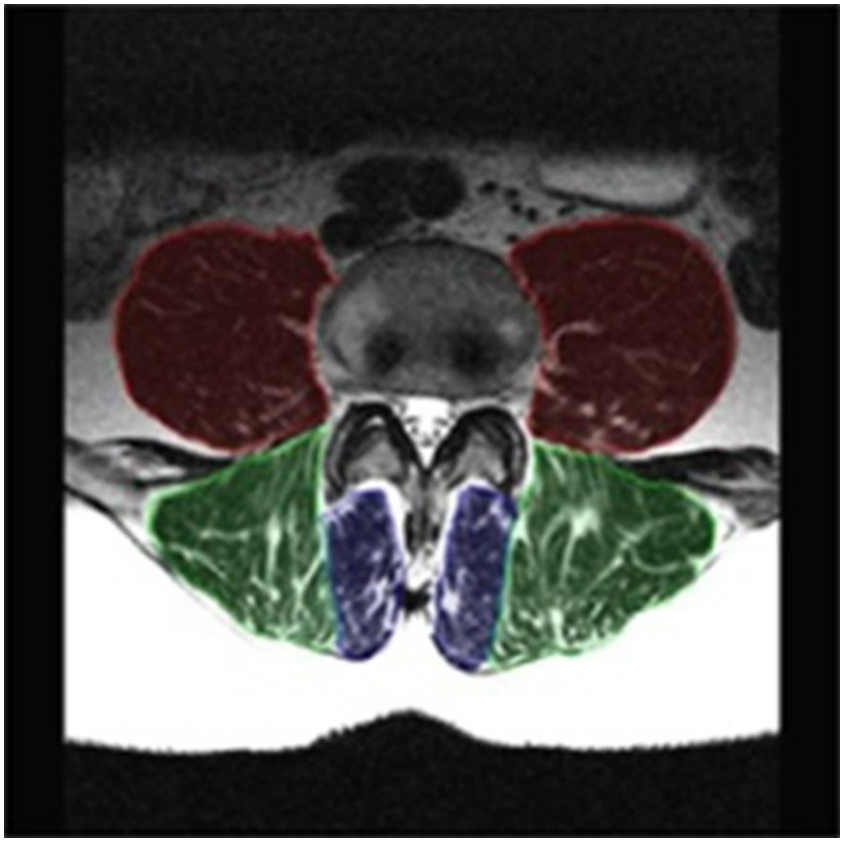

Figure 1.3: Sample MR slice at L4 with bulk muscle boundaries highlighted. The highlighted muscles include the bilaterally occurring: psoas (red), erector spinae (green), and multifidi (blue). (Note: the image has been brightened and contrasted to enhance relevant tissue boundaries.

transverospinales, while the other three muscles are in general short and only span 1-3 vertebrae. The multifidi are situated in the lamina groove formed by the large transverse and spinous process where the form fascicles that originate at the mammillary processes of inferior vertebrae and insert to the spinous process of the vertebrae one-three vertebrae above. In addition, both the rotatores and the interspinales have insertion points along the spinous processes of the lumbar. Accurate differentiation of these three muscles proximal to insertion is difficult due to insufficient resolution or clear definition of boundary lines. However, what are clearly delineated are the lateral boundaries of the erector spine and the anterior and medial borders of the boney anatomy of the transverse and spinous processes. Since it can be difficult to differentiate the three muscles, it is common practice for researchers to simply label them collectively as the multifidus. This technique reduces misclassifications, while also simplifying biomechanical modeling and analysis. The last muscle of the transversospinal system in the lumbar region, the intransversarii, is a short muscle occupying the space between the transverse processes of adjacent vertebrae (laterales) and also the space from the accessory process of a vertebra to the mammillary process of the vertebra below (mediales). Gray proposed that the function of the multifidi were to assist the vertebrae in extension, lateral flexion, and rotation [14]. Later studies have indicated that when viewed as singular muscles their role appears to produce stabilizing forces rather than function as the primary drivers of motion [35]. Further studies were conducted via electromyography which supported that the multifidi's primary function is to control intersegmental motion [36-38].

(3) **Erector Spinae (Figure 1.3, green)**: a large muscle that encompasses the cervical, thoracic, and lumbar regions of the spine [14]. Somewhere along the upper lumbar region the muscle splits into three columns: a lateral (iliocostalis), an intermediate (longissimus) and a medial (spinalis). In the lumbar bulk regions of the erector spinae lie in the groove formed by the transverse and spinous processes, lateral to the multifidus. This is where they are additionally divided into the superficial (longissimus) and the deep (iliocostalis) layers. Specific location of branching of the erector spinae is largely ambiguous between subjects in a given population. Thus, for comparative analysis it is more robust to simply classify the corresponding muscular and tendinous regions by a single, clearly defined anatomical feature (i.e. the outer fascial boundary). Unilaterally flexion of the erector spinae bends the spine laterally, while bilateral flexion generates posterior sagittal rotation. In this respect, they supplement the multifidi by resisting flexion caused by the abdominal muscles during rotation of the trunk [39].

It is important to note the absence of the quadratus lumborum in the description of the transversospinal system provided above. The quadratus lumborum is not used in the analysis provided by this manuscript; still it remains a significant contributor to the functionality of the lumbar [14, 18]. The quadratus lumborum originates at the iliac crest and inserts to the apices of the transverse processes of the first four lumbar vertebrae [14]. It has been suggested that it assists with lateral flexion of the upper four lumbar vertebrae [40, 41].

A wide range of diagnoses reveal themselves during diagnosis of LBP. It turns out that some discernable etiologies stem from mechanical changes such as deformities in the boney anatomy or visible degeneration of the intervertebral discs. Imaging studies can quickly reveal definitive diagnoses for which unambiguous algorithms are applied for treatment. However, what is perhaps of greater interest are patients that present with non-specific low-back pain that lack any definitive diagnostic evidence. In these situations, the most likely culprit could be soft tissue injuries or anomalies that do not present during imaging.

Low back pain is distinguished broadly as being either acute or chronic – depending on the length of time that pain has persisted. Acute back pain is defined as not lasting for longer than 4 weeks, while sub-acute back pain is defined as lasting between 4–12 weeks. Acute low back pain with sudden onset is the most common and often stems from a single traumatic event such as a fall or improperly lifting a large load. Treatment of acute back injury is most often approached with conservative measures and may include: over-the-counter analgesics or non-steroidal anti-inflammatory drugs (NSAIDs), heat, limited exercise, time, or osteopathic manipulation. However, though a majority of low back pain will resolve without further intervention, approximately 20% of acute injuries progress beyond the sub-acute phase and into the chronic period, with symptoms often persisting for over a year after onset.

Chronic back pain is defined by persistence for 12 weeks or longer. Similar to acute injury, chronic pain can be initiated with a single traumatic event; however, chronic symptoms often stem from structural defects that take years, if not decades, to present. Therefore, the

pathogenesis of chronic pain is mechanical in nature and will require some sort of interventional therapy. Since a plethora of sources are responsible for the chronicity of low back pain, it is perhaps clearer to present it as a function of treatment rather than diagnosis. Depending on the post-acute diagnosis, treatments can be broadly grouped into two types of treatments:

(1) **Non-surgical** therapies are performed during the sub-acute period and are preferred to their counterpart, as non-surgical methods are: cheaper, less invasive, and have shown significant efficacy for specific diagnoses. Some common non-surgical treatments include: physical therapy and OMT, where improved joint flexibility and muscle strength are sought; narcotic and non-narcotic analgesics, NSAIDs, and muscle relaxants; traction, where the vertebrae are pulled slightly apart to alleviate compression of the intervertebral discs; behavioral modification; and, lastly, epidural steroid injections at the nerve root, which are the most invasive of the non-surgical treatments. Often 2-dimensional (2D) radiographic studies provide better insight and can be used to rule out any specific systemic etiologies. In terms of overall patient care, clinicians may utilize a shotgun approach that will include a combination of any of these therapies. Finally, upon successful remediation of pain, treatment is halted.

(2) **Surgical** therapy remains as the final and most drastic option for alleviating chronic back pain. Some common diagnoses that elicit surgical intervention include: herniated disc, spinal stenosis, spondylolisthesis, vertebral fractures, and discogenic disease. In addition to the radiographs captured during the sub-acute period, pre-operative imaging includes

either computed tomography (CT) or magnetic resonance imaging (MRI) studies. The imaging modality utilized is specific to the diagnosis and may even be patient specific (e.g. metallic arthroplasty hardware or implanted devices may prevent use of an MRI). In addition, intraoperative imaging, such as real-time fluoroscopy, is also an option.

One unfortunate aspect to a shotgun approach is that the specific etiology is never determined and the incidence is quite simply classified as “non-specific”.

Though the occurrence and fiscal impact of LBP has been well documented in the literature, no clear consensus has been reached on the pathogenesis or etiology. This is particularly evident with nonspecific LBP where diagnosis, treatment, and prognosis can be highly variable. Ultimately this lack of a clear medical outcome is a result of a lack of understanding behind the complex nature of LBP and progressive nature of degenerative spine diseases.

Chapter 2

Literature Review

Drawing correlation from soft tissue data of the lumbar spine to any specified pathology first requires accurate segmentation of anatomically relevant features. Through the use of sophisticated medical image processing algorithms this information is readily available in the form of MR images, but from the perspective of implementation, accessing that data is problematic due to inhomogeneous signal intensity inherent to MRI technology. Developing a reliable algorithm for accurately segmenting this data requires a thorough formulation of the noise models that confound MRI of the lumbar spine.

2.1 Noisy MRI Data

A common form of corruption encountered in magnetic resonance imaging (MRI) of the lumbar spine is spatial inhomogeneity of the signal intensity, or intensity inhomogeneity. Intensity inhomogeneity presents difficult obstacles for many modern noise filters, primarily due to an

assumption of a globally homogenous signal-to-noise ratio (SNR). In the presence of local intensity perturbations, static filtering can lead to over or under filtering that spatially mimics the local shifts in SNR. Therefore, when confronted with intensity inhomogeneity it may often be more desirable to utilize adaptive filtering that is sensitive to local changes in signal strength.

2.1.1 Intensity Inhomogeneity Correction (IIC)

The development of IIC methods is not a trivial task and has been the principle focus of researchers for decades [42-44]. Though a number of sources have been attributed to the spatial intensity degradation of MR images such as inappropriate coil tuning, gradient eddy currents, radio frequency (RF) standing wave effects, and RF penetration effects – the predominant source of corruption has generally been linked to spatial inhomogeneity sensitivity of the RF receiver coils [45]. Intensity inhomogeneity manifests when the receiver coil suffers from a steep decline in sensitivity in the direction of increasing tissue depth.

A plethora of IIC methods have been proposed in the literature, [46] provides a review and classification of these methods. Two IIC philosophies that have garnered recent attention are: information minimization, [47, 48]; and histogram matching [49, 50]. A hybridization technique was proposed by Salvado et al. [51], which merges ideas from information minimization and histogram matching to form what they term *Local Entropy Minimization with bicubic Splines* (LEMS). LEMS was utilized in the ideas proposed in this dissertation and more detailed description of the algorithm in Section 2.1.

2.1.2 Noise Correction (NC)

The general consensus behind electronic noise is that it arises during the reformulation process of the raw analog signal. MR images are most commonly reconstructed from their raw analog signals by computing the inverse discrete Fourier transform (DFT) of the k-space [52-54]. Because of difficult to control phase errors in the acquisition process, the signal measurements contains both real and imaginary parts of the DFT. Therefore, it has been shown that the noise component of the raw image signal arises from two independent Gaussian random variables (originating from the real and imaginary channels) [52-54]. By definition, the noise originating from two independent Gaussian random variables has Rician distribution.

Similar to IIC, a multitude of methods and philosophies have been developed over the years for MR NC. A review by Milanfar [55] provides a clear elucidation of the similarities that exist between most modern image filters. The author points out that the distinction is found merely in the specific formulation of the kernel function, while all other parameters and function remain more or less the same. In this regard, the kernel function is essentially an evaluation of pixels between their: spatial distance, photometric distance, or spatial and photometric distance. Our analysis focuses on the non-local means (NLM) filter, which uses photometric distances to develop the kernel [56, 57]. However, the selection of NLM was mostly arbitrary and the analysis could theoretically be extended to any of the kernels listed above. A more detailed discussion of the NLM filter is presented in Section 4.1.2.

The development of the IIC and NC methods were by in large designed as independent entities tailored to one specific artifact or the other. However, the crux of our analysis focuses on the situation when confronted with both artifacts in the same image, and, more specifically, how to perform filtering with global spatial intensity non-uniformity. Various techniques have been proposed in the literature to combat such a scenario, such as coupling a filtering parameter calculated from local noise statistics with an anisotropic filter [58]. Another study developed a variational approach that utilized a half-quadratic optimization [59] NC method by adaptively fitting the extracted bias field through simultaneous body and surface coil images [60]. Kervrann and Boulanger [61] developed a spatially adaptive filter, similar in structure to the NLM filter, that optimizes the weighting function by iteratively growing the search window and adaptively updating the filtering parameter.

In developing an algorithm that encompasses both IIC and NC it is also important to consider the order in which the corrections methods are applied. That is, is it more optimal to perform IIC first and then apply NC, or is the inverse procedure more optimal? Previous studies have concluded that IIC should precede NC, since IIC can disrupt the homogeneity of the embedded image noise [62, 63]. However, it is worth noting that these studies employed denoising methods with static filter strength parameters and, thus, did not investigate the relevance of adaptive noise filters in the image processing sequences.

A common measurement model for describing the correlation between the observed noisy image and the latent signal of the true image:

$$y(i) = x(i)b(i) + n(i) \quad \textbf{(1)}$$

where, at the i-th pixel, $y(i)$ is the observed signal, $x(i)$ is the signal of interest, $b(i)$ is the intensity inhomogeneity, and $n(i)$ is noise.

Solving for (10) elicits the question of which correction method should be performed first to achieve the most optimal results. Previous studies have concluded that IIC should precede NC, since IIC can disrupt the homogeneity of the embedded image noise [63, 64]. However, these studies employed static denoising methods and, thus, did not investigate the relevance of adaptive filtering.

The aims of this study were twofold: first was to modify a non-local means filter to adaptively update its kernel weight with values based on estimates of the underlying bias noise. Using this technique, the strength of filtering is adaptively modulated by local SNR values, where low SNR regions experience more filtering and high SNR regions incur less filtering. The second aim was to determine the order in which LEMS and adaptive NLM should be performed. We evaluated the performance of both aims with a synthetic phantom and MR images of the lumbar spine.

2.2 The Kernel

A filtering kernel, sometimes referred to as a convolution matrix, is a useful tool frequently employed in image processing for smoothing, blurring, sharpening, and embossing, to name a few. It is a small matrix that is convolved (Appendix A) with the original image to achieve the

desired effect, such as a filtering (i.e. smoothing). Returning to Equation (1) and for the purposes of the formulizing the smoothing kernel, the bias signal, $b(i)$, and the latent signal of interest, $x(i)$ are combined by elementwise multiplication such that $x(i)b(i) = z(i)$. The algorithm most frequently employed for solving for the true signal of interest is the weighted least squares (WLS) approximation. Adopting the notation used by Milanfar [55], the WLS approximation of the true signal $z(v_j)$ at position v is defined as:

$$\hat{z}(v_j) = \arg\min_{z(v_j)} \sum_{i=1}^{n} [y_i - z(v_j)]^2 K(v_i, v_j, y_i, y_j). \quad \textbf{(2)}$$

where $y_i - z(v_j)$ is the error and $K(\cdot)$ is the kernel, or weighting function, defined with respect to the indices i and j. $K(\cdot)$ Is a symmetric, unimodal function that measures the similarity between the samples y_i and y_j at locations v_i and v_j, respectively. This representation of the WLS is extremely convenient as it demonstrates how the kernel is clearly parsed from the error expression so as to simplify and promote a discussion of how the kernel affects the performance of the filter. Milanfar goes on to construct four generalized types of filters that differ only in the design of the kernel, which utilize the spatial and/or photometric distances [55]. Here, three of those kernel models are modified and juxtaposed against the adaptive counterparts proposed later in this dissertation (Table 1).

Chapter 3

Research Aims

The primary goal of this dissertation was to investigate the etiologic and pathogenic characteristics of LBP, in addition to searching for possible links that would expose LBP as a progressive disease that eventually degrades overtime leading to a host of more serious diagnoses and treatments. In order to determine a causal link between LBP and progressive degeneration, spines with degenerative disease were also investigated in this work.

The best and most accurate *in vivo* method for drawing etiological inference of soft tissue involves extracting information via clinical imaging. Numerous hypotheses suggest a possible correlation between musculature and adipose tissue of the lumbar spine [65-69]. Differentiation between these tissues is clearly visible in T_2 MRIs where they appear with deeply contrasted edges, as muscle appears dark, while adipose is nearly white. However, one significant drawback to T_2 MRI of the lumbar spine is the presence of a globally inhomogeneous signal-to-noise ratio

(SNR). This results in deleterious distortion of image quality, which in turn severely hinders the fidelity of intensity based segmentation algorithms.

One noteworthy aspect of segmentation in medical imaging is that there is generally a preprocessing step that involves smoothing electronic noise which significantly boosts the performance accuracy. This does not pose any hindrance in most applications of medical imaging. However, a serious problem is invoked when an image is confounded with both intensity inhomogeneity and electronic noise. Neither IIC nor NC methods were originally designed to compensate for both forms of distortion. Therefore, it is the primary aim of this dissertation to propose, test, and validate a method for reliably smoothing electronic noise when in the presence of global intensity inhomogeneity. Whereby developing an algorithm that successfully merges the two concepts (IIC and NC) into a more singular image processing technique. Namely, the design of a filter that utilizes information gained from IIC process and utilizes it to enhance filtering for NC.

The proposed filter was designed with adaptive properties that determine local SNR values based on the underlying intensity bias and adjusts the strength of filtering accordingly. The framework of this dissertation is built around the non-local means (NLM) filter, a smoothing technique that is popular in the medical imaging community. Thusly, the technique proposed in this manuscript is termed *bias adaptive non-local means* (baNLM) filtering. The efficacy of baNLM is measured against traditional NLM filtering by way of testing with a synthetic phantom, a common black-and-white photograph (Lena), and T_2 MRIs of the lumbar spine.

Another primary aim of this dissertation was to elucidate any etiologic and/or pathogenic information of LBP and spine degeneration within the study population. As was previously stated, the objective was not to make sweeping generalization about LBP, but rather to take a small sample size, acquire copious amounts of data on each subject, and attempt to expound correlations that in the future will provide some clarity with *where* to look and *what* to look for in larger population based studies. Because, as it currently stands, neither *where* nor *what* have been explicitly defined with any degree of certainty that would be acceptable in the clinical setting.

A secondary, and perhaps coincidental, objective of this dissertation was to investigate differences that exist between male and female lumbar spines. Though sexual dimorphisms are not radically profound in the world of biomechanics, they do represent a particularly noticeable exemption in the literature surrounding the spine. While not wholly devoid of research, any conclusions that may have been drawn are insufficient for in the clinical setting [19, 70-74].

Finally, it would be a mistake to assume that the scope of this dissertation alone could be extended to the diagnostic algorithms currently employed by clinicians, or that it might draw sweeping conclusions regarding some "grand cure" to low back pain. Such an undertaking would indeed require extensive studies with a wide population base and shrewd analytical techniques (the type reserved for most of the serious scientific and clinical enquires out there that use legitimate population statistics to posit generalized inferences). Rather, the breadth and scope of this dissertation aims to utilize a smaller sample size, with which a wealth of data was collected,

and distill out clinically relevant findings as they pertain to soft tissues of the lumbar. This is ultimately done with the hope that the ideas developed here may later be extended to someday enhance treatments and clinical outcomes of lumbar spine related disease.

Chapter 4

Adaptive Filtering

As was demonstrated in Chapter 2, MR images of the lumbar spine quite often suffer from a globally inhomogeneous signal-to-noise ratio (SNR) throughout the field of view, which can severely obstruct the robustness of modern post-processing imaging techniques often implemented for filtering and segmentation. The primary premise of the research presented in this manuscript hinges on the accuracy of segmenting muscle data from MRIs, thus it becomes of vital importance to develop a robust filtering algorithm for medical images that can aptly manage globally inhomogeneous SNRs that inherently hinder intensity based methods for segmentation. This chapter focuses on the theory, implementation, and testing of the development processes behind designing an image filter that can adaptively and optimally smooth regions based on their local SNR values.

4.1 Understanding the Noise Model

T_2 MRI of the lumbar spine notoriously suffers from a shift in intensity that tends to decrease laterally and anteriorly from most posterior midline of the back (Figure 4.1(a, b)). As previously mentioned, this shift is due to a one sided coil that is situated between the subject and the scanning table. The purpose of the coil (or any MR coil for that matter) is to amplify tissue signals to make for clearer reading by clinicians, yet since the lumbar coil is only one-sided, the quality of the signal quickly declines the further away the tissue is from the coil. It is worth noting that coils used in, for example, brain and knee scans completely encompass the anatomical region of interest and thus avoid problems associated with inhomogeneous signal strengths. Nonetheless, the typical lumbar MR scan will quite reliably have high SNR the nearest the most posterior midline aspect of the lumbar, by the spinous processes, and will quickly fade moving anteriorly and laterally (Figure 4.1(b)).

The effect of inhomogeneity severely distorts relative intensity values between tissue types. For example, the image used in Figure 4.1(c) shows fatty tissue captured in the posterior spine (labeled “High SNR”) with near maximum pixel values of around 255, whereas tissue captured anterior to the spinal column (labeled “Low SNR”) show fatty tissue with pixel values around 20. Ideally these fatty tissues would have similar pixel values to make it easier to distinguish from other tissue types. Alas, it is understandable why this does not pose a significant problem in the clinical setting as the human brain is readily capable of interpolating variable tissue signals by recognizing that though their intensity may appear different in different locations, the tissues in both regions are in fact the same, the underlying signal strength has simply undergone some

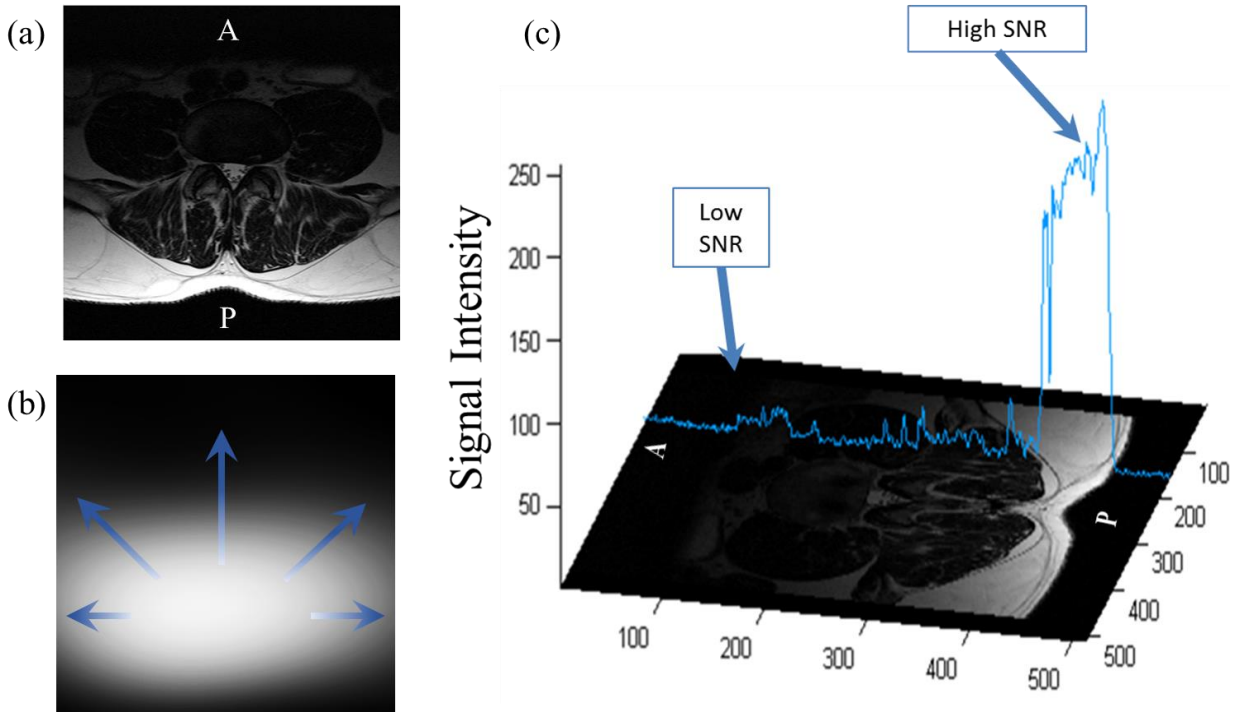

Figure 4.1: MR images of the lumbar characteristically exhibit globally inhomogeneous signals with high SNR values near the posterior midline of the lumbar (a). The signal dissipates radially from that point along a smooth gradient, the extracted bias images resembles that of vignetting distortion observed in most optical system (b). A graph of the intensity in the out-of-plane axis (z-axis) reveals the distorted nature of spatially sensitive signal intensities (c).

degradation along a gradient. It turns out that through visual inspection a clinician is able to quickly and efficiently segment relevant tissue boundaries to conclude a diagnoses. However, the unfortunate realization quickly settles in that intensity based segmentation algorithms are not nearly as clever as the human brain.

Formulizing the noise model associated with MR imaging is necessary for understanding how best to approach image reconstruction when faced with this sort of distortion. The first step is to investigate the nature of MR based inhomogeneous signals and how that correlates with relative SNR values. The discussion begins with a brief introduction on the method that was utilized in this research for intensity inhomogeneity correction (IIC). As was described in Chapter 2, numerous methods for correcting images with globally inhomogeneous signals have been proposed in the literature; however, LEMS has shown promise in the domain of MR imaging and also provides a practical framework from which to build upon.

4.1.1 Local Entropy Minimization by Bicubic Splines (LEMS)

LEMS is an optimization algorithm that determines the global background bias field through noise estimates in local sub-regions of the global image. The authors posit that the entropy of an image is a relevant analogy by which to represent noisy image, and that by estimating the entropy in smaller sub-regions of the image, these entropies can be ranked to generate a piecewise representation of local SNR values. This idea is formalized by adopting the variable definitions outlined in Eq. (1) for a noisy image, where the entropy, S, is given by: $S = -\sum_{\mathcal{N}\in\{gray\ levels\}} PDF_X(\mathcal{N}) log[PDF_X(\mathcal{N})]$, where $PDF_X(\cdot)$ is the probability density function

of X and $\mathcal{N}$ is a window containing a targeted local sub-region. The values of $PDF_X(\mathcal{N})$ are approximated by a histogram of the pixels in the current region divided by the number of pixels in that region. LEMS optimizes the entropy in a piecewise fashion by calculating the entropy in local square neighborhoods (or knots). The knots are then ranked based on their SNR. At the first knot, the bias field is optimized by minimizing the entropy in the local region. Next, the second highest SNR knot is optimized for both the current and previous regions; the third knot is then optimized for the first, second, and third regions and so on, until all knots have been optimized. Concurrently, the local bias fields are stitched together with a bicubic spline that smooths the gradient of the global bias field. The final result is an estimation of the intensity bias the distortion observed the original image, which is also an estimation of $b(i)$ from Equation (1) and the restored image is determined by: $x_{LEMS}(i) = y(i)/b(i)$.

At this point, the most obvious question is: why not now simply pass the undistorted image $x_{LEMS}(i)$ through one of the filtering algorithms that is commonly found in the medical image processing repertoire? The answer requires some insight into the theoretical basis of image filtering and the fundamental assumptions used to formulize them.

4.1.2 Non-local Means (NLM) Filtering

As was described in Chapter 2, this manuscript largely addresses kernel guided filters that utilize the weighted least squares framework outlined in Equation (2). That is not to say that other such frameworks would not be applicable in this context, the weighted least squares simply casts a wide net over many modern filters. In addition, the least squares framework also allows for one

to readily utilize photometric and/or geometric features to populate the kernel, Milanfar discusses this implementation in greater detail [55]. Nonetheless, one popular filter that uses this framework is the non-local means (NLM) algorithm [57, 75]. The NLM filter prepares a patched non-local similarity kernel to compute the photometric distance between other local patches, while ignoring the geometric distance between those patches. Again, adopting the variables defined in Equation (1), the formulization of NLM is detailed below.

Given a noisy image $y = \{y(i)|i \in I\}$, the true signal is estimated by a weighted average of all pixels in the image, $NLM[y](i) = \sum_{j \in I} w(i,j)y(j)$, where $\{w(i,j)\}_j$ represents the family of weights that depend on the similarity between pixels i and j. This measures the grayscale similarity between the vectors $y(\mathcal{N}_i)$ and $y(\mathcal{N}_j)$, where $\mathcal{N}_k$ denotes a square neighborhood (or the *similarity window*) centered at the kth pixel. The similarity is a measurement of the photometric distance, $\|y(\mathcal{N}_i) - y(\mathcal{N}_j)\|_{2,a}^2$, where $a > 0$ is the standard deviation of the Gaussian kernel. The kernel is defined as:

$$w(i,j) = \frac{1}{Z(i)} e^{-\frac{\|y(\mathcal{N}_i) - y(\mathcal{N}_j)\|_{2,a}^2}{h^2}}, \quad \textbf{(3)}$$

where $Z(i)$ is the normalizing constant:

$$Z(i) = \sum_j e^{-\frac{\|y(\mathcal{N}_i) - y(\mathcal{N}_j)\|_{2,a}^2}{h^2}}. \quad \textbf{(4)}$$

The parameter h dictates the strength of filtering by controlling the exponential decay of the kernel function. Larger values of h flatten the decay of the weighting function, thus causing greater filtering. Smaller values of h sharpen the decay, forcing the weights to zero and causing less filtering. The parameterization of h is the principle point of interest in our analysis and provides the connection between image denoising and intensity inhomogeneity.

Ideally NLM compares the similarity of neighborhoods throughout the entire image. However, from the perspective of practicality such a process would be far too computationally expensive. A smaller *search window* is employed that encompasses the similarity. The idea is that the search window should be large enough to determine non-local neighborhoods, while still small enough to yield reasonable computation times. As was suggested by Baudes, the search window was set to be 21 x 21 and the similarity window was set to 7 x 7. These window sizes were utilized throughout the entirety of this dissertation. With the 512 x 512 MR images used in this research, the iterative complexity of the final algorithm is calculated as 512^2 x 21^2 x $7^2 \sim 5.7$ x 10^9. This brings to light one of the primary drawbacks to NLM that even in its abridged form it is still computationally impractical for many uses.

Nonetheless, one benefit to the NLM algorithm is that it does not require many external parameter definitions by the user, in fact only three. Two of which are the search and similarity window sizes. Those are not the primary focus of this analysis and so were left unchanged from what the original authors proposed. The last, which is of perhaps more interest and pertinence to this manuscript are the properties and assumptions associated with tuning the parameter h.

The assumption behind h is that it will be parameterized with a value proportional to the noise level present in the image. So, an image with more noise will require an increased value of h that flattens the decay of the kernel function. However, a singular value of h assumes a globally homogeneous noise model throughout the image, which is certainly not the case in the present lumbar MR images and thus poses a significant predicament with regards to efficacy. This dilemma is best understood by understanding the changes that occur during the LEMS reconstruction process. Begin, for example, with a blank median image that is corrupted (by vector multiplication) with a realistic inhomogeneity bias (Figure 4.2(a)) and then 9% Rician noise is added to the image (Figure 4.2(b)). Now say the intensity inhomogeneity bias is corrected for using the LEMS method (Figure 4.2(c)). At this point, what is of particular interest is the noise model present in the post-processed image. The effects of LEMS can be observed by displaying the output image as a map of local SNR values (Figure 4.2(d)).

The goal now becomes determining how to filter the output image in Figure 4.2(c) for optimal denoising. From the image provided in Figure 4.2(d), it is quite apparent that image suffers from a wide range of SNR. Thus, if a static value of h were utilized in NLM most local regions would not experience optimal smoothing and invariably undergo over- or under-filtering. Finally, we observe why the basic assumption of NLM (and almost all non-SNR adaptive filters) is inappropriate for usage with lumbar MRIs; and, further, why over/under-filtering would be far less than optimal in the context of accurate segmentation. It would rathe be significantly more desirable if a system were devised that accurately controls the decay of the kernel to reflect local SNR values. This could be accomplish by designing h such that it was modifiable during -

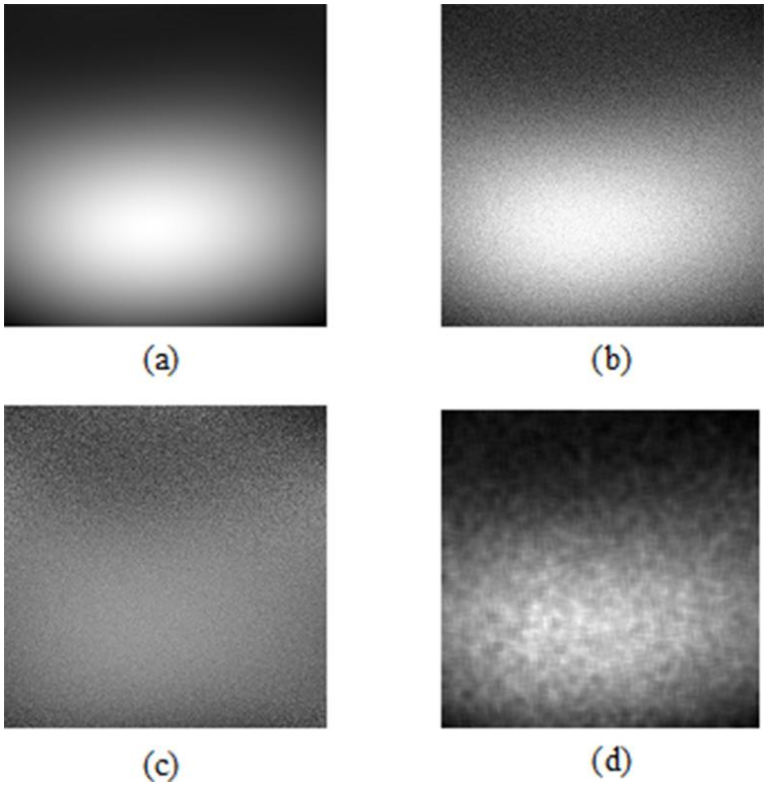

Figure 4.2: A realistic bias field (a) is corrupted with 9% Rician noise (b), and then corrected for the intensity inhomogeneity with LEMS (c). Finally, local SNR values are calculated (d). As expected, the SNR image follows a similar trend to the intensity of the extracted bias field (a).

runtime, and if it was also intelligently sensitive to local changes in SNR and adapt accordingly.

Figure 4.2(d) shows that the SNR bears close resemblance to the original bias image observed in Figure 4.2(a). In fact, the research presented in this manuscript, posits that the bias and SNR images bear such close resemblance that the underlying bias image provides a detailed map to adaptively parameterizing NLM (i.e. h) such that it accurately reflects local SNRs. From an algorithmic point of view, the system might proceed as follows: 1) the image of interest is passed to LEMS which extracts the underlying bias, 2) the bias image is passed along to some filtering scheme, and 3) the filter adaptively adjusts h using the bias image as a map. The resultant would be an image with smoothed regions that correspond with local SNR values.

4.1.3 The Coupled Algorithm: Bias Adaptive Non-Local Means (baNLM) Filtering

As mentioned in Section 4.1.2, the constant h in (3) and (4) is a tuning parameter that adjusts the level of smoothing imposed by the filter (lower values incur less smoothing, while higher values experience more smoothing). In various applications of NLM, h is generally parameterized as a function of the global standard deviation of the noise. This is, however, impractical for lumbar MRIs where: (1) the noise variance is not explicitly known and (2) there is poor spatial homogeneity of the SNR. In the context of external inputs, the extracted bias field is directly applicable to the parameterization of NLM (Figure 4.3). This notion is illustrated by Figure 4.3 which shows the correlation between the probability density of the kernel and local SNR values. The green similarity window indicates an approximate mean value of the image and perhaps encompasses some supposed theoretical region with “average” intensity characteristics.

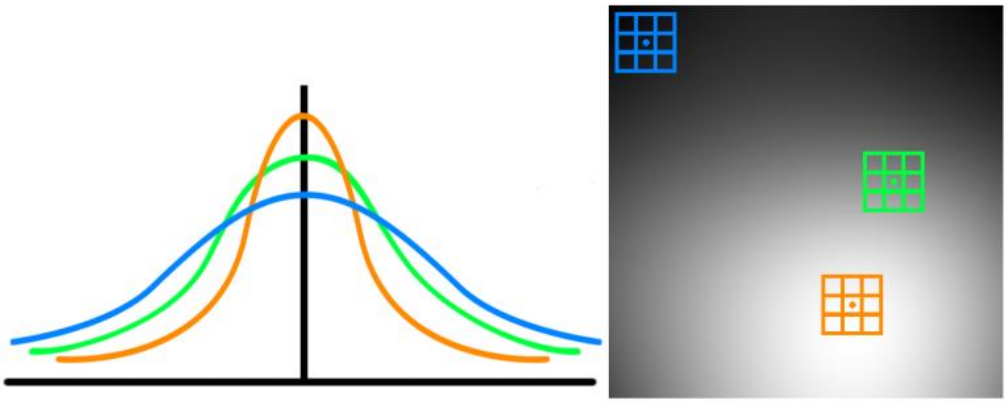

Figure 4.3: Three similarity windows aligned at various coordinates along the bias image (right). The mean pixel intensity contained in each window reflects the probability density (left) for parameterizing the adaptive kernel function.

That is to say that the signal is neither too weak nor too saturating and that the noise is at some intensity that would typically be expected. For the sake of brevity, the proposed theoretical framework will suppose that NLM is tuned to this specific region. Now, if we move to the orange similarity window, which spans the brightest region of the image and likewise correlates to the highest SNR region, we could correctly assume that this region requires less filtering than that of the green. If we were to visualize the probability densities of each region, we observe that has a sharper curve, which is directly analogous to the decay of the optimal kernel (Figure 4.3 (left), orange vs. green curve). Conversely, the blue similarity window reflects a region with low SNR and where heavier smoothing would likely be desired. The noise model in that region would be represented with widen and flatten the density curve (Figure 4.3 (left), blue vs. green curve). Similar to as before, the relative sharpness of the density function is directly analogous to the optimal decay of the kernel. Thus, the blue region will require a higher value of h than in the

green region to achieve optimal smoothing.

It has now been demonstrated how the bias image can be used as a map to guide smoothing; however, one last modification must be clarified. The bias image represents the compliment (or negative) of what is required to properly augment h. As previously mentioned, h effectively controls the decay of the exponential function of the kernel. Since it is the denominator of a negative exponential the function decay slows with large values, while hastening with small values. The exact opposite effect is achieved with the extracted bias image (e.g. Figure 4.2(a) and Figure 4.3(right)), thus, necessitating the inversion of the image. The compliment of the bias image is found by:

$$\hat{b}(i) = \arg\max_{B}\big(b(i)\big) - b(i) \quad \textbf{(5)}$$

where $\hat{b}(i)$ is the compliment of bias image and $\arg\max_B(\cdot)$ is the maximum gray value of the bit depth range of the bias image, B.

Maintaining the notation developed in Section 4.1.2 for the similarity window and center pixel, we define the adaptive filtering parameter as the mean of $\hat{b}$ within the corresponding similarity window about the current center pixel i:

$$\beta(i) = \frac{1}{m}\sum_{l=1}^{m} \hat{b}(\mathcal{N}_i(l)) \quad \textbf{(6)}$$

where $\beta(i)$ is the adaptive weighting parameter, l is the current pixel in the similarity window, and m is the size of the similarity window. Consequently, $\beta(i)$ modulates the filtering intensity by dictating the local SNR within the similarity window. In this view, $\beta(i)$ is an instantaneous scalar multiplier to h, which when combined with (3) is:

$$w(i,j) = \frac{1}{Z(i)} e^{-\frac{\|y(\mathcal{N}_i) - y(\mathcal{N}_j)\|_{2,a}^2}{\beta(i)h^2}} \quad \textbf{(7)}$$

and with (4) as:

$$Z(i) = \sum_j e^{-\frac{\|y(\mathcal{N}_i) - y(\mathcal{N}_j)\|_{2,a}^2}{\beta(i)h^2}} \quad \textbf{(8)}$$

As shown in equations (7) and (8), the weighting values modulate the strength of filtering. In our analysis of the two test images we optimized h for NLM and then initialized baNLM with the same value of h to ensure an unbiased evaluation of each filters' performance.

4.1.4 Exploring the Adaptive Kernel in Various Filters

The adaptive properties of baNLM are readily extended to other filtering techniques. The Gaussian filer (Appendix B), for example, is easily modified to exhibit the adaptive properties of baNLM. The Gaussian function (Equation (16) from Appendix B) is adapted by modulating the standard deviation term of the kernel with values from extracted in the bias image, $\beta(i)$ (Defined in Equation (6)):

$$G_{\beta}(i) = \frac{1}{\sqrt{2\pi}h} e^{-\frac{i^2}{\beta(i)2h^2}} \quad \textbf{(9)}$$

Similar to Equation (8), $\beta(i)$ in Equation (9) adaptively modulates the denominator of the exponential to sharpen or flatten the probability density curve of the noise model. The adaptive form of the Gaussian filter is notated here as *baG*.

Similar to the Gaussian function, the adaptive characteristics of baNLM can be trivially extended to a different class of filters based on partial differential equations (PDEs). One of the more common PDE based smoothing algorithms is centered around anisotropic diffusion (AD) (Appendix C), also called Perona-Malik diffusion [76], which is. An adaptive form of the diffusion equation is developed by manipulating Equation (21) similarly to the derivation of baNLM and baG. In this case the bias term is utilized to modulate the time step coefficient, τ, at each κth iteration in the discretized form of the diffusion equation:

$$z(i)^{\kappa+1} = z(i)^{\kappa} + \tau\big(\beta(i)\big)\big(A(i)^{\kappa} * z(i)^{\kappa}\big). \quad \textbf{(10)}$$

Thus, the use of $B(i)$ is analogous to a weighting function that determines the magnitude of diffusion that is applied a specific regions. This is illustrated in the bias gradient image where low pixel values (darker regions) induce less diffusion, while high pixel values (lighter regions) induce greater diffusion.

Table 1: Example of four frequently utilized filtering kernels and a version that includes the bias adaptive modification proposed in this dissertation.

Filer	Kernel	Bias Adaptive Kernel
Non-Local Means	$K(v_i, v_j, y_i, y_j) = e^{\frac{-\|y_i - y_j\|^2}{h_y^2}}$	$K(v_i, v_j, y_i, y_j) = e^{\frac{-\|y_i - y_j\|^2}{\beta(i)h_y^2}}$
Gaussian	$K(v_i, v_j, y_i, y_j) = e^{\frac{-\|v_i - v_j\|^2}{h_v^2}}$	$K(v_i, v_j, y_i, y_j) = e^{\frac{-\|v_i - v_j\|^2}{\beta(i)h_v^2}}$
Bilateral	$K(v_i, v_j, y_i, y_j) = e^{\frac{-\|v_i - v_j\|^2}{h_v^2} + \frac{-\|y_i - y_j\|^2}{h_y^2}}$	$K(v_i, v_j, y_i, y_j) = e^{\frac{-\|v_i - v_j\|^2}{\beta(i)h_v^2} + \frac{-\|y_i - y_j\|^2}{\beta(i)h_y^2}}$
Anisotropic Diffusion	$K(v_i, v_j, y_i, y_j)^{\kappa+1} = \tau\big(A(i)^k\big)$	$K(v_i, v_j, y_i, y_j)^{\kappa+1} = \tau\big(\beta(i)\big)(A(i)^{\kappa})$

As shown in Table 1, the bias adaptive modification is relatively straightforward in all four of the kernels. The modification ultimately centered on either adaptively tuning the decay of the exponential function (baNLM, baG, baBilateral) or tuning the "speed" of the time step coefficient of the PDE (baAD).

4.2 Testing of the Adaptive Filters

The efficacy of baNLM was validated by a variety of image processing experiments. However, before exploring how the filter performed with MR images, it was first tested in a more controlled environment that allows for rigorous evaluation of performance. Two images were employed to test the algorithm: (1) a synthetic phantom generated from a modified two-toned apollonian gasket (Figure 4.4(a–f)); (2) the familiar image of "Lena" (Figure 4.4(g–l)), which is popular and well recognized within the medical image processing community.

The literature surrounding medical image processing employs multiple phantoms for testing filter efficacy. None of these, however, seemed suitable for testing for scenarios when confronted with both intensity inhomogeneity and electronic noise. What was desirable was an image that exhibited the bi-tonal nature of muscle and adipose pixels observed in T_2-weighted MR images. That is to say, an image was required with numerous smaller boundaries that could be randomly marked as either adipose or muscle, and then configured in such a way that these boundaries would both span significant portions of the bias field, but also be tested in more local regions of the bias. For example, an image a single "muscle" boundary that spans a large portion of bias field (both high and low SNR), but also smaller "muscle" boundaries that are seated entirely in a high or low SNR local sub-region.

The Apollonian gasket provides a practical framework for establishing such an image that exhibits these properties. The gasket is a fractal image based on a sequence of repeating circles that are tangent to each other. Construction of the gasket begins with three mutually tangent

circles; Apollonius of Perga discovered that there exists at least two additional, non-intersecting circles that have the characteristic of being mutually tangent to the original three circles. Fractal-like properties of the circles ensue by the addition of smaller and smaller inner circles that are tangent to at least two outer circles. The inner and successively smaller circles are an analogous representation of fatty infiltrations and sparsely located adipose tissue dispersed within muscle. The phantom was generated by randomly seeding each circle with an 8-bit grayscale value of either 50 or 150. These values were intended to simulate values of either muscle (50) or fat (150) commonly observed in lumbar T_2-weighted MR images. Throughout the remainder of this manuscript, any reference to the image generated by the Apollonian gasket will simply be referred to as the *synthetic phantom*.

In addition to the synthetic phantom, the popular image of Lena was also used for testing (Figure 4.4(g)). Lena provides a more complex image than that of a two-toned phantom. It is often utilized in the literature and is likewise recognized for possessing some of the more desirable qualities of a photograph such as variable textures, wide range of shades, and contrasted features. Throughout the duration of this manuscript references to this image are termed “Lena”.

4.2.1 Images and Simulated Corruption

A 512 x 512 image of a quaternary form of the synthetic phantom was generated for testing (Figure 4.4 (a, d)). A gasket of this form has four equivalent circles encompassed by the primary outer circle (five total), with subsequently smaller circles filling the inner spacing between the original five (Figure 4.4 (a, d)). In addition to the gasket, an image of Lena was obtained which

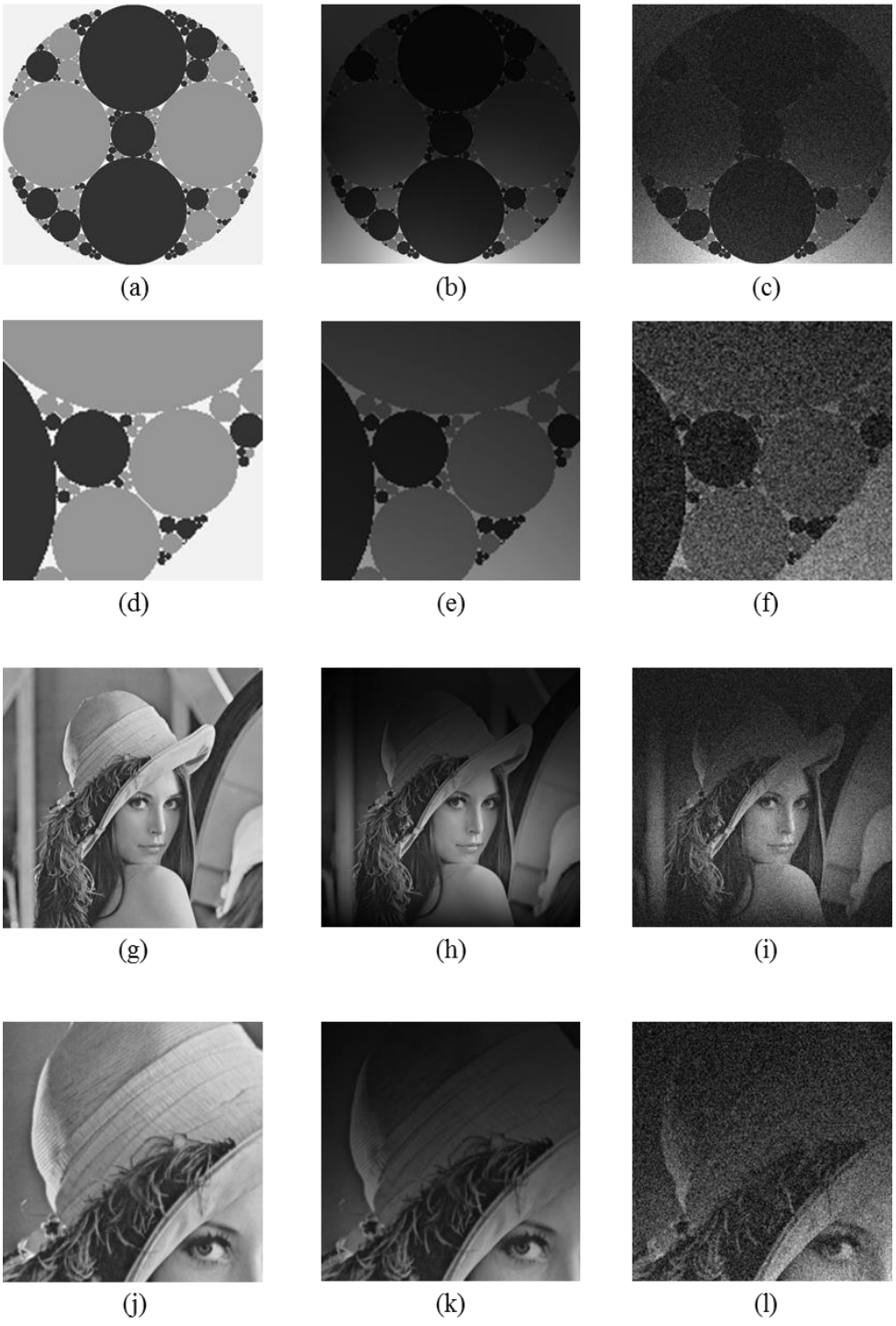

Figure 4.4: The phantom and Lena (a, g) corrupted with a realistic inhomogeneity bias field (b, h) and 9% Rician noise (c, i). A zoomed in section of the original images (d, j), with the added field (e, k), and 9% Rician noise (f, l).

likewise had a resolution of 512 x 512 pixels (Figure 4.4 (g, j)). Corruption of the phantom and Lena were carried out in the exact same manner and are detailed below.

4.2.2 Images and Simulated Corruption

A 512 x 512 image of a quaternary form of the synthetic phantom was generated for testing (Figure 4.4 (a, d)). A gasket of this form has four equivalent circles encompassed by the primary outer circle (five total), with subsequently smaller circles filling the inner spacing between the original five (Figure 4.4 (a, d)). In addition to the gasket, an image of Lena was obtained which likewise had a resolution of 512 x 512 pixels (Figure 4.4 (g, j)). Corruption of the phantom and Lena were carried out in the exact same manner and are detailed below.

A realistic bias field was first extracted from a lumbar MR image suffering from the characteristic intensity inhomogeneity. It was then added to the test images by element-wise multiplication as is suggested by Equation (1) (Figure 4.4 (b, e, h, k)). Rician noise was then added to the bias-corrupted images (Figure 4.4 (c, f, i, l)) at five different noise levels: 1, 3, 5, 7, and 9%. Results were averaged over 10 iterations at each of the five Rician levels. The same two phantom images were used throughout testing.

4.2.3 Notation Convention

Two different configurations were tested for both NLM and baNLM, four configurations total. The distinction is observed by if filtering is performed before or after correcting for intensity inhomogeneity. They are denoted using the following convention: NLM as Π^{η}, bias-adaptive

NLM as Π^{β}, and LEMS as Π^{λ}. The combined orders of operation are thus tabulated in Table 2.

Table 2: Notation for orders of operation.

Orders of Operation	Notation
Bias-adaptive NLM→LEMS	$\Pi^{\beta\lambda}$
NLM→LEMS	$\Pi^{\eta\lambda}$
LEMS→Bias-adaptive NLM	$\Pi^{\lambda\beta}$
LEMS→ NLM	$\Pi^{\lambda\eta}$

Two algorithmic paths for implementing these order of operations for baNLM is visualized in Figure 4.5. In summary, the red (top) path performs IIC once and then performs filter; while the green (bottom) path performs IIC twice, once to extract the bias image for baNLM and then again a second time after filtering for intensity correction (Figure 5).

4.2.4 Performance Measures

Evaluation of the baNLM algorithm was performed by four quantitative performance measures. We first used the Root Mean Square Error ($RMSE$), which measures the difference between pixel values of a reference image and the denoised (corrected) image:

$$RMSE = \left(\frac{1}{N_R} \sum_{i=1}^{N_R} \left(x_{reference}(i) - x_{corrected}(i) \right)^2 \right)^{1/2} \quad \textbf{(11)}$$

where $x(i)$ is the ith pixel of the reference or corrected image and N_R is the number of pixels in the reference or corrected image. Lower values of $RMSE$ indicate better performance.

The second measure was Pratt's Figure of Merit (FOM) [77]. FOM is a quantitative measure of edge preservation and enhancement, and is calculated from the Canny edge map of the denoised and referenced images:

$$FOM = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{1+{d_i}^2\gamma}, \quad \textbf{(12)}$$

where N is $MAX\{N_{corrected}, N_{reference}\}$, $N_{corrected}$ is the number of detected edge pixels in the corrected image and $N_{reference}$ is the number of detected edge pixels in the reference image, d_i is the distance between the ith edge pixel of the corrected image's Canny map and the next closest edge pixel in the reference map, and γ is a scaling constant defined to be 1/9 as indicated by [77]. The index of $FOM \in [0,1]$ where higher values indicate better performance.

The third performance measure used was the structural similarity index ($SSIM$), which is a quantitative measure of the structural similarities between the reference and denoised images [78]:

$$SSIM = \frac{1}{N}\sum_{i=1}^{N}\frac{(2\tilde{u}_i u_i + c_1)(2\hat{\sigma}_i + c_2)}{(\tilde{u}_i{}^2 + u_i^2 + c_1)(\tilde{\sigma}_i{}^2 + \sigma_i^2 + c_2)} \quad \textbf{(13)}$$

The SSIM was computed over an 11 x 11 window centered around the i-th pixel; μ_i and σ_i are the local mean and standard deviation of the original image, respectively; $\tilde{u}_i$ and $\tilde{\sigma}_i$ are the local mean and standard deviation of the denoised image, respectively; $\hat{\sigma}_i$ is the local covariance

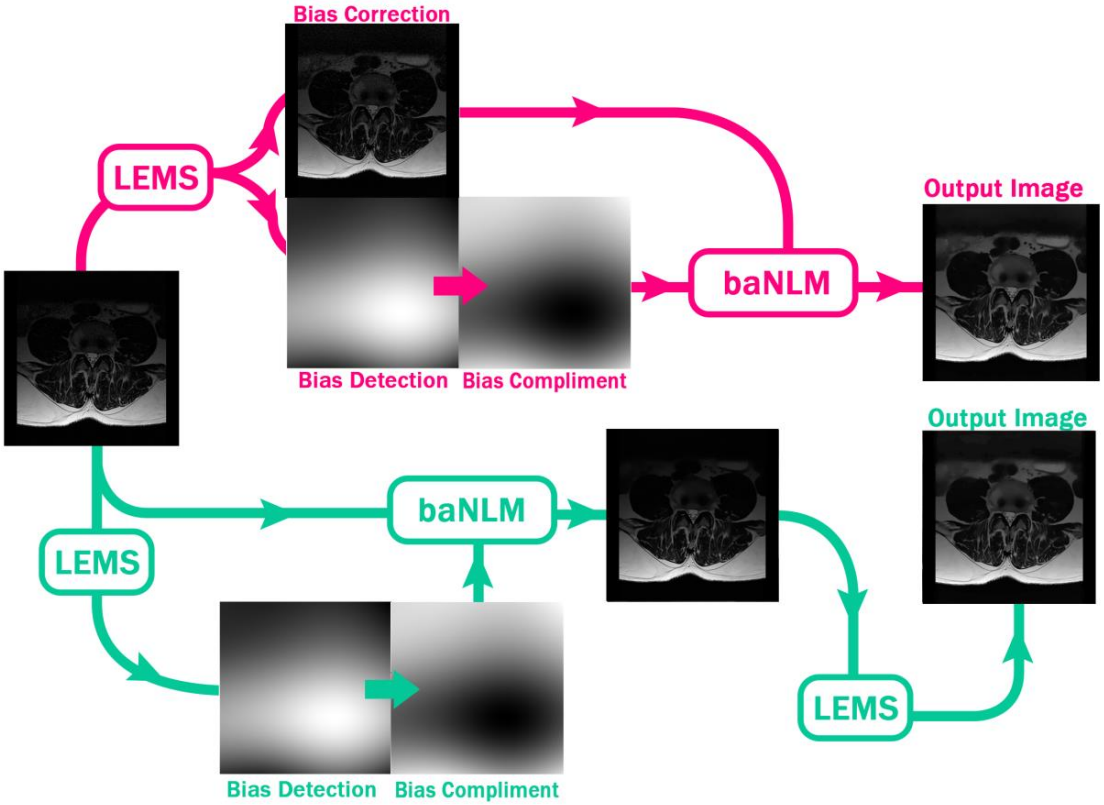

Figure 4.5: Two paths utilized for the analysis of baNLM an example MRI. The method for $\Pi^{\beta\lambda}$ (top, red) employs LEMS prior to filtering, while the method for $\Pi^{\lambda\beta}$ (bottom, green) employs LEMS before and after filtering.

between the original and denoised image; and the two constants were chosen to be $c_1 = 0.01$ and $c_2 = 0.03$ as recommended in [78]. Values of $SSIM \in [-1,1]$, where larger values indicate better structural similarity.

The SSIM was computed over an 11 x 11 window centered around the i-th pixel; μ_i and σ_i are the local mean and standard deviation of the original image, respectively; $\widetilde{u_i}$ and $\widetilde{\sigma_i}$ are the local mean and standard deviation of the denoised image, respectively; $\widehat{\sigma_i}$ is the local covariance between the original and denoised image; and the two constants were chosen to be $c_1 = 0.01$ and $c_2 = 0.03$ as recommended in [78]. Values of $SSIM \in [-1,1]$, where larger values indicate better structural similarity.

4.3 Optimization and Parameterization of h

Preliminary testing of the filter was required to optimize the parameters of the filter for the image sets being utilized. As stated in Section 4.2, the image sets used for testing and validation were a fractal image generated from an Apollonian gasket and Lena. However, before proceeding with testing the performance of the filter, proper parameterization of the filter must be established. As shown in equations (3), (4), (7), and (8), both NLM and baNLM (and in general any kernel based filter) utilize the tuning parameter h to dictate the strength of smoothing, which is ideally a function of the noise level present in the image. Since several noise levels were artificially introduced to the test images, individual values of h must be determined that optimize the targeted restoration metrics (i.e. the optimization of FOM, SSIM, and RSME which were presented in Section 4.2.4).

Sections 4.1.2 and 4.1.3 described the function of h in the NLM framework. Ideally h is parametrized to the relative level noise present in the image. In the context of baNLM h should be optimized and then local changes in B can be used to guide the smoothing.

The three metrics described in Section 4.2.4 all provide different methods for evaluating how well an image has been restored relative to the uncorrupted version of that image. Multiple performance measures were used because no one metric provides a definitive evaluation of the restorative process. Rather, the aim here was to utilize several different metrics in a combined fashion to generate a more encompassing evaluation of the filter. Since values of RMSE frequently fell within the range of [0.1, 0.3], it can be conveniently inverted such that $iRMSE = 1 - RMSE$, where the new values fall in the range of [0.7, 0.9]. This simple conversion allows for a "higher is better" interpretation like that of FOM and $SSIM$. Using this approach the three metrics were averaged together over ten iterations, denoted the mean metric value (MMV). The optimal value of h were determined by brute force maximization of MMV at each noise level (Figure 4.6 (a, c)).

They reflect values of h that returned the best performances for $\Pi^{\eta\lambda}$ and $\Pi^{\lambda\eta}$. They were likewise used to parameterize the two adaptive filters to ensure equivalent testing.

4.3.1 Result of test images

Filtering results of the synthetic phantom showed that baNLM generally outperforms NLM, particularly at lower levels of noise (Table 3). It was observed that $\Pi^{\beta\lambda}$ achieved the better results than the three other filters for all FOM and RMSE measures (with the exception of 9% RMSE) and SSIM at 1 and 3% noise; while the results for SSIM at 5, 7, and 9% noise were better for $\Pi^{\lambda\beta}$. 9% RMSE was the only test where a non-bias adaptive filter ($\Pi^{\eta\lambda}$) did not achieve the best score.

4.3.2 Results with Various Kernels

Outlined in Section 4.1.4 were the kernels for two other popular filters, namely: the Gaussian (G) and anisotropic diffusion (AD) filters. The filters were tested against the modified bias adaptive version of each in addition to NLM/bgNLM.

The performance for each filter was measured and averaged over ten different instances of the noise model. To ensure equivalent comparisons were being performed the same noise corrupted image was processed by all six filters at each iteration. In regards to the order of operation, the $\Pi^{\lambda\beta}$ configuration was used in all of the adaptive filters over $\Pi^{\beta\lambda}$ since the later post-process the image with LEMS. This was likewise performed to avoid any equivocation regarding the results.

Parameterization of h for AD and G filtering were optimized similarly to NLM/baNLM (Section 4.3). The values were determined by brute force optimization of the MMV metric to one

significant digit of h. As was expected, the Gaussian filters required values of h approximately two to three orders of magnitude greater than was needed for AD and NLM.

Lastly, results that are captured at separate runtimes are not necessarily directly comparable due to variations in the random number generator. For example, the results for the same filter in Table 3 and Table 4 may reveal differing performances. Testing showed that these variable results arise from a number of different factors, including such things as variable noise models (inherent expectation), bit depth, and datatype can all have significant effects on the performance measures.

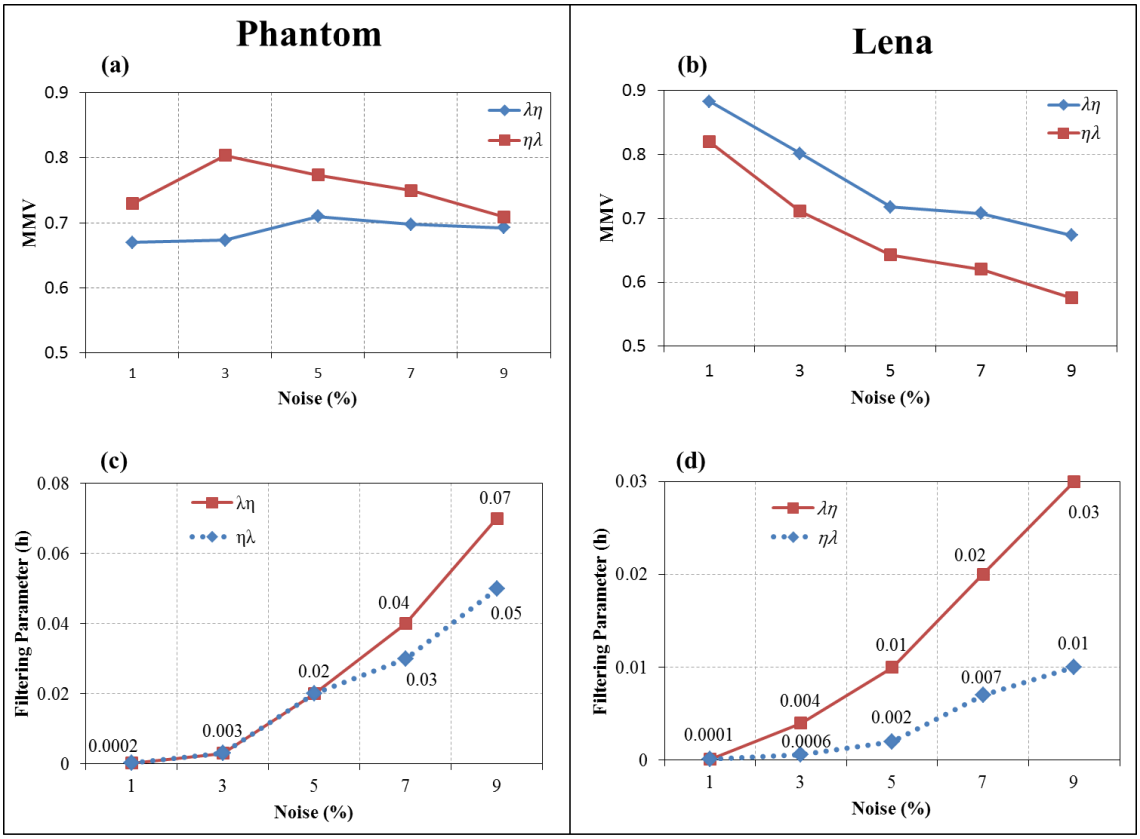

Figure 4.6: The mean values over 10 iterations of the optimized image performance measures, FOM, SSIM, and *i*RMSE, at the various indicated noise levels for $\Pi^{\lambda\eta}$ and $\Pi^{\eta\lambda}$ for the synthetic phantom and Lena (a and c, respectively). The filer strength parameter, *h*, that achieved the highest mean metric measurements over 10 iterations provided an indication of the optimal value at the given noise levels for the synthetic phantom and Lena (a and c, respectively).

Table 3: Average performance of the four filters on the synthetic phantom and Lena with various levels of noise. Statistically significant (p<0.01) performance values for a matched pairs t-test are indicated with the appropriate symbol (footnote). Bolded values indicate the filter that obtained the best average performance and that a test of significance was performed (see footnotes).

Noise	Metric	Synthetic Phantom				Lena				h
		$\Pi^{\lambda\eta}$	$\Pi^{\eta\lambda}$	$\Pi^{\lambda\beta}$	$\Pi^{\beta\lambda}$	$\Pi^{\lambda\eta}$	$\Pi^{\eta\lambda}$	$\Pi^{\lambda\beta}$	$\Pi^{\beta\lambda}$	
1%	SSIM	0.7877	0.7828	0.7888	**0.7892**†	0.8211	0.7555	**0.8248***†	0.7523	
	FOM	0.4988	0.6881	0.5554	**0.7313***†	**0.9295**	0.8625	0.9289	0.8605	0.0002
	RMSE	0.2767	0.2821	0.2769	**0.2625***†	0.1024	0.1581	**0.1023***†	0.1581	
3%	SSIM	0.8094	0.8323	0.8202	**0.8424***†	0.6877	0.6167	**0.6937***†	0.6168	
	FOM	0.4534	0.8062	0.5334	**0.8199***†	**0.8232***†	0.6731	0.8174	0.6827	0.003
	RMSE	0.2427	0.2280	0.2378	**0.2168***†	0.1053	0.1565	**0.1047***†	0.1583	
5%	SSIM	0.8618	0.8720	**0.8761***†	0.8641	0.6009	0.4995	**0.6075***†	0.5023	
	FOM	0.4500	0.6013	0.5347	**0.6354***†	0.6596	0.5622	**0.6630***†	0.5691	0.02
	RMSE	0.1827	0.1531	0.1616	**0.1469***†	0.1063	0.1339	**0.1061***†	0.1368	
7%	SSIM	0.8651	0.8681	**0.8754***†	0.8596	0.5414	0.4521	**0.5456***†	0.4547	
	FOM	0.3739	0.5060	0.4440	**0.5620***†	**0.6916***†	0.5407	0.6848	0.5441	0.03
	RMSE	0.1468	0.1243	0.1270	**0.1235**	0.1100	0.1318	**0.1096***†	0.1353	
9%	SSIM	0.8403	0.8410	**0.8434***†	0.8389	0.4811	0.3952	**0.4840***†	0.3985	
	FOM	0.3764	0.4142	0.4132	**0.4486***†	0.6573	0.4599	**0.6821***†	0.4785	0.05
	RMSE	0.1388	**0.1276**	0.1374	0.1285	0.1177	0.1277	**0.1168***†	0.1307	

*Statistically different (p<0.01) than the next closest filter for the same noise level and performance measure.

†Statistically different (p<0.01) than the result of the complementing filter (e.g. the complement of $\Pi^{\beta\lambda}$ is $\Pi^{\eta\lambda}$, and the complement of $\Pi^{\lambda\eta}$ is $\Pi^{\lambda\beta}$).

Table 4: Average (N=10) performances of the non-local means (NLM), anisotropic diffusion (AD), and Gaussian (G) filtering, and the bias adaptive (ba) counterparts on the Apollonian test phantom. Statistically significant (p<0.01) performance values for a matched pairs t-test are indicated with the appropriate symbol (footnote). Bolded values indicate the filter that obtained the best average performance and that a test of significance was performed (see footnotes).

Noise	Metric	baNLM	NLM	h	baAD	AD	h	baG	G	h
1%	SSIM	**0.7194***†	0.7176		**0.7138***	0.7131		**0.6905***	0.6878	
	FOM	**0.7492***	0.6871	0.0002	**0.8783***†	0.8557	0.007	**0.6198***	0.5889	4
	RMSE	**0.3268***†	0.3269		**0.3278***	0.3278		**0.3303***	0.3306	
3%	SSIM	**0.6920**†*	0.6878		**0.6777***	0.6766		**0.6530***	0.6508	
	FOM	**0.6399****	0.5418	0.003	**0.7931***†	0.6781	0.03	0.4414	**0.4571**	7
	RMSE	**0.3374***†	0.3375		**0.3389***	0.3390		**0.3421***	0.3423	
5%	SSIM	**0.6945**†	0.6938		**0.6907***	0.6891		**0.6623***	0.6595	
	FOM	**0.7005***	0.6170	0.02	**0.7245***†	0.5611	0.04	0.4775	**0.5229**	8
	RMSE	0.3284	**0.3284**†		**0.3290***	0.3291		**0.3334***	0.3337	
7%	SSIM	**0.7352***†	0.7338		**0.7272***	0.7249		**0.6942***	0.6909	
	FOM	**0.5620***†	0.4672	0.03	**0.3822**	0.3758	0.05	**0.5583**	0.5580	10
	RMSE	**0.3008***†	0.3009		**0.3011***	0.3011		**0.3075***	0.3078	
9%	SSIM	**0.7049***†	0.7032		**0.6994***	0.6972		**0.6778***	0.6750	
	FOM	**0.4994***	0.4251	0.05	0.3637	**0.3718**	0.06	0.5123	**0.5587**†	10
	RMSE	**0.3140***	0.3142		**0.3133***†	0.3134		**0.3192***	0.3194	

*Statistically significant (p<0.01) performance measure than the opposing filter from which the kernel was derived.

†Statistically significant (p<0.01) performance measure than when independently compared to all other filter performances.

Chapter 5

MRI of the Lumbar Spine

As previously mentioned, the research presented in this manuscript aimed to investigate three groups: (1) healthy with no recent history of low back pain and no radiological evidence of degenerative pathology; (2) recent history of low back pain, but otherwise healthy with no radiological evidence of degenerative pathology; and (3) recent history of low back pain and presented with at least one of the following radiological findings: degenerative disc disease, spondylthesis, Schmorl's Nodes, disc bulging both with and without canal or foraminal stenosis, disc osteophyte complexes, decreased height and fluid signal in the intervertebral disc, or facet hypertrophy. Overall, the population was comprised of 11 female and 15 male subjects and no significant differences were observed between the study groups in age, weight, weight, and body mass index (BMI) (Table 5). The objective of this analysis was to search for clues in the adipose and musculature data that would indicate some pathogenic or etiologic source for LBP and/or degenerative spine disease. The following discussion outlines the processes for acquisition,

image processing, and post-analyses that were perform make inferences regarding the pathological nature of LBP and degenerative spine disease.

5.1 Acquisition and Processing

The lumbar paraspinal muscles of 26 human subjects were evaluated for this study (Table 5). The participants were volunteers recruited from the staff, student, and patient populations at Vanderbilt University Medical Center (VUMC) and the University of Tennessee (UT), Knoxville. Institutional Review Board (IRB) approval was obtained as well as informed consent for all subjects participating in this study (UT IRB #7393; VUMC IRB #070883). No statistical differences were observed in demographic data of the study population groups (Table 5).

Table 5: Population demographics and statistics

Group	Total (N)	Female (N)	Male (N)	Population Means			
				Age (yrs)	Height (cm)	Weight (kg)	BMI
Healthy	10	5	5	42.0	175.6	86.8	27.9
LBP	7	3	4	46.0	173.3	80.8	26.8
Degen.	9	3	6	39.8	176.7	84.0	27.1
Intrapopulation ANOVA				p-value ($\alpha = 0.05$)			
Healthy vs. LBP vs. Degen				0.4600	0.7553	0.8868	0.9878
Healthy vs. LBP				0.2872	0.6198	0.7114	0.9433
Healthy vs. Degen				0.9015	0.7886	0.9365	0.9275
LBP vs Degen				0.2733	0.4553	0.5753	0.8829
Healthy + LBP vs. Degen				0.6631	0.5751	0.7683	0.8824
Health vs. LBP + Degen				0.4824	0.9272	0.8765	0.9692

Axial T_2-weighted images of the target lumbar muscles were acquired on a 1.5 T MR[†] unit ([†]Achieva; Phillips Medical Systems, Amsterdam, Netherlands) with a SENSE-spine[†] phased array coil. The images were captured using a turbo spin echo scanning sequence with: 512 x 512 resolution, 4 s repetition time, 120 ms echo time, 4 mm slice thickness, 90° flip angle, and a 225 x 225 mm^2 field of view. A stack of 50 images were acquired for each subject spanning approximately T12 to S1.

Three bilateral paraspinal lumbar muscles were evaluated for testing: the psoas, erector spinae, and multifidi (Figure 5.1).

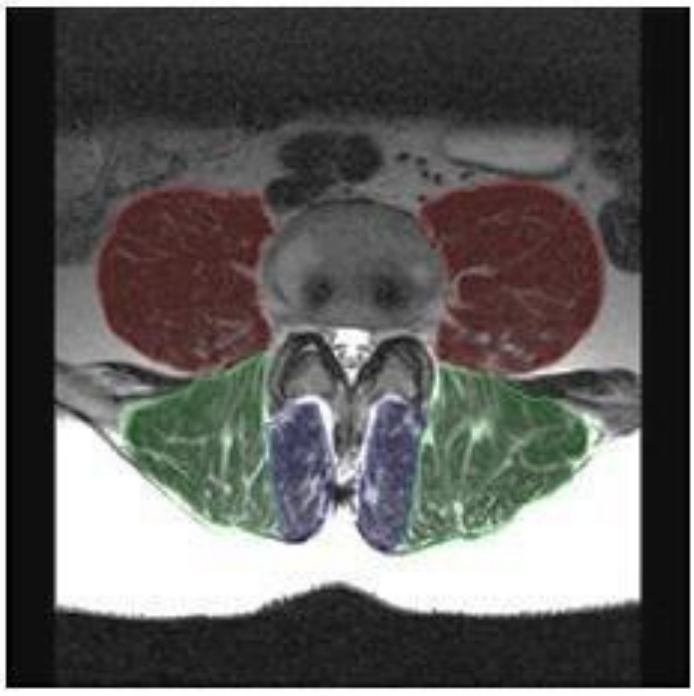

Figure 5.1: Sample MR slice at L4 with bulk muscle boundaries highlighted. The bilaterally occurring muscles are the psoas (red), erector spinae (green), and multifidi (blue). Note that the image has been brightened and the contrast increased to enhance the tissue boundaries.

The psoas muscles are positioned lateral to the vertebral body and anterior to the transverse processes; the erector spinae (superficial and deep) lie in the groove of the vertebral column and transverse processes; and the multifidi sit medially to the erector spinae and in the groove of the spinous and transverse processes (Figure 5.1). T_2 weighting displays pixel values of muscle as dark gray to black, fat/adipose as white, and bone as light gray (Figure 5.1).

5.1.1 Semi-Automatic Segmentation

Bulk muscle measurements were performed by manually segmenting the outer fascial boundaries of each of the six muscles of a single image at L1, L3, and L5. A gold standard was generated with manual segmentation, blindly performed by two different people. Inter-muscular adipose tissue was further segmented by fuzzy c-means (FCM) 2-cluster classification [79, 80].

Segmentation accuracy was evaluated against the manually segmented gold standard by way of Dice's coefficient. This is a comparative metric used for testing the similarity between two images [81], calculated by:

$$Dice = (x_{FCM} \cap x_{GS})/N \quad \textbf{(14)}$$

where x_{FCM} is the semi-automatically segmented image, x_{GS} is the manually segmented gold standard, and N is the number of pixels in one image. Calculations of $Dice \in [0,1]$, where higher values indicate more accurate segmentation.

5.1.2 Validation of baNLM on MRI Datasets

Spatially varying filter strength was achieved with real MR data (Figure 5.2). This was performed by beginning with an axial MRI of the lumbar (Figure 5.2(a)) and extracting the underlying intensity inhomogeneity bias (Figure 5.2(c)). As hypothesized, less smoothing was observed in the posterior muscle region in the erector spinae and multifidus, while more was in the anterior region (psoas). Figure 5.2(d) shows a zoomed region of the original MRI that has not undergone processing, while Figure 5.2(h) has been processed by baNLM, yet, little smoothing has occurred. In comparison, when a zoomed posterior region of the original MRI in region Figure 5.2(c) is juxtaposed against the same region filtered with baNLM in Figure 5.2(g) the difference of smoothing are quire stark. The controlled smoothing observed in Figure 5.2(g) and Figure 5.2(h) highlight the adaptive nature of baNLM.

The performance of baNLM was compared to that of NLM using real MR images of the lumbar spine. It would be difficult to perform a coherent comparative analysis between patients for an entire MR stack. As stated earlier, three images were selected to be processed. The processed and segmented images were compared against a set of gold standard images and analyzed via paired t-test. The reported results reflect the accuracy of the FCM tissue classification compared to the gold standard (e.g. Figure 5.3), quantized by the Dice index (Equation (14)).

Figure 5.3 illustrates two examples of the classification accuracy for muscle tissue. Relatively high accuracy by both filters (Figure 5.3(a)), though baNLM had approximately half as many mutually exclusive misclassifications as NLM (Figure 5.3(d)). Even greater disparity was

observed between the filters in Image 2, where baNLM had approximately seven times less misclassified pixels as NLM (Figure 5.3(d)). Additionally, the classifications of baNLM of both test Images 1 and 2 were more accurate than NLM (Figure 5.3 and Table 6). baNLM outperformed NLM in the erector spinae and multifidus muscles based on mean paired differences of the Dice coefficient (Table 7).

The left erector spinae at L5 was the exception, as no significant differences were observed for both filter comparisons. No differences were observed at the psoas muscle, except at L1. Using the total segmented area, the baNLM algorithm returned higher Dice index values than NLM at all three vertebral levels and both configurations.

Table 6: Classification accuracy percentages of test Images 1 and 2 (Figure 5.3). Reported are the percentages of pixels classified as: true positive (TP), true negative (TN), false positive (FP), false negative (FN), and percent accuracy

Image	Filter	TP	TN	FP	FN	Accuracy
1	baNLM	88.53	6.43	2.92	2.12	94.96
	NLM	89.55	4.39	4.96	1.09	93.94
2	baNLM	76.66	7.19	1.23	14.92	83.85
	NLM	70.12	7.53	0.89	21.47	77.64

Lastly, a residual image is a subtraction of the filtered image from original. It provides a qualitative means for investigating a filter's ability to perform noise reduction without losing fine features important to the image. The idea is that if an image were wholly filtered of 100% of noise, then the difference of the filtered and unfiltered image of this hypothetical perfect filter would yield a purely noisy image that contained no discernable features of the original image.

Table 7: Dice values for the mean differences over twenty-six subjects at each vertebral level and muscle group for baNLM against NLM

Mean difference	Level	Psoas		Erector		Multifidus		Full Seg.*
		Right	Left	Right	Left	Right	Left	
$\Pi^{\beta\lambda} - \Pi^{\eta\lambda}$	L1	0.0095	0.0112	0.0077	0.0091	0.0164	0.0148	0.0096
	L3	0.0011	0.0012	0.0044	0.0053	0.0126	0.0123	0.0051
	L5	0.0046	-0.0012	0.0139	0.0500	0.0080	0.0082	0.0059
$\Pi^{\lambda\beta} - \Pi^{\lambda\eta}$	L1	0.0240	0.0112	0.0141	0.0123	0.0242	0.0195	0.0145
	L3	0.0020	0.0004	0.0055	0.0060	0.0143	0.0141	0.0057
	L5	-0.0011	-0.0009	0.0240	0.0534	0.0085	0.0116	0.0063

*Full Seg. refers to the collective segmentation of all six muscle groups. Bolded values indicate a significant difference ($p < 0.05$) for a paired t-test.

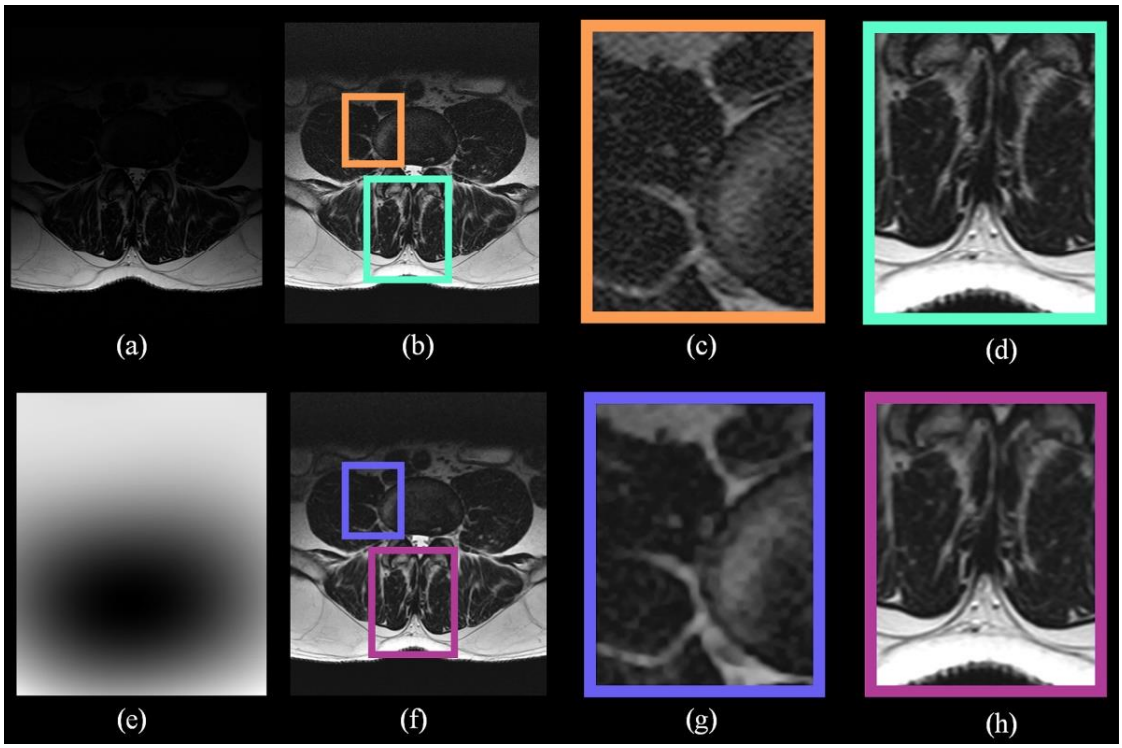

Figure 5.2: The original MR image (a) and compliment of the extracted bias field (e). Shown are the unbiased image (b) and the unbiased image with bias adaptive NLM (f). Zoomed regions of the unbiased image are shown in (c) and (d), while (g) and (h) are the same zoomed regions for the bias guided filtering. Greater filtering was applied at (g) compared to (h), while little filtering occurred at (d) compared to (h).

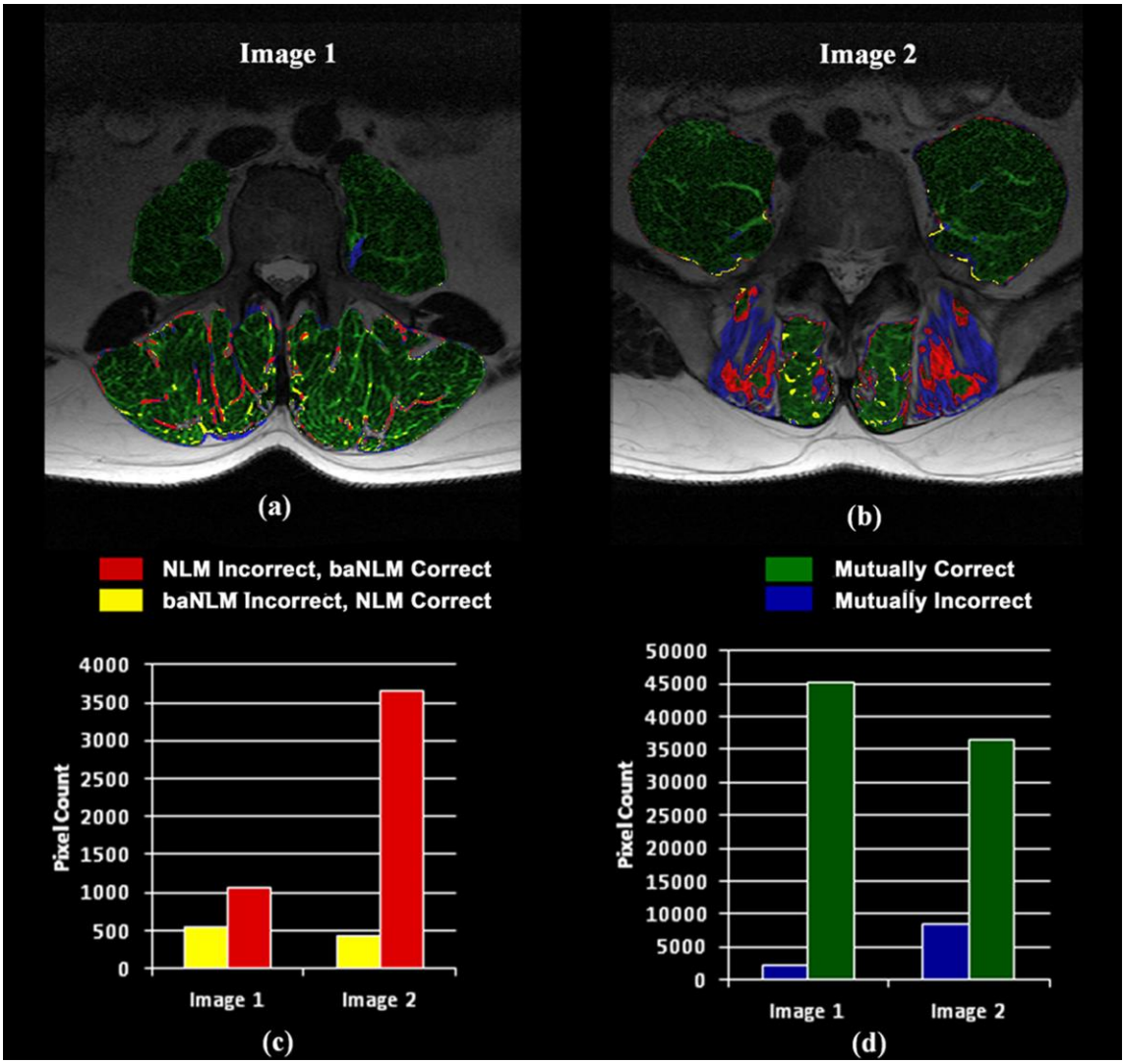

Figure 5.3: Automatic segmentation accuracy of muscle tissue for two different MR images (a, b). Each image was processed independently with baNLM and NLM, and then classified with fuzzy c-means clustering. The classification results of the images from both filters are overlayed on a single image and color coded based on correspondence of the classification results. Green depicts regions where baNLM and NLM were mutually correct, while blue indicates where they were mutually incorrect (b, d). Red indicates baNLM was correct and NLM was incorrect; while yellow indicates NLM was correct and baNLM was incorrect (a, c). Incorrectly classified pixels were identified as both false-positives and false-negatives, while correctly identified pixels were associated with only true-positive muscle tissue. The classifications from both images show baNLM (yellow) had significantly less mutually exclusive misclassifications than NLM (red) in both images (c).

An example MRI residual of baNLM (Figure 5.4(a)) shows better fine feature retention than the same image filtered with NLM (Figure 5.4(b)), particularly in the posterior region.

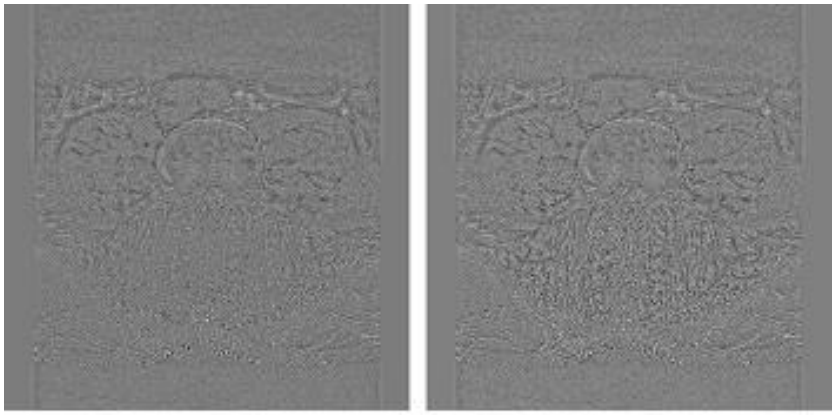

Figure 5.4: Dice values for the mean differences over twenty-six subjects at each vertebral level and muscle group for baNLM against NLM.

Performance was markedly improved in the posterior region at the muscle-fat interface. The anterior region, however, appeared to have similar qualitative properties. These results further confirm the findings from Table 7 that showed more accurate segmentation in the posterior region, while the segmentation in the anterior region was not significantly different.

5.2 Population Trends

The MRIs acquired for this study were typically comprised of 50 axial images incrementally stacked along the transverse plain of the lumbar spine. Each image in the stack was filtered with baNLM using the $\Pi^{\beta\lambda}$ configuration detailed in Table 2. The filtered images were then processed

by automatically segmenting the adipose tissue from the bulk muscle boundaries by FCM classification. Individual images revealed a segmented region for the muscular and adipose tissue that represents a snapshot of their cross-sectional area (CSA) in the axial plane. The CSAs are determined by summing the number of pixels in a specific region and multiplying by the area of an individual pixel. These data offer valuable insight into the structural makeup up of the lumbar spine; for example, numerous studies in the literature have drawn correlation between the biomechanical function of the paraspinal muscles with the corresponding CSA. The 2D segmentations can also be used in three dimensions by constructing surface volumes with the voxel data over the length of the scan (Figure 5.5 and Figure 5.9).

The individual bulk muscle volumes tend to follow a predictable pattern along the vertebral column. The volume of the psoas major gradually decreases as it ascends from the lower extremity along the lumbar vertebrae towards L1 and T1 where the muscle eventually tapers to tendinous connective fibers (Figure 5.5). The erector spinae is also consistent with expected anatomy as the muscle volume gradually increases in the superior direction from L5 towards L1 (Figure 5.5). The multifidi tended to decrease in volume as it extends superiorly towards L1, similar to the psoas major but with less drastic intervertebral changes (Figure 5.5).

From the perspective of biomechanics, it stands to reason that symmetric active forces would foster more favorable conditions in the lumbar spine. To this end, one might postulate that bilateral symmetry of the musculature may in fact promote healthier motion of the spine and, thus, we may observe higher muscle symmetry in healthy spines. However, in this study,

bilateral differences between similar muscles were largely negligible (Figure 5.5) with the exception of the multifidus and the erector spinae at L4 ($p \leq 0.05$ for a matched pair t-test). In these instances, the left multfidus was observed to contain less volume than on the right side, while less volume on the right erector spinae was observed than on the left side.

The argument for bilateral differences can be a bit spurious, especially without taking into consideration of the "handedness" (right of left) of the study participants, which was not tracked in this study. Nonetheless, the handedness of the subjects can be normalized by evaluating the absolute value of the differences between the right and left sides. After correcting for this bias no significant differences were observed in any muscle and at any level.

In an effort to better realize the role bilateral muscle volume plays in the pathology of the lumbar spine the percent adipose ($(adipose/(adipose + muscle)) * 100\%$) of each left/right muscle pair was plotted against the subject's BMI (Figure 5.6). These results revealed some interesting information (Figure 5.6): First, the trend lines of the plots show that the three muscle groups clearly cluster in different regions. That is, the multifidus has higher adipose content than the psoas, while the erector spinae falls somewhere in between. Second, the regression lines of both the right erector spine and right psoas have noticeably less slope than the left sides of their counterparts. Third, there appears to be no bilateral differences in the psoas muscles.

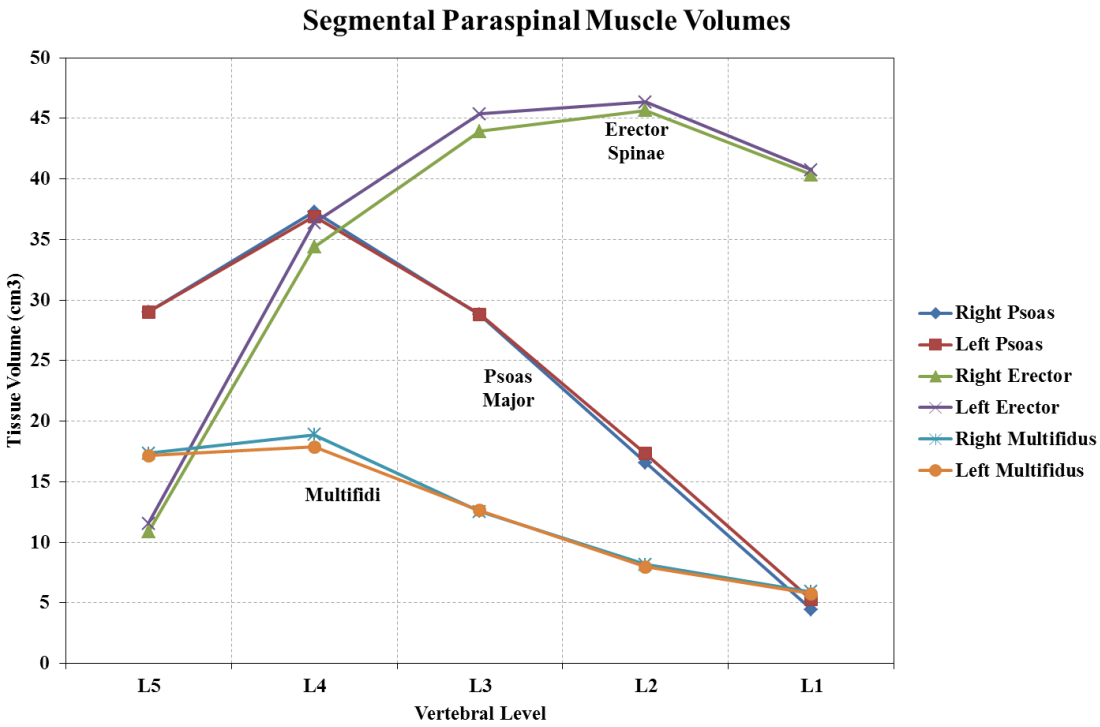

Figure 5.5: Population average of the paraspinal muscle volumes at specific lumbar vertebral levels.

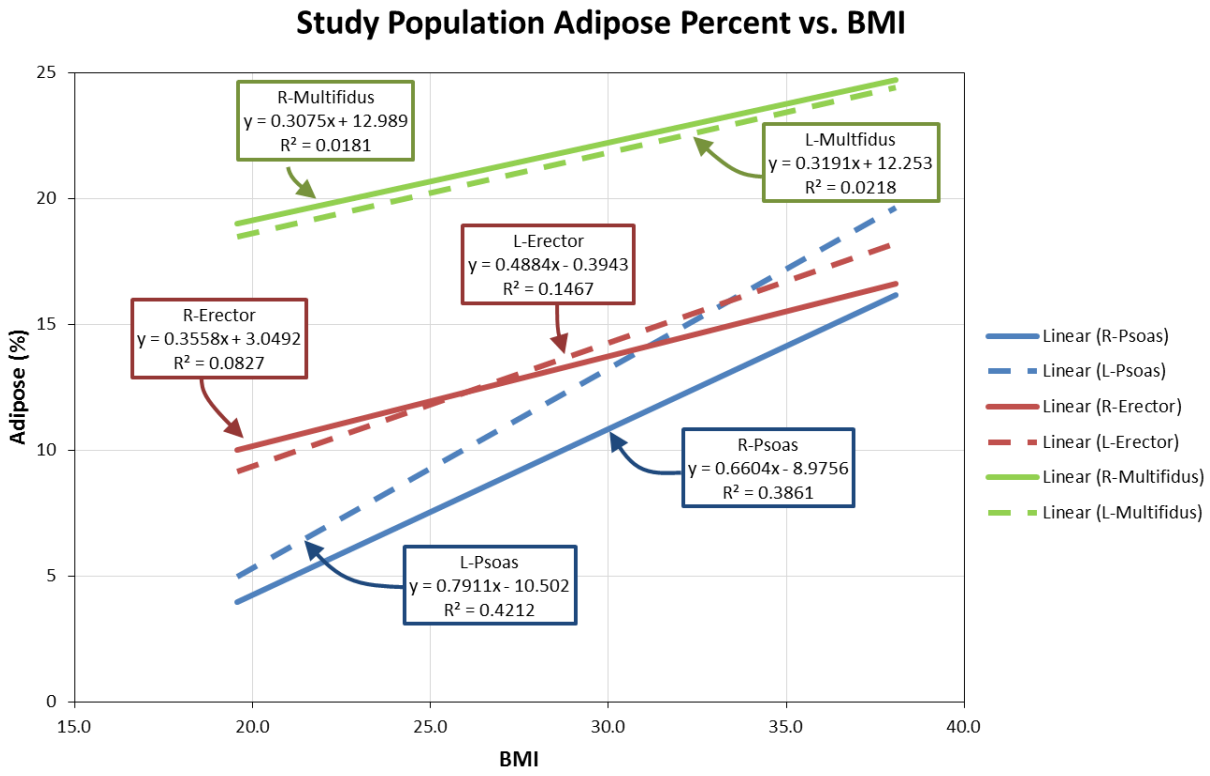

Figure 5.6: Gender specific relationships between percent adipose tissue and BMI. Percent adipose is a reflection of the content across the entire paraspinal column of the lumbar. Female subjects have positive correlation of adipose to BMI while males exhibited a negative correlation.

Upon investigation, the most noticeable feature driving observed differences in muscle volume is its correlation with gender. The muscle volume of the male population was unanimously and statistically ($p < 0.0001$) larger than the female population (Figure 5.7, left). No significant differences were observed for the raw adipose volume (Figure 5.7, right). The observed gender differences in muscle volume instantly confound any inferences that are to be made regarding pathology and, without normalization, would necessitate separating the data based on gender. Thus, the notion of sexually dimorphic biomechanics is explored in detail later (Section 5.3).

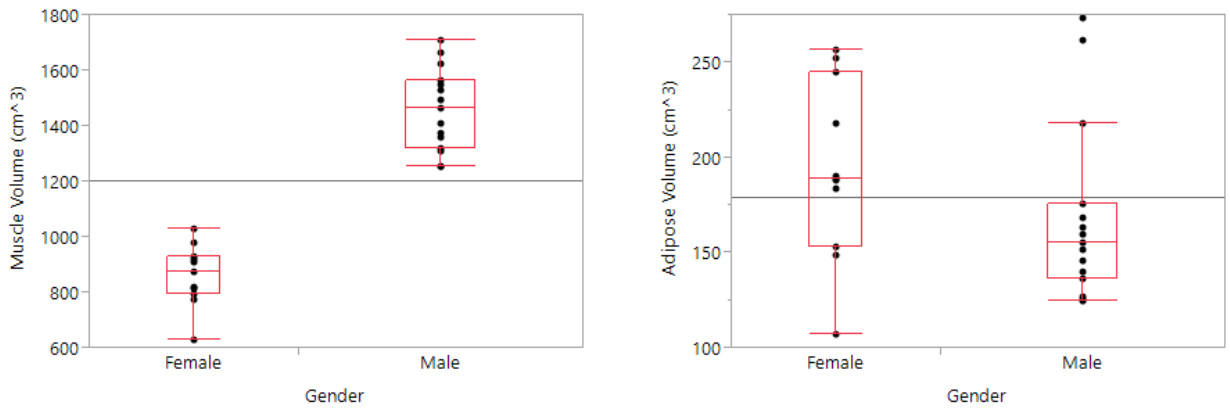

Figure 5.7: Observed gender differences in the muscle and adipose volume. Muscle volume was statistically less for females than males (left), while no significant difference was observed in the adipose volume (right).

The next most logical and ubiquitous demographic feature to investigate would be body mass index (BMI). Even though recent literature brings into question the efficacy of BMI as a metric of health, it is still a reasonable feature for analyzing biomechanic function from the perspective of active forces and muscle volume/CSA. A Kruskal-Wallis test of significance of the subject's

BMI separated by gender revealed no significant differences between each other (Figure 5.8). However, the vast difference in muscle between genders is pursuant to further investigation.

5.3 Gender Bias

The statistics of the population as a whole seem to indicate that muscle volume is the largest predictor of gender. Neither adipose volume nor BMI were shown to hold any statistical differences. Since the discrepancy of muscle volume between genders is so vast the discussion then naturally transitions to sexual dimorphisms that may be biomechanically relevant. That is, at least, before continuing with the discussion on pathology.

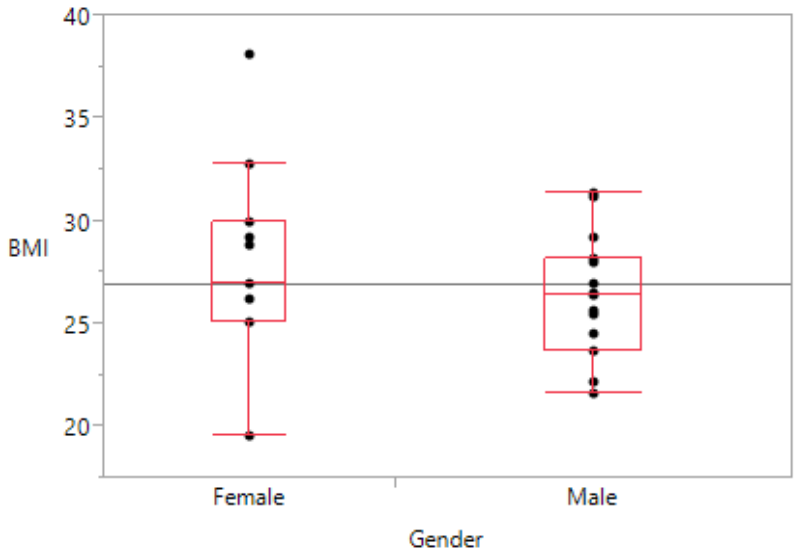

Figure 5.8: Boxplot of the BMI separated by gender. No significant differences were observed between genders.

The representations of the 3-dimensional surface models displayed in Figure 5.9 provide a convenient tool for qualitatively evaluating how the adipose tissue is dispersed throughout the muscular tissues. Inspection of the volume data reveals that a similar discrepancy begins to emerge with a significant bias towards gender (Figure 5.9). The most noticeable observation is the muscle girth between genders, where the male erector spinae and psoas muscles are clearly larger in the male subject compared to the female. In addition, female subjects routinely showed more adipose content than males; particularly in the posterior region of the lumbar where the erector spinae and multifidus muscle boundaries are observed. It is fairly well established that numerous sexually dimorphic characteristics exist between human females and males. In particular is the difference in physique between genders, where the male stature is generally larger than that of females. These traits were positively confirmed in this study (Table 8).

Table 8: Average volumetric (cm^3) data across the study population of three bilateral lumbar muscle groups shows a significantly different cross-gender comparison in all lean muscle groups for a Kruskal-Wallis test. In the adipose tissue, only the multifidus showed significant differences.

Muscle	Lean Vol.		Adipose Vol.	
	Female	Male	Female	Male
	cm^3 (± S.D.)			
Right Psoas	126.7 (11.0)*	261.8 (9.4)*	16.8 (2.0)†	19.0 (1.7)†
Left Psoas	127.3 (9.3)*	263.8 (8.0)*	21.2 (2.6)†	24.8 (2.2)†
Right Erector	207.3 (11.7)*	333.5 (10.0)*	41.2 (3.6)†	36.5 (3.1)†
Left Erector	210.8 (13.2)*	344.3 (11.3)*	43.1 (4.6)†	37.9 (3.9)†
Right Multifidus	95.5 (5.3)*	128.6 (4.6)*	37.0 (3.1)*	25.4 (2.7)*
Left Multifidus	93.4 (5.2)*	127.6 (4.4)*	34.8 (3.1)*	25.0 (2.6)*

* Significant at $\alpha \leq 0.05$ probability level.
† Not significantly different.

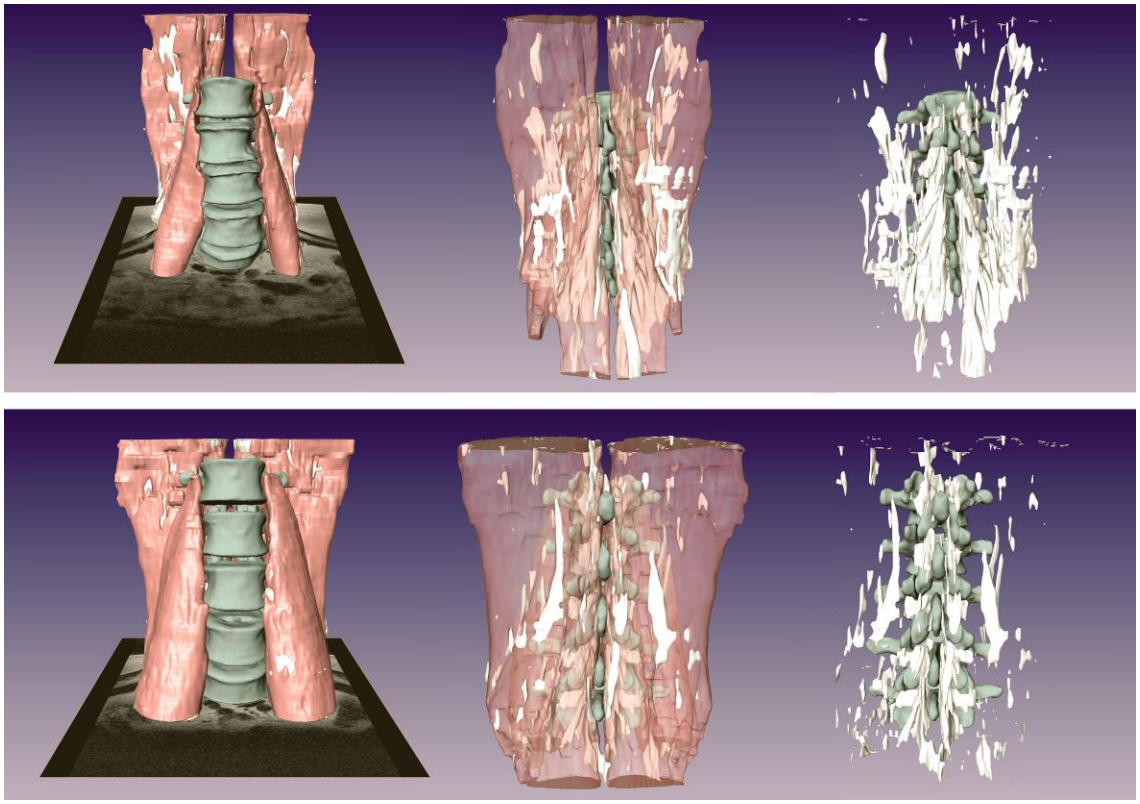

Figure 5.9: Three dimensional volumetric reconstruction of the lumbar spine and paraspinal musculature of a female (top) and male (bottom). The surface models show a significant gender differences in both the muscle girth and adipose tissue content.

The distribution between genders was of particular note as the female population had a wider spread of lean muscle percentages. However, perhaps the simplest explanation for this observation was that while male population mean hovered near 90%, the upper tail of the data cannot exceed 100%, so, there is a natural tightening of the distribution nearer the statistical limits.

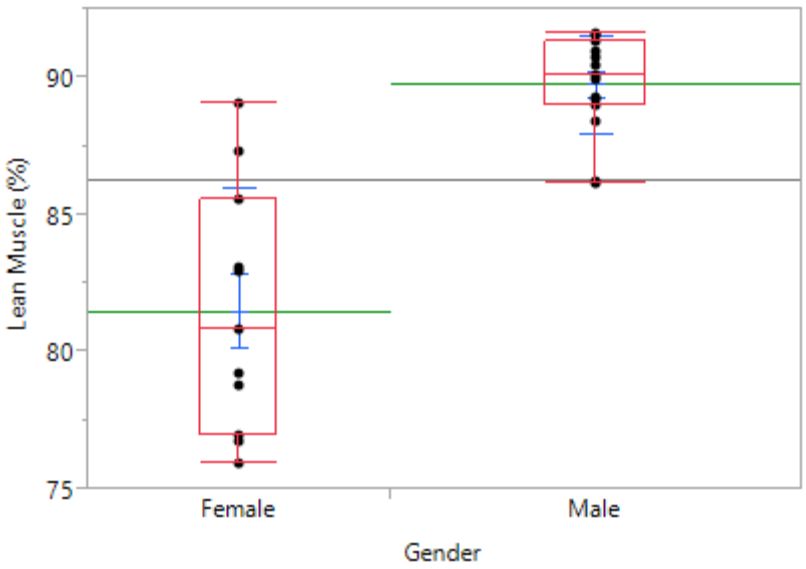

Figure 5.10: Box plots depicting gender specific percent lean muscles values. Male individuals clearly have more lean paraspinal muscles than females.

A similar gender-centric trend was observed when plotting percent adipose content versus BMI, which showed that while the genders shared fairly similar regression slopes (males: 0.3546; females: 0.2258), the y-intercepts (percent adipose) were significantly higher for females than males (Figure 5.11). Females saw an intercept of 12.19 and males an intercept of 0.8973. These

results further confirm the correlation between percent adipose (inversely proportional to percent lean muscle) and gender (Figure 5.11).

The analysis is further distilled by plotting individual muscles against the subject's BMI (Figure 5.12). Plots of all three muscle groups again showed that female musculature had increased percent adipose content as compared to males. However, what is perhaps of more interest is the between-gender trend that the individual muscles reveal.

The most striking of the trends in Figure 5.12 was observed in the multifidus. Here, the regression line showed a negative slope for females (-0.25) and positive slope for males (0.50). The implication is that female subjects with higher BMI tend to have lower percent adipose in the multifidus (Figure 5.12(b)). This is counterintuitive as the expectation would be for percent adipose to rise accordingly with BMI. These results are, however, complimentary to the volume totals outlined in Table 8, which showed a significant difference between in the adipose volume of the multifidus, but not the other two muscles. Thus, it is not unreasonable to conclude that the multifidus is functioning differently for females than for males.

An additional observation is noted in the differences between slopes in the psoas muscle (Figure 5.12(a)). The slope of the female psoas (0.86) is considerably steeper than that of the male (0.16), which is nearly flat. This data suggests that the percent adipose found in the male psoas remains unchanged regardless of BMI; whereas with the female population, there appears to be a stronger correlation of increased percent adipose and higher BMI.

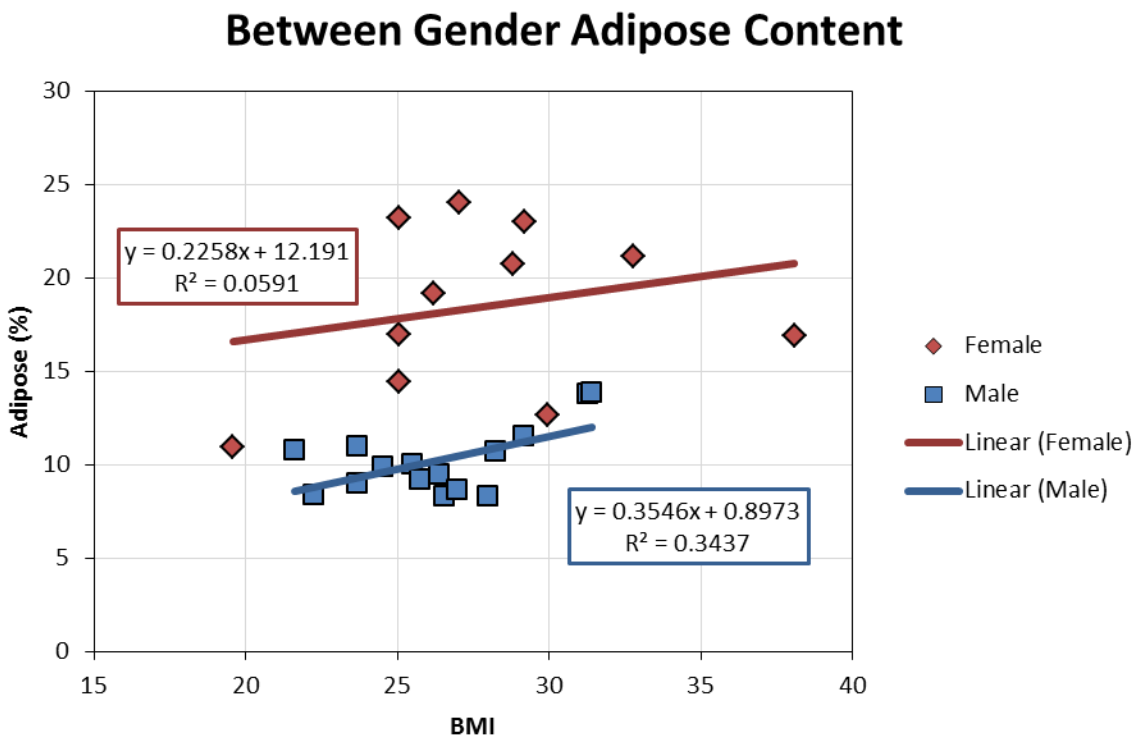

Figure 5.11: Between gender percentages of adipose tissue in the paraspinal muscles of L1–L5. Female percentages were significantly greater than that of males and showed a slightly less steep regression line.

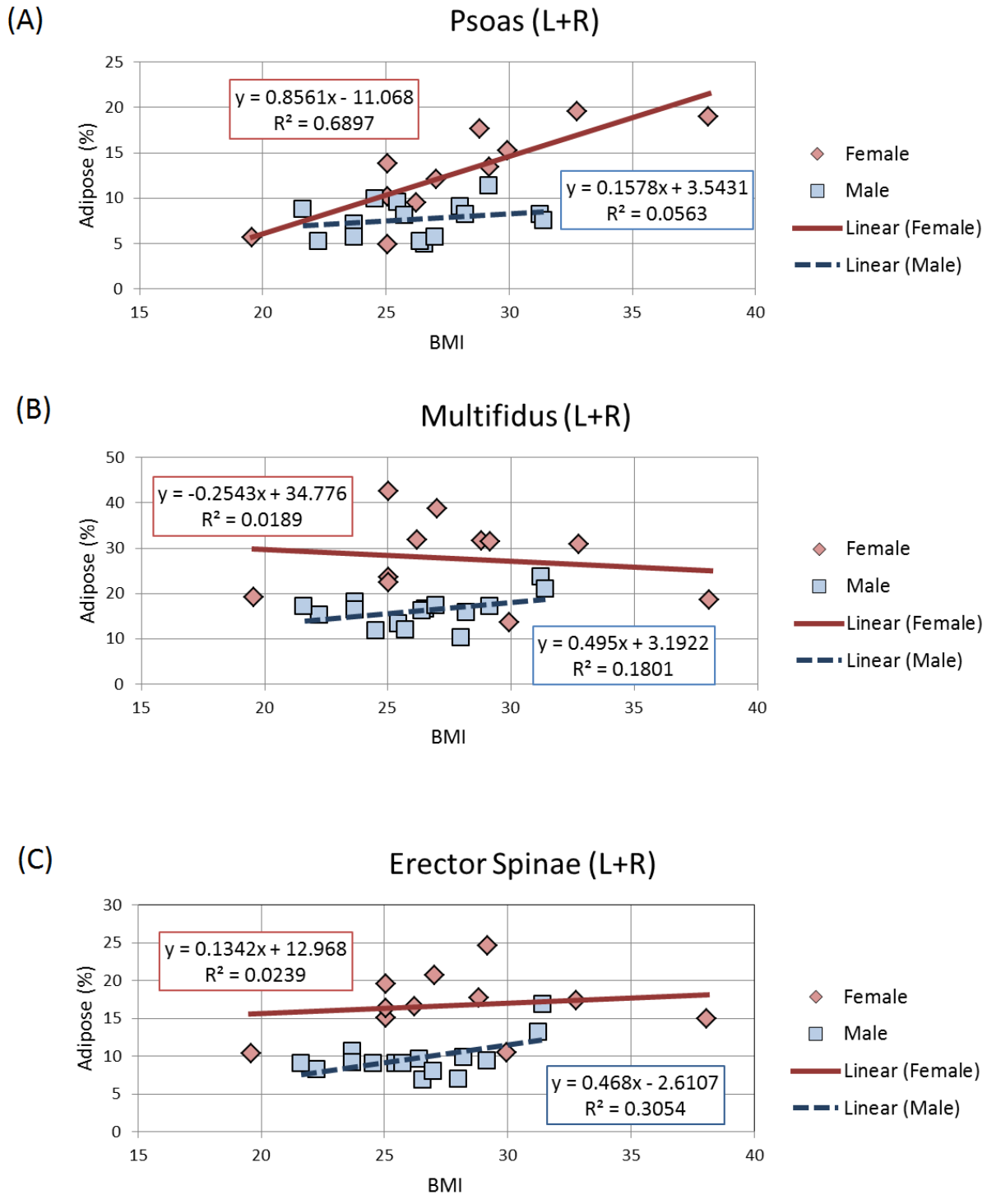

Figure 5.12: Combined bilateral (left and right) percent adipose tissue versus BMI with corresponding regression correlation lines for the psoas (a), multifidus (b), and erector spinae (c) muscles.

This observation is then swapped between genders for the erector spinae, where the slope for females (0.13) is more flat than males (0.47).

Gender differences between anatomy and musculature are certainly not a novel concept. Take for example the observed differences of the so called *q-angle* in the knee joint. Owing to a naturally wider pelvis, women on average exhibit a greater q-angle than males. As a result, the biomechanics of the knee are universally different between genders. This idea is solidified by the fact that orthopaedic companies have widely sought out to develop gender specific total knee replacements, because they exhibit far better fit and wear patterns than if they were to a single joint across both genders. The concept of this sexual dimorphism can be easily extended to the psoas muscle, which spans the pelvis. It is not out of the question to assume that the psoas functions differently between genders; or at the least that the biomechanic forces have differing lines of action; or that the psoas work in concert with other active muscles differently between genders. The exact functionality of the psoas exemplifies the symbiotic nature and compensatory characteristics that the biomechanics that muscles must exhibit in order to provide humans with stable posture and locomotion.

The results presented here demonstrate to a clear difference of musculature between genders. This is relevant to numerous ideas surrounding the lumbar spine. First, the unfortunate reality is that the difference in raw muscle volume necessarily splits the population in this study by gender. Second, it would be prudent to develop future dynamic and biomechanic models with gender specificity in mind and those models should explore possible gender differences in the

muscle activation energies and forces that promote posture and locomotion. Third, this study clearly established the existence of volumetric differences, but it may be even more fruitful to extend these ideas to investigate gender specific shape differences in girth and musculature, possibly with the aid of a type of flexible statistical shape modeling. Lastly, if the gender differences are indeed observed in the biomechanics of the lumbar, there could be considerable evidence available to infer that the pathogenic nature of degenerative spine diseases occurs along a path that is in fact gender specific; a radical notion that would shift the whole paradigm of how LBP and degeneration are understood and treated. There is indeed compelling reason to pursue a deeper understanding of how gender affects the biomechanics and pathology of the lumbar spine.

Chapter 6

Classification of Pathology

The ideas established in Chapter 5 identify a clear relationship between gender and the paraspinal musculature and its corresponding adipose content. While these gender specific findings are curious from the perspective of biomechanics, they do unfortunately muddle the statistical clout of the population. The volumetric differences between genders are so vast that the any analysis aimed to utilize the raw values essentially requires the data be separated by gender. This is in fact detrimental to some of the analyses. For example, the low back pain (LBP) group had only three female participants. – making most statistical inferences into the group largely meaningless. This dilemma prompted an exploration into various other sources of data and data manipulations that may elucidate the analyses.

Between group muscle volume differences were negligible and certainly not nearly as stark as was the observed dichotomy between genders (Figure 6.1).

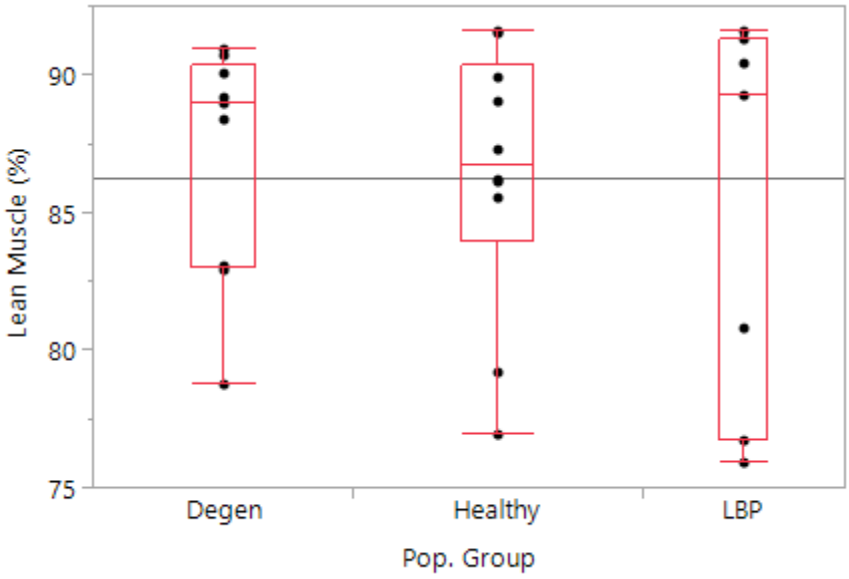

Figure 6.1: Distribution of lean muscle values across the three subject groups. Distributions and means are not significantly different from each other.

Some hypotheses in the literature posit that the mechanism of low-back pain is inherently progressive, and that eventual degeneration is a likely outcome. Johnson et al. [82] observed a correlation between poor *in vivo* kinematics (i.e. out-of-plane motion) and the pathological classification of the lumbar. They found subjects with increasingly severe diagnoses had increased incidences of out-of-plane motion during in-plane activities. The results seemed to indicate that the mechanics are progressive in nature, meaning that an individual, who is currently at the fusion stage, has likely already passed through the LBP and degenerative phases. This notion could be extended to the intervertebral discs, which act as a cushion between the vertebral bodies and promote fluid motion between the facet joints. Many researchers theorize that poor mechanical properties of the discs can cause a myriad of cascading effects resulting in spinal degeneration. A study was devised to investigate the compression ratio of discs from

positions of supine to standing between healthy lumbar spines verses those with degenerative abnormalities. The hypothesis was that degenerative spines would experience more translational movement than healthy spines, as the degenerative subjects' discs are theoretically more compromised. This was tested by evaluating relative *in vivo* distances at adjacent lumbar levels and the overall L1-to-L5 distance. The intervertebral spacing was measured for two test groups: 16 subjects with healthy lumbar spines (LBP group was joined with the healthy) and 9 subjects classified with a specified degenerative lumbar pathology. Each patient underwent fluoroscopic and computed tomography (CT) examination. CT imaging provided 3D bone models, which served as the supine frame of the patient (Figure 6.2). The extracted 3D surface models are also registered to fluoroscopy videos by overlaying and optimizing the orientation of the models.

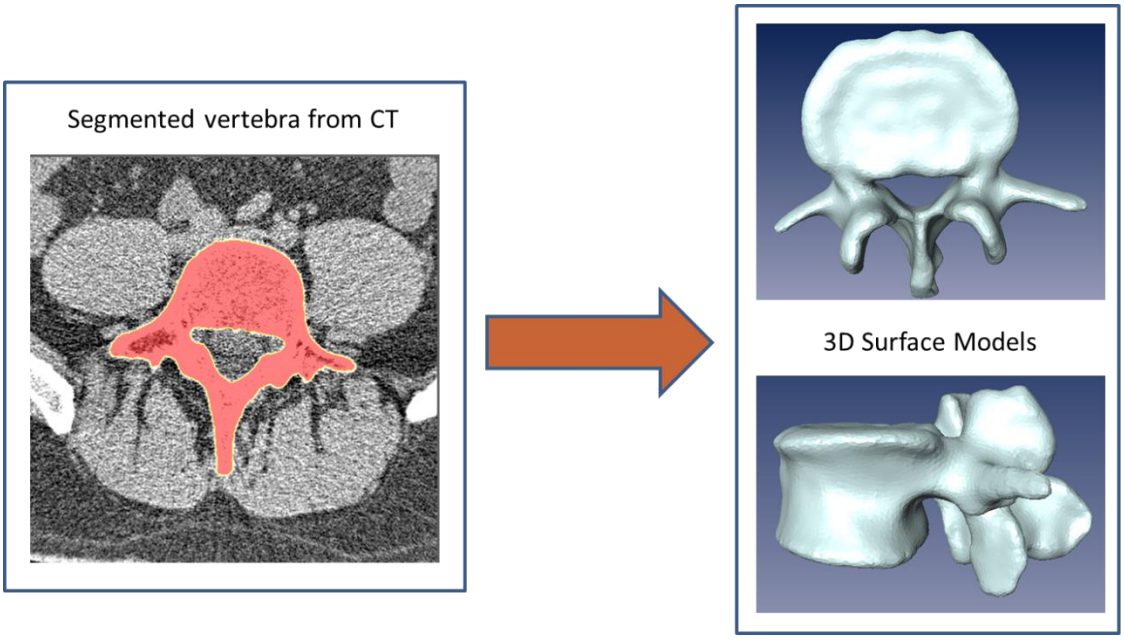

Figure 6.2: Boundaries of the bones are segmented in slices of a CT scan. Adjacent segmentations are joined together to generate 3D model of the vertebrae.

A C-Arm fluoroscopic unit captured vertebral images of the subject while in an erect, standing position (Figure 6.3). Measurements in the erect position were acquired by overlaying 3D bone models over the fluoroscope images using a 3D-to-2D registration technique [83] (Figure 6.3). At both supine and standing positions a 3D curve was generated by stitching together the centroids of L1 through L5 using a piecewise cubic Hermite spline (Figure 6.4).

Reported are the average supine-to-standing differences normalized by the measured L1–L5 length of the lumbar (Table 9). The results reflect a calculation of the functional spine unit compression ratio: $CR = \left(\left(dist._{supine} - dist._{standing} \right) / dist._{L1-L5} \right) * 100\%$, where $dist._{supine}$ is the intervertebral spline distance in the supine position, $dist._{standing}$ is the intervertebral spline distance in the standing position, and $dist._{L1-L5}$ is the spline distance in the supine position between L1 and L5. As formulized the CR allows for negative values – when standing distance is greater than the supine – which implies a lengthening rather than a compression. The extrema on the CR values were: [-2.54,6.27]. Negative values explain for the large standard deviations compared to the means (Table 10). Across the population a matched pair t-test of the CR ratio was performed with $\alpha<0.01$. The CR was found to be significant at all levels but L1–L2.

The test depicted in Table 9 confirmed the primary hypothesis that a loaded spine will generally experience measureable compression when compared to the unloaded spine. Supine measurements were significant less than standing for the L1–L5 distance and all individual vertebral levels except for L1–L2.

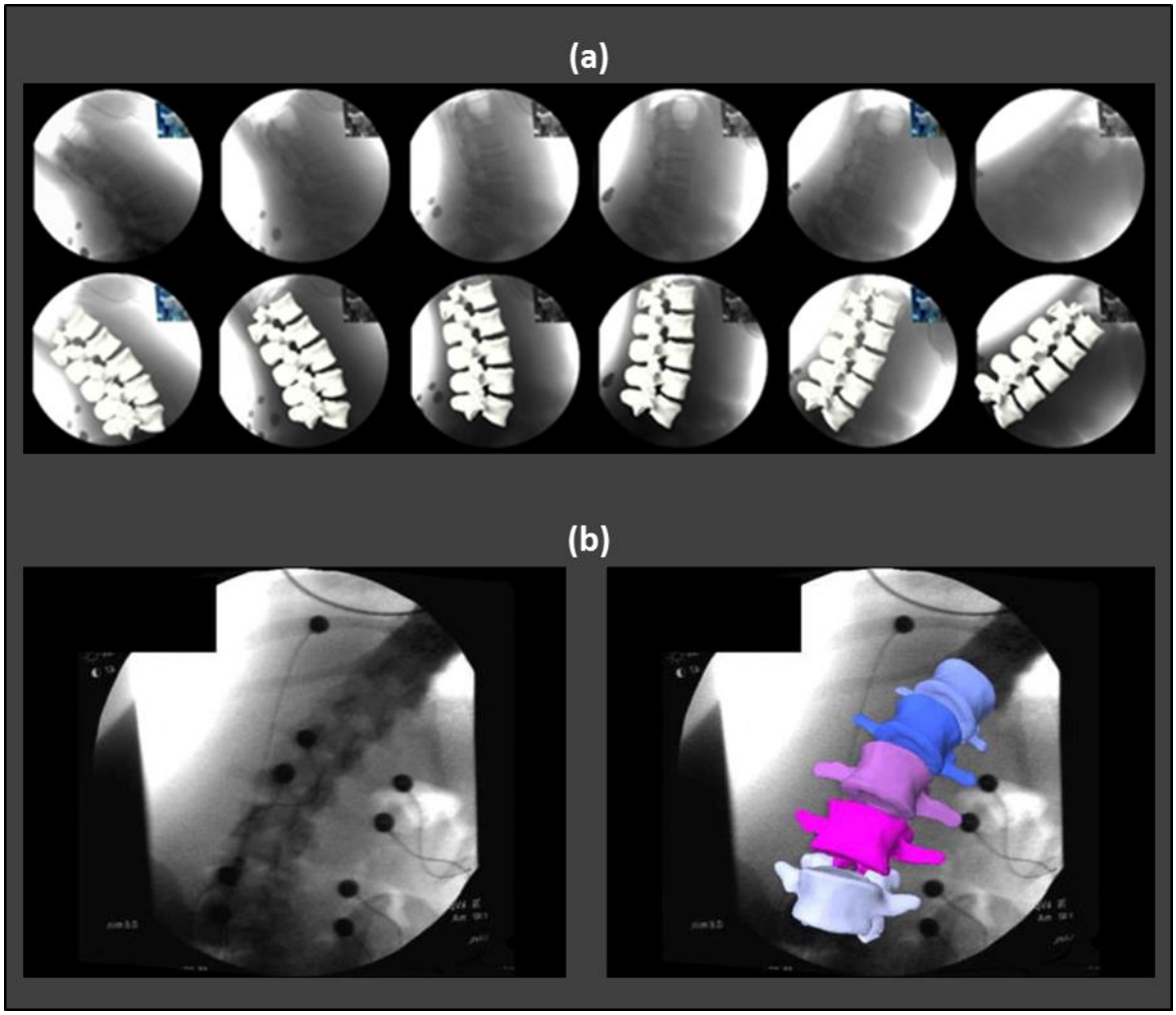

Figure 6.3: Overlay of 3D models on fluoroscopic imaging allowing for in vivo tracking of the vertebrae with six degree of freedom resolution during a flexion-extension activity (a). Overlay with separately colored surface models are tracked to determine the relative intervertebral kinematics during a lateral bending activity (b).

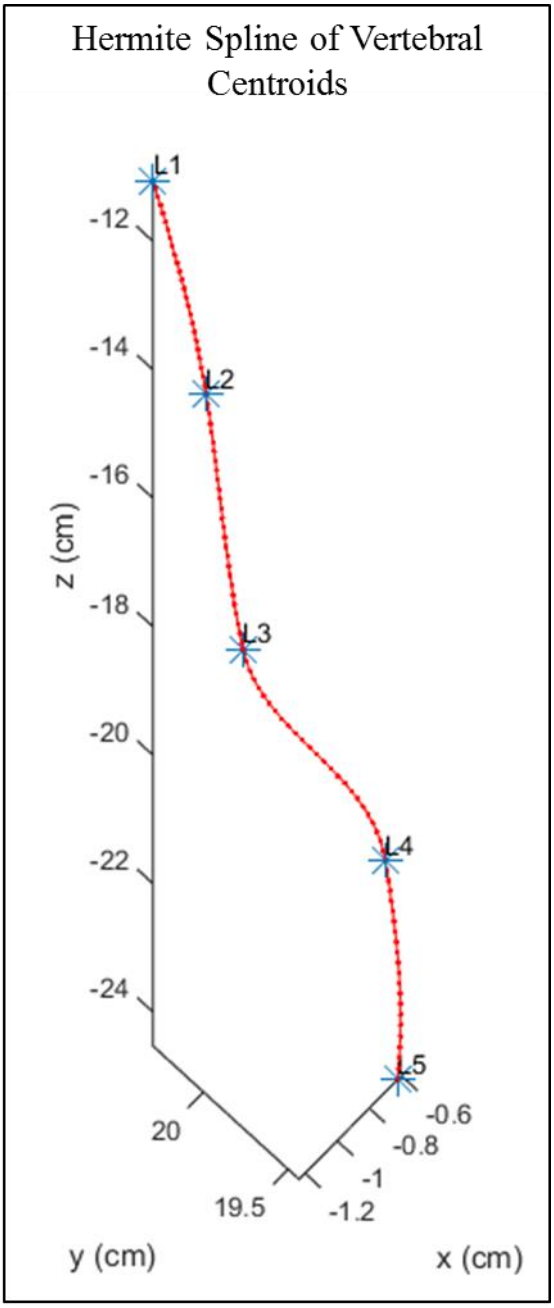

Figure 6.4: Hermine spline connecting the vertebral centroids by their 3D surface models.

Table 9: Compression rations of the intervertebral spacing from the supine-to-standing positions. Bolded values indicate significantly different values for $\alpha = 0.01$.

Level	Supine*	Standing*	p-value
L1–L2	3.46	3.44	0.1621
L2–L3	**3.53**	**3.43**	**0.0011**
L3–L4	**3.43**	**3.22**	**<0.0001**
L4–L5	**3.15**	**3.06**	**0.0007**
L1–L5	**13.58**	**13.16**	**<0.0001**

* Dimensionless measurements

When comparing between the healthy and degenerative groups, two of the adjacent vertebrae (L2–L3 and L4–L5) and at the overall distance (L1–L5) exhibited a significant difference for a Kruskal-Wallis test (Table 10).

Table 10: Means, standard deviations, and corresponding p-values of the intervertebral spacing for the healthy plus LBP groups versus the degenerative group. Bolded values indicate significant differences with a Kruskal-Wallis test.

Level	Healthy + LBP*	Degenerative*	p-value
L1–L2	2.122 ± 7.80	6.41 ± 11.77	0.1929
L2–L3	**3.504 ± 9.16**	**14.98 ± 7.91**	**0.0092**
L3–L4	9.124 ± 8.72	12.66 ± 9.51	0.2821
L4–L5	**2.181 ± 9.33**	**10.69 ± 9.53**	**0.0230**
L1–L5	**16.930 ± 20.60**	**44.74 ± 22.64**	**0.0055**

* Dimensionless measurements

A difference at L4–L5 is not terribly surprising as it is often thought to be the pathogenic beginning for lumbar pathology (along with L5–S1). A significant difference at L2–L3, however, is a bit more unexpected as degeneration generally begins in the lower lumbar and progressively

moves cranially through adjacent vertebral levels. Thus, if differences were observed at L2–L4, then we might expect differences at L3–L4 as well.

Nonetheless, the analysis shows greater translational motion in the degenerative spine than the healthy. A number of reasons could explain the observed motion in the degenerative spine: ligament laxity, weakening of surrounding musculature, poor posture, or weakening of the annulus wall, to name a few.

6.1 Multivariate Analysis

The compression ratios offered a new source of data that provided some separation of the clinical groups – given that the healthy and LBP groups were combined. Since the cross-sectional area (CSA) and volumetric data of the lean muscle and adipose tissues were drawing insignificant finding, the idea was to prepare a set of features that adequately provided some separation of the clinical groups.

Thus far, the data presented in this manuscript has evaluated several aspects of the soft tissue in the lumbar spine. These *in vivo* measurements include: the cross-sectional areas of six paraspinal muscles, the adipose content of those six muscles, and the relative motion experienced by vertebrae in the erect and supine positions. In view of the litany of measurements that could be garnered from the lumbar spine, there seems to be overwhelmingly large volumes of data, which can be exceedingly difficult to glean any discernable information from as the list of different measurements grows. For example, when added together, 17 measurements were extracted from

the current dataset (six lean muscle volumes, six adipose content volumes, and five vertebral translations), which ultimately represents a small sample of the innumerable measurements that could be performed. But what information does this reveal?

Multivariate analysis is an exceedingly interesting field and one that is seeing a resurgence in application due in part to increased computational capacity of everyday computers (i.e. home computer processing speeds are finally able to perform the analyses in a practical timeframe). While numerous subfields of multivariate analysis exist, the focus here will be on a branch sometimes referred to as supervised machine learning. Supervised machine learning is in itself a large, complex subfield with an exceptional amount of varying techniques and methods whose details exceed the scope and purpose of this dissertation. There are, however, some excellent resources available (e.g. [84, 85]). One of the more prominent and fundamental machine learning techniques is artificial *neural networks*. In the traditional sense, neural networks represent a set of supervised learning algorithms that were originally modeled after memory characteristics exhibited by neurons (i.e. the neuron model) [86]. A superb toolbox is available through the MATLAB software package that readily allows the user to explore the vast world of artificial neural networks (MATLAB and Neural Network Toolbox Release 2016a, The MathWorks, Inc., Natick, Massachusetts, United States). The complex flexibility of the Neural Network Toolbox provides an advantageous platform for performing pattern classification of the 17-feature dataset [87] (Appendix D).

In addition to the neural network classifier, two other popular classification techniques were employed to validate the veracity of neural network system. These were the naïve Bayes classifier [88] and support vector machines [89]. These are both reliable methods and the fact that they do not require much (if any) user defined parameterization means they are often used as a baseline classification techniques for testing and validation.

It can be tempting to utilize all available features of a system in order to better describe it, but in terms of classification and separability the addition of more features is not always beneficial. For example, an attempt to classify the lumbar based on the 17-features previously described. The data was evaluated with three learning algorithms (neural network, naïve bayes, and support vector machines) via leave-one-out cross validation for classification accuracy (Table 11).

Table 11: Prediction accuracy of all 17-features used to classify the population.

Feature Size	Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)
$N = 17$	Neural Network	65.4	75.0	50.0
$N = 17$	Naïve Bayes	57.7	41.2	88.9
$N = 17$	Support Vector Machine	69.2	76.5	55.6

The results from Table 11 show a maximum classification accuracy of less than 70%, hardly a desirable efficiency for accurate prediction. This highlights a dilemma frequently encountered in the field of machine learning coined the “curse of dimensionality” (Figure 6.5).

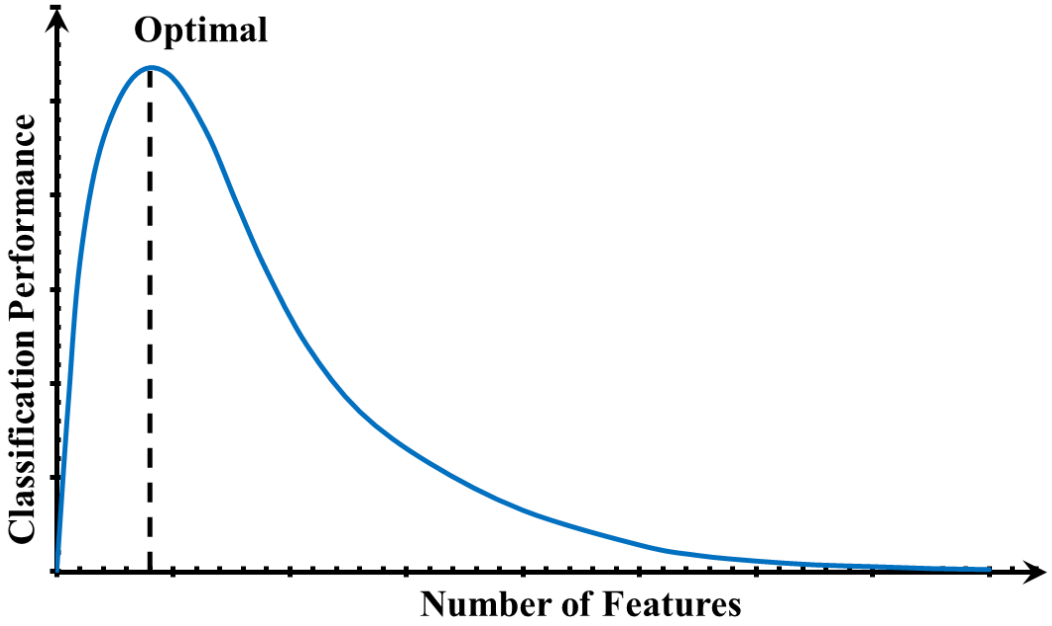

Figure 6.5: The Curse of Dimensionality suggests that there exists an optimal number of features for classification; however, after exceeding this arbitrarily optimal number the performance experiences a sharp decline in accuracy.

The curse of dimensionality is a phenomenon that arises in the presence of high-dimensional feature spaces, such as 17-feature dataset used to acquire the results in Table 11. This so called "curse" describes the steep decline often observed in classification performance after exceeding some optimal number of features in the input space (Figure 6.5). This effect was contrasted in Table 11 and Table 12, where 17 features lead to decidedly worse classification accuracy than just a single, highly separable feature. Granted the single feature (L1–L5 distance) was also present in the 17 feature analysis, the robustness of that single feature did not distill through concentrated enough to dilute the noise created by the other features. The task thus becomes an optimization problem of the number of features that maximizes the prediction accuracy.

The results from the 17-feature set are in stark contrast to if only a single classifier were used, for example the L1-L5 CR (Table 12).

Table 12: Prediction accuracy of a single feature (L1–L5 distance) used to classify the population.

Feature Size	Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)
$N = 1$	Neural Network	76.9	86.7	63.6
$N = 1$	Naïve Bayes	88.5	88.2	88.9
$N = 1$	Support Vector Machine	80.8	82.4	77.8

A more tactful approach must be devised that will ensure more accurate prediction (otherwise the multivariate analysis is a moot point). The obvious solution would be to perform dimension reduction that eliminated redundant and arbitrary features so that only the most important features that aid in classification remain. A popular dimension reducing technique that is often employed is called principal component analysis (PCA).

PCA involves mapping the original data with an orthogonal transformation that converts the set of possibly highly correlated variables into a new set of observations that are linearly uncorrelated. The variables in the orthogonal basis are called *principal components*. Each principal component is a linear combination of the original variables. The transformation is configured in such a way that the first principal component accounts for the largest possible variance in the data, while subsequent principal component progressively account for less and less variance (Figure 6.6). PCA was performed on the same 17 features used to tabulate the results in Table 11 (Figure 6.6).

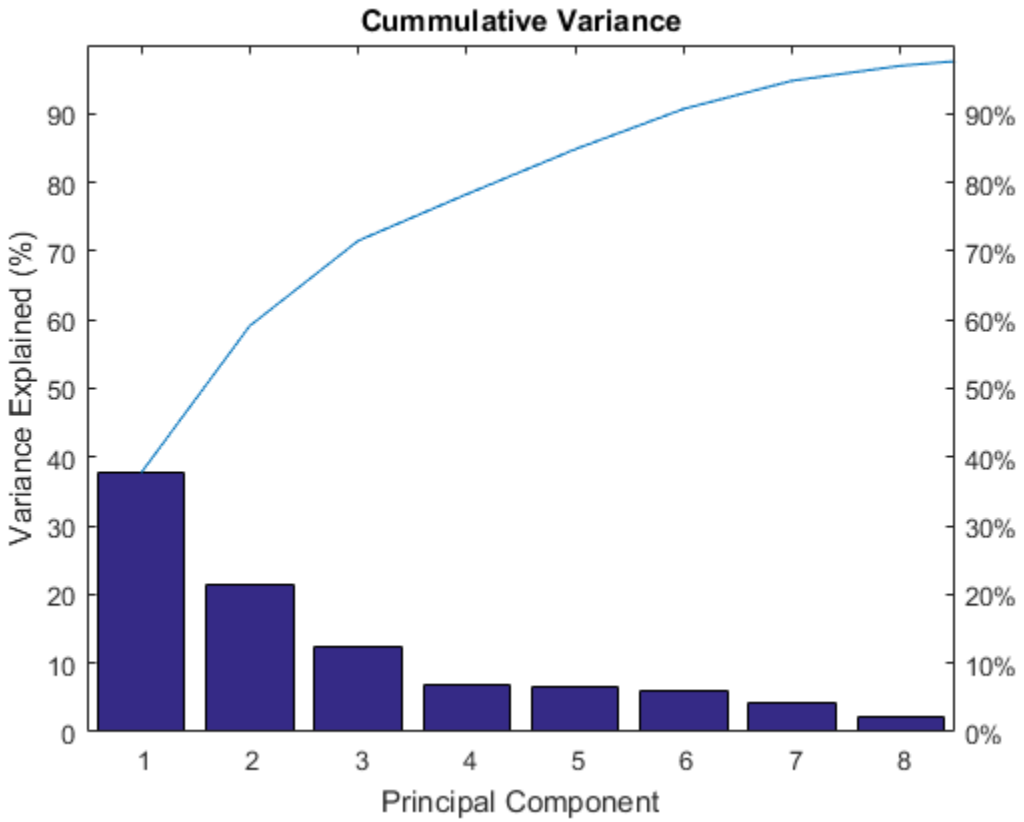

Figure 6.6: Variance of individual principal components (bar) the cumulative sum of the first eight principal components (line).

The first eight principal components account for 97.0% of the explained variance in the transformed basis. However, using the new coordinate system, the first 8 variables of the transformed data was again run through the pattern classification, but this time with an accuracy of 65.8% – approximately similar to 17-feature space in the original basis.

Instead of using the transformed basis we must turn to statistical significance for delineating between the healthy and pathologic group. This reduction is less mathematically rigorous and decidedly *ad hoc*, but nonetheless necessary. Each was feature was individually analysis using a Kruskal-Wallis test of significance. It was determined that of the original 17 features, five were statically difference between the healthy and degenerative groups. These five features were: the L2–L3, L4–L5, and L1–L5 distance; and the raw adipose volume in *only* the psoas muscles. Again running the reduced data with the three different classifiers unquestionably different results begin to emerge with the accuracies in the 80 and 90 percent range. To maximize the possible accuracy all possible (non-repeating) permutations of these five features were analyzed (Table 13). Analysis of the reduced dataset shows that the neural network and support vector machine classifiers can achieve up to 92% prediction accuracy between healthy and degenerative lumbar spines (Table 13).

Table 13: Maximum classification accuracy using the optimized features of the reduced (5-feature) space.

Features	Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)
Right/Left Psoas, L2–L3, L4–L5	Neural Network	92.0	87.5	100.0
Right Psoas, L1–L5	Naïve Bayes	84.0	92.9	72.7
Right Psoas, L2–L3, L4–L5; Right/Left Psoas, L2–L3, L4–L5	Support Vector Machine	92.0	87.5	100.0

These results are considerably different than what was achieved during the first classification process with 17 features. Though the proposed was decidedly performed *ad hoc*, the complexity of the data in some regards necessitated it. Nonetheless, 92% accuracy is an exceptional result that sheds light on the separability of healthy and degenerative spines, where the most prominent differences appear to be in the adipose content of the psoas muscles and the laxity of the intervertebral discs.

Chapter 7

Contributions and Future Work

In summary, this dissertation focused on delivering a system for analyzing the pathologic nature of the lumbar spine. In an exceedingly difficult feature space that the lumbar presents from the countless active and passive forces that provide stabilization and locomotion, this problem presents difficult challenge that require ingenuity and manipulation in order to discern pathologic differences.

Contributions:

1) Development of a spatially sensitive image filter that enhances tissue segmentation while in the presence of intensity inhomogeneity. This was established through testing on two phantom images and real MR images and validated by rigorous medical image processing metrics that measure the performance of processing. The proposed kernel was also implemented and tested in two different types of filters (Gaussian and anisotropic

diffusion) that showed significant performance increases over the original filters.

2) Develop a significant correlation between musculature/adipose and pathology of the lumbar spine. Five features were identified as provided statistical separation between pathologic and non-pathologic spines. These results are some of the most accurate predictor of healthy and degenerative spines in the literature.
3) Develop a significant correlation between musculature/adipose and sexual dimorphisms in the lumbar spine. These results were unintended but undeniably obvious and as such required more formal investigation. Though these results are preliminary they certainly indicate sexually dimorphic characteristics indicated by the stark muscular differences. These differences almost certainly point to proportionally different dynamics that propel the spine and keep it stabilized.
4) Design of a novel technique for analyzing higher dimensional spaces of the lumbar spine to draw classification of relevant pathology. To the author's knowledge, these set of features have yet to be comprised for evaluation of the lumbar spine.

Future work:

1) Fully generalize the baNLM algorithm for any filter. Though the proposed modification was extended to three different types of filters, the idea was generalized for any kernel based filter, it was not rigorously generalized to any filter.
2) Explore new ways to populate $\beta(i)$ for the baNLM kernel. The method proposed in this dissertation used the mean of the search window to populate $\beta(i)$. It would certainly be

an interesting exercise to test the effect of other values to populated the kernel. For example, the mean of the similarity window could be used, or the center values of either the search or similarity windows. Numerous possibilities exist.

3) Explore new methods for analyzing MRI/CT data by generating a dynamic neural network framework that uses slices in place of discrete time sampling
4) Investigate how baNLM performs with ultrasound scans and other forms of medical imaging. Ultrasound suffers from a similar spatially sensitive degradation of the target signal. baNLM May provide improved denoising of the images.

References

[1] D. I. Rubin, "Epidemiology and risk factors for spine pain," *Neurol Clin,* vol. 25, pp. 353-71, May 2007.
[2] G. B. Andersson, "Epidemiology of low back pain," *Acta Orthop Scand Suppl,* vol. 281, pp. 28-31, Jun 1998.
[3] J. W. Frymoyer, "Back pain and sciatica," *N Engl J Med,* vol. 318, pp. 291-300, Feb 4 1988.
[4] B. Vällfors, "Acute, subacute and chronic low back pain: clinical symptoms, absenteeism and working environment," *Scandinavian journal of rehabilitation medicine. Supplement,* vol. 11, pp. 1-98, 1984.
[5] J. L. Kelsey and A. A. White, 3rd, "Epidemiology and impact of low-back pain," *Spine (Phila Pa 1976),* vol. 5, pp. 133-42, Mar-Apr 1980.
[6] T. S. Carey, J. M. Garrett, A. Jackman, and N. Hadler, "Recurrence and care seeking after acute back pain: results of a long-term follow-up study. North Carolina Back Pain Project," *Med Care,* vol. 37, pp. 157-64, Feb 1999.
[7] T. R. Stanton, N. Henschke, C. G. Maher, K. M. Refshauge, J. Latimer, and J. H. McAuley, "After an episode of acute low back pain, recurrence is unpredictable and not as common as previously thought," *Spine (Phila Pa 1976),* vol. 33, pp. 2923-8, Dec 15 2008.
[8] J. L. St Sauver, D. O. Warner, B. P. Yawn, D. J. Jacobson, M. E. McGree, J. J. Pankratz, *et al.*, "Why patients visit their doctors: assessing the most prevalent conditions in a defined American population," *Mayo Clin Proc,* vol. 88, pp. 56-67, Jan 2013.
[9] B. S. Webster and S. H. Snook, "The cost of 1989 workers' compensation low back pain claims," *Spine (Phila Pa 1976),* vol. 19, pp. 1111-5; discussion 1116, May 15 1994.
[10] J. N. Katz, "Lumbar disc disorders and low-back pain: socioeconomic factors and consequences," *J Bone Joint Surg Am,* vol. 88 Suppl 2, pp. 21-4, Apr 2006.
[11] R. Chou, A. Qaseem, D. K. Owens, and P. Shekelle, "Diagnostic imaging for low back pain: advice for high-value health care from the American College of Physicians," *Ann Intern Med,* vol. 154, pp. 181-9, Feb 1 2011.
[12] R. Chou, A. Qaseem, V. Snow, D. Casey, J. T. Cross, Jr., P. Shekelle, *et al.*, "Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society," *Ann Intern Med,* vol. 147, pp. 478-91, Oct 2 2007.
[13] R. A. Deyo and J. N. Weinstein, "Low back pain," *N Engl J Med,* vol. 344, pp. 363-70, Feb 1 2001.
[14] H. Gray, "Anatomy of the human body," C. D. Clemente, Ed., 30th American ed. / edited by Carmine D. Clemente.. ed. Philadelphia: Philadelphia : Lea & Febiger, 1985.
[15] N. H. Kim and K. S. Suk, "Clinical and radiological differences between traumatic and idiopathic coccygodynia," *Yonsei Med J,* vol. 40, pp. 215-20, Jun 1999.
[16] F. Postacchini and M. Massobrio, "Idiopathic coccygodynia. Analysis of fifty-one operative cases and a radiographic study of the normal coccyx," *J Bone Joint Surg Am,* vol. 65, pp. 1116-24, Oct 1983.

[17] C. Leboeuf-Yde, J. Nielsen, K. O. Kyvik, R. Fejer, and J. Hartvigsen, "Pain in the lumbar, thoracic or cervical regions: do age and gender matter? A population-based study of 34,902 Danish twins 20-71 years of age," *BMC Musculoskelet Disord,* vol. 10, p. 39, 2009.
[18] L. Hansen, M. de Zee, J. Rasmussen, T. B. Andersen, C. Wong, and E. B. Simonsen, "Anatomy and biomechanics of the back muscles in the lumbar spine with reference to biomechanical modeling," *Spine (Phila Pa 1976),* vol. 31, pp. 1888-99, Aug 1 2006.
[19] M. J. Jorgensen, W. S. Marras, K. P. Granata, and J. W. Wiand, "MRI-derived moment-arms of the female and male spine loading muscles," *Clin Biomech (Bristol, Avon),* vol. 16, pp. 182-93, Mar 2001.
[20] M. R. Pierrynowski and J. B. Morrison, "A physiological model for the evaluation of muscular forces in human locomotion: theoretical aspects," *Mathematical Biosciences,* vol. 75, pp. 69-101, 1985/07/01 1985.
[21] M. Yoshio, G. Murakami, T. Sato, S. Sato, and S. Noriyasu, "The function of the psoas major muscle: passive kinetics and morphological studies using donated cadavers," *J Orthop Sci,* vol. 7, pp. 199-207, 2002.
[22] A. D. Skyrme, D. J. Cahill, H. P. Marsh, and H. Ellis, "Psoas major and its controversial rotational action," *Clin Anat,* vol. 12, pp. 264-5, 1999.
[23] L. Penning, "Psoas muscle and lumbar spine stability: a concept uniting existing controversies. Critical review and hypothesis," *Eur Spine J,* vol. 9, pp. 577-85, Dec 2000.
[24] A. Nachemson, "The possible importance of the psoas muscle for stabilization of the lumbar spine," *Acta Orthop Scand,* vol. 39, pp. 47-57, 1968.
[25] A. Nachemson, "Electromyographic studies on the vertebral portion of the psoas muscle; with special reference to its stabilizing function of the lumbar spine," *Acta Orthop Scand,* vol. 37, pp. 177-90, 1966.
[26] P. Le Floch-Prigent, "[The role of the psoas muscle: apropos of the dissection of the muscles from 10 adults and 10 newborn infants]," *Bull Assoc Anat (Nancy),* vol. 67, pp. 213-8, Jun 1983.
[27] M. A. Johnson, J. Polgar, D. Weightman, and D. Appleton, "Data on the distribution of fibre types in thirty-six human muscles. An autopsy study," *J Neurol Sci,* vol. 18, pp. 111-29, Jan 1973.
[28] A. C. Hooper, "The role of the iliopsoas muscle in femoral rotation," *Ir J Med Sci,* vol. 146, pp. 108-12, Apr 1977.
[29] S. Gracovetsky, H. F. Farfan, and C. Lamy, "The mechanism of the lumbar spine," *Spine (Phila Pa 1976),* vol. 6, pp. 249-62, May-Jun 1981.
[30] S. Gracovetsky, "An hypothesis for the role of the spine in human locomotion: a challenge to current thinking," *J Biomed Eng,* vol. 7, pp. 205-16, Jul 1985.
[31] H. F. Farfan, "Form and function of the musculoskeletal system as revealed by mathematical analysis of the lumbar spine. An essay," *Spine (Phila Pa 1976),* vol. 20, pp. 1462-74, Jul 1 1995.
[32] N. Bogduk, M. Pearcy, and G. Hadfield, "Anatomy and biomechanics of psoas major," *Clin Biomech (Bristol, Avon),* vol. 7, pp. 109-19, May 1992.

[33] J. V. Basmajian, "Electromyography of iliopsoas," *Anat Rec,* vol. 132, pp. 127-32, Oct 1958.
[34] E. Andersson, L. Oddsson, H. Grundstrom, and A. Thorstensson, "The role of the psoas and iliacus muscles for stability and movement of the lumbar spine, pelvis and hip," *Scand J Med Sci Sports,* vol. 5, pp. 10-6, Feb 1995.
[35] E. W. Donisch and J. V. Basmajian, "Electromyography of deep back muscles in man," *Am J Anat,* vol. 133, pp. 25-36, Jan 1972.
[36] J. Cholewicki, M. M. Panjabi, and A. Khachatryan, "Stabilizing function of trunk flexor-extensor muscles around a neutral spine posture," *Spine (Phila Pa 1976),* vol. 22, pp. 2207-12, Oct 1 1997.
[37] G. L. Moseley, P. W. Hodges, and S. C. Gandevia, "Deep and superficial fibers of the lumbar multifidus muscle are differentially active during voluntary arm movements," *Spine (Phila Pa 1976),* vol. 27, pp. E29-36, Jan 15 2002.
[38] M. Solomonow, B. H. Zhou, M. Harris, Y. Lu, and R. V. Baratta, "The ligamento-muscular stabilizing system of the spine," *Spine (Phila Pa 1976),* vol. 23, pp. 2552-62, Dec 1 1998.
[39] N. Bogduk, *Clinical anatomy of the lumbar spine and sacrum*: Elsevier Health Sciences, 2005.
[40] J. W. Cossette, H. F. Farfan, G. H. Robertson, and R. V. Wells, "The instantaneous center of rotation of the third lumbar intervertebral joint," *J Biomech,* vol. 4, pp. 149-53, Mar 1971.
[41] R. J. Last, *Anatomy : regional and applied*. Edinburgh: Churchill Livingstone, 1978.
[42] J. Haselgrove and M. Prammer, "An algorithm for compensation of surface-coil images for sensitivity of the surface coil," *Magnetic Resonance Imaging,* vol. 4, pp. 469-472, 1986.
[43] R. M. Henkelman, "Measurement of signal intensities in the presence of noise in MR images," *Med Phys,* vol. 12, pp. 232-3, Mar-Apr 1985.
[44] E. R. McVeigh, M. J. Bronskill, and R. M. Henkelman, "Phase and sensitivity of receiver coils in magnetic resonance imaging," *Med Phys,* vol. 13, pp. 806-14, Nov-Dec 1986.
[45] D. W. McRobbie, E. A. Moore, M. J. Graves, and M. R. Prince, *MRI from Picture to Proton*. New York: Cambridge University Press, 2006.
[46] U. Vovk, F. Pernus, and B. Likar, "A review of methods for correction of intensity inhomogeneity in MRI," *IEEE Trans Med Imaging,* vol. 26, pp. 405-21, Mar 2007.
[47] U. Vovk, F. Pernus, and B. Likar, "MRI intensity inhomogeneity correction by combining intensity and spatial information," *Phys Med Biol,* vol. 49, pp. 4119-33, Sep 7 2004.
[48] U. Vovk, F. Pernus, and B. Likar, "Intensity inhomogeneity correction of multispectral MR images," *Neuroimage,* vol. 32, pp. 54-61, Aug 1 2006.
[49] D. W. Shattuck, S. R. Sandor-Leahy, K. A. Schaper, D. A. Rottenberg, and R. M. Leahy, "Magnetic resonance image tissue classification using a partial volume model," *Neuroimage,* vol. 13, pp. 856-76, May 2001.

[50] M. Styner, C. Brechbuhler, G. Szekely, and G. Gerig, "Parametric estimate of intensity inhomogeneities applied to MRI," *IEEE Trans Med Imaging,* vol. 19, pp. 153-65, Mar 2000.
[51] O. Salvado, C. Hillenbrand, S. Zhang, and D. L. Wilson, "Method to correct intensity inhomogeneity in MR images for atherosclerosis characterization," *IEEE Trans Med Imaging,* vol. 25, pp. 539-52, May 2006.
[52] M. A. Bernstein, D. M. Thomasson, and W. H. Perman, "Improved detectability in low signal-to-noise ratio magnetic resonance images by means of a phase-corrected real reconstruction," *Med Phys,* vol. 16, pp. 813-7, Sep-Oct 1989.
[53] H. Gudbjartsson and S. Patz, "The Rician Distribution of Noisy MRI Data," *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine,* vol. 34, pp. 910-914, 1995.
[54] A. Macovski, "Noise in MRI," *Magn Reson Med,* vol. 36, pp. 494-7, Sep 1996.
[55] P. Milanfar, "A Tour of Modern Image Filtering: New Insights and Methods, Both Practical and Theoretical," *Signal Processing Magazine, IEEE,* vol. 30, pp. 106-128, 2013.
[56] S. P. Awate and R. T. Whitaker, "Unsupervised, information-theoretic, adaptive image filtering for image restoration," *IEEE Trans Pattern Anal Mach Intell,* vol. 28, pp. 364-76, Mar 2006.
[57] A. Buades, B. Coll, and J. M. Morel, "A non-local algorithm for image denoising," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, 2005, pp. 60-65 vol. 2.
[58] A. A. Samsonov and C. R. Johnson, "Noise-adaptive nonlinear diffusion filtering of MR images with spatially varying noise levels," *Magn Reson Med,* vol. 52, pp. 798-806, Oct 2004.
[59] D. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 14, pp. 367-383, 1992.
[60] A. Fan, W. M. Wells, J. W. Fisher, M. Cetin, S. Haker, R. Mulkern, *et al.*, "A unified variational approach to denoising and bias correction in MR," *Inf Process Med Imaging,* vol. 18, pp. 148-59, Jul 2003.
[61] C. Kervrann and J. Boulanger, "Optimal spatial adaptation for patch-based image denoising," *IEEE Trans Image Process,* vol. 15, pp. 2866-78, Oct 2006.
[62] A. Montillo, J. K. Udupa, L. Axel, and D. N. Metaxas, "Interaction between noise suppression and inhomogeneity correction in MRI," pp. 1025-1036, 2003.
[63] D. Palumbo, B. Yee, P. O'Dea, S. Leedy, S. Viswanath, and A. Madabhushi, "Interplay between bias field correction, intensity standardization, and noise filtering for T2-weighted MRI," *Conf Proc IEEE Eng Med Biol Soc,* vol. 2011, pp. 5080-3, 2011.
[64] A. Montillo, J. K. Udupa, L. Axel, and D. N. Metaxas, "Interaction between noise suppression and inhomogeneity correction in MRI," in *Medical Imaging 2003*, 2003, pp. 1025-1036.

[65] P. Kjaer, T. Bendix, J. S. Sorensen, L. Korsholm, and C. Leboeuf-Yde, "Are MRI-defined fat infiltrations in the multifidus muscles associated with low back pain?," *BMC Med,* vol. 5, p. 2, 2007.
[66] B. Mengiardi, M. R. Schmid, N. Boos, C. W. Pfirrmann, F. Brunner, A. Elfering, *et al.*, "Fat content of lumbar paraspinal muscles in patients with chronic low back pain and in asymptomatic volunteers: quantification with MR spectroscopy," *Radiology,* vol. 240, pp. 786-92, Sep 2006.
[67] A. Pezolato, E. E. de Vasconcelos, H. L. Defino, and M. H. Nogueira-Barbosa, "Fat infiltration in the lumbar multifidus and erector spinae muscles in subjects with sway-back posture," *Eur Spine J,* Mar 31 2012.
[68] P. Suri, A. L. Fry, and A. C. Gellhorn, "Do Muscle Characteristics on Lumbar Spine Magnetic Resonance Imaging or Computed Tomography Predict Future Low Back Pain, Physical Function, or Performance? A Systematic Review," *Pm r,* May 5 2015.
[69] A. J. Teichtahl, D. M. Urquhart, Y. Wang, A. E. Wluka, P. Wijethilake, R. O'Sullivan, *et al.*, "Fat infiltration of paraspinal muscles is associated with low back pain, disability, and structural abnormalities in community-based adults," *The Spine Journal,* vol. 15, pp. 1593-1601, 7/1/ 2015.
[70] A. M. Alshami, "Prevalence of spinal disorders and their relationships with age and gender," *Saudi Med J,* vol. 36, pp. 725-30, Jun 2015.
[71] E. E. Ekin, H. Kurtul Yildiz, and H. Mutlu, "Age and sex-based distribution of lumbar multifidus muscle atrophy and coexistence of disc hernia: an MRI study of 2028 patients," *Diagn Interv Radiol,* Apr 1 2016.
[72] O. P. Gautschi, M. V. Corniola, N. R. Smoll, H. Joswig, K. Schaller, G. Hildebrandt, *et al.*, "Sex differences in subjective and objective measures of pain, functional impairment and health-related quality of life in patients with lumbar degenerative disc disease," *Pain,* Jan 5 2016.
[73] T. Kienbacher, E. Fehrmann, R. Habenicht, D. Koller, C. Oeffel, J. Kollmitzer, *et al.*, "Age and gender related neuromuscular pattern during trunk flexion-extension in chronic low back pain patients," *J Neuroeng Rehabil,* vol. 13, p. 16, 2016.
[74] F. Stromqvist, B. Stromqvist, B. Jonsson, and M. K. Karlsson, "Gender differences in patients scheduled for lumbar disc herniation surgery: a National Register Study including 15,631 operations," *Eur Spine J,* vol. 25, pp. 162-7, Jan 2016.
[75] S. P. Awate and R. T. Whitaker, "Higher-order image statistics for unsupervised, information-theoretic, adaptive, image filtering," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, 2005, pp. 44-51 vol. 2.
[76] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 12, pp. 629-639, 1990.
[77] W. K. Pratt, *Digital Image Processing*. New York, NY, USA: Wiley, 1978.

[78] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans Image Process,* vol. 13, pp. 600-12, Apr 2004.
[79] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algoritms*. New York: Plenum Press, 1981.
[80] J. C. Dunn, "A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters," *Journal of Cybernetics,* vol. 3, pp. 32-57, 1973/01/01 1973.
[81] L. R. Dice, "Measures of the Amount of Ecologic Association Between Species," *Ecology,* vol. 26, pp. 297-302, 1945.
[82] J. M. Johnson, M. Mahfouz, N. V. Battaglia, A. Sharma, J. S. Cheng, and R. D. Komistek, "Clinical and statistical correlation of various lumbar pathological conditions," *J Biomech,* vol. 46, pp. 683-8, Feb 22 2013.
[83] M. R. Mahfouz, W. A. Hoff, R. D. Komistek, and D. A. Dennis, "A robust method for registration of three-dimensional knee implant models to two-dimensional fluoroscopy images," *IEEE Trans Med Imaging,* vol. 22, pp. 1561-74, Dec 2003.
[84] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*: Wiley, 1973.
[85] S. Theodoridis and K. Koutroumbas, "Pattern recognition," 2003.
[86] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The bulletin of mathematical biophysics,* vol. 5, pp. 115-133, 1943.
[87] M. T. Hagan, H. B. Demuth, M. H. Beale, and O. De Jesús, *Neural network design* vol. 20: PWS publishing company Boston, 1996.
[88] A. McCallum and K. Nigam, "A comparison of event models for naive bayes text classification," in *AAAI-98 workshop on learning for text categorization*, 1998, pp. 41-48.
[89] N. Cristianini and J. Shawe-Taylor, *An introduction to support vector machines and other kernel-based learning methods*: Cambridge university press, 2000.
[90] M. T. Hagan, H. B. Demuth, and M. H. Beale, *Neural network design* vol. 1: Pws Boston, 1996.
[91] L. Medsker and L. C. Jain, *Recurrent neural networks: design and applications*: CRC press, 1999.
[92] M. I. Jordan, "Generic constraints on underspecified target trajectories," in *Neural Networks, 1989. IJCNN., International Joint Conference on*, 1989, pp. 217-225.
[93] S. Chen, S. Billings, and P. Grant, "Non-linear system identification using neural networks," *International journal of control,* vol. 51, pp. 1191-1214, 1990.
[94] K. S. Narendra and K. Parthasarathy, "Identification and control of dynamical systems using neural networks," *Neural Networks, IEEE Transactions on,* vol. 1, pp. 4-27, 1990.
[95] J. L. Elman, "Finding structure in time," *Cognitive science,* vol. 14, pp. 179-211, 1990.
[96] M. F. Møller, "A scaled conjugate gradient algorithm for fast supervised learning," *Neural Networks,* vol. 6, pp. 525-533, // 1993.

Appendices

Appendix A

Convolution

A key operation utilized in image processing is convolution. In general terms, convolution is a mathematical operation between two functions that produces a third function that closely resembles that of one of the two original functions. In terms of image processing, convolution is performed between an image of interest and a convolution matrix, or kernel, which can have many desirable effects, such as, smoothing, sharpening, and edge detection, to name a few. The convolution operator is typically denoted with an asterisk, where the convolution of functions f and g are written as: $f * g$. In image processing the convolution of a kernel, H, with a given image, I, is defined as

$$H * I \stackrel{\text{def}}{=} \sum_{i \in R_H} I(u - i) \cdot H(i). \quad \textbf{(15)}$$

Much of the image processing material presented in this dissertation centers around the development and use of the filtering kernel and its influence on smoothing techniques utilized to reduce noise. Though the kernel the is often the primary focus of many discussions, it important to recognize that convolution is the mathematical relationship that relates the original with the kernel being discussed.

Appendix B

Gaussian Smoothing

One method for performing image noise reduction is to convolve the original image with a kernel of a low-pass filter such as is described by the Gaussian function. Given a noisy image $y = \{y(i)|i \in I\}$, the Gaussian function is written as:

$$G(i) = \frac{1}{\sqrt{2\pi h^2}} e^{\frac{i^2}{2h^2}} \quad \textbf{(16)}$$

where i is the pixel location and h is the standard deviation of the Gaussian distribution. Smoothing is performed by convolution of the Gaussian kernel with the original image:

$$y(i) = I(i) * G(i) = I(i) * \frac{1}{\sqrt{2\pi h^2}} e^{\frac{i^2}{2h^2}} \quad \textbf{(17)}$$

Appendix C

Anisotropic Diffusion

Anisotropic diffusion is frequently employed for image smoothing, edge detection, and coherence-enhancing filtering. Anisotropic diffusion is performed by solving for the partial differential of the diffusion equation. Given a noisy image $y = \{y(i)|i \in I\}$, the diffusion equation, in general, can be written as:

$$\frac{\partial}{\partial t}Y(i,t) = \nabla \cdot (C(i,t)\nabla Y(i,t)) \quad \textbf{(18)}$$

where Y is the density of the diffusing material at location i and time t. $C(i,t)$ is the tensor-based Perona–Malik diffusivity function:

$$C(i,t) = 1/(1 + \frac{|\nabla Y_\sigma|^2}{K^2}) \quad \textbf{(19)}$$

where Y_σ is the regularization of the image and K is a generalization of the Rayleigh distribution. The discretized form of the diffusion equation is solved by replacing the spatial derivatives with the central differences and $\partial u/\partial t$ with a forward difference approximation, which is explicitly represented as:

$$y(i)^{k+1} = y(i)^k + \tau(A(i)^k * y(i)^k) \quad \textbf{(20)}$$

where $y(i)^k$ is the approximation of $y(i,t)$ at pixel i and time, t, is given as $k\tau$ at the kth iteration with time step τ. $A(i)^k * y$ is the discretization of the right-hand side of the diffusion equation, where $A(i)^k$ is a 3x3 kernel that with space which is convolved with the image at the kth iteration. The discretized term of the diffusion function is then

$$z(i)^{k+1} = z(i)^k + \tau(A(i)^k * z(i)^k). \quad \textbf{(21)}$$

.

Appendix D

Pattern Classification

D.1 Neural Networks

The field of artificial neural networks (NN) consists of a large, and still growing, area of knowledge that contains numerous architectures and techniques that can be used in a wide range of applications. In the most basic sense, the architecture of a NN contains three phases: the input, learning, and output phases.

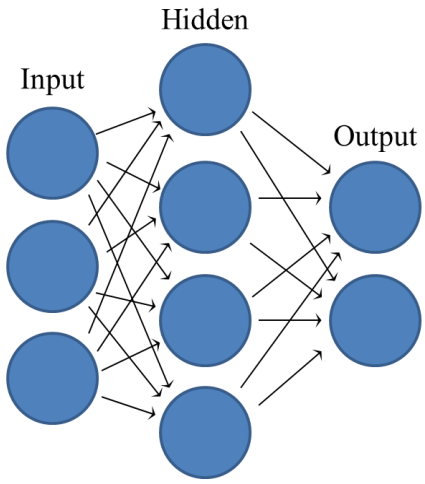

Figure D.1: An example 3-layer neural network architecture showing feedforward interconnections between the nodes of the input, hidden, and output layers.

Thus, it follows that a typical system is designed as a three layer network with an input, hidden, and output layers. It should be noted that it is possible to train with multiple hidden layers; however, it has been shown that a single hidden layer is sufficient to fully classify any system [ref]. The most commonly utilized training technique is feedforward with backpropagation, which implies that the inputs are introduced and initially forward propagated through the network (feedforward) and then optimization is iteratively performed between the hidden and output layers (backpropagation) until some predetermined criterion is met (Figure D.1)

D.2.1 Neuron Architecture

A single-layer is defined by five network elements: the input(p), weighting matrix (W), bias (b), transfer function (f), and output (a). A diagram of a single-layer of neurons is shown in Figure D.2.

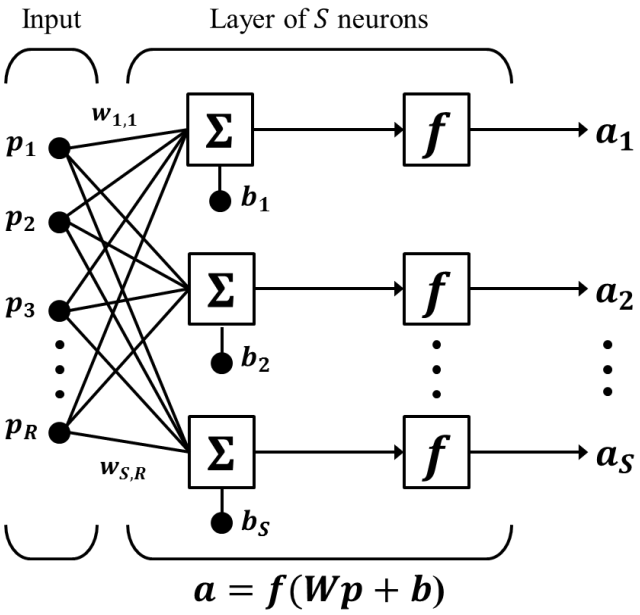

Figure D.2: A layer of multiple neurons and multiple inputs.

As shown in Figure D.2, the input vector has R elements, while the neuron layer has$S$ neurons, biases, and outputs. Connected between the input vector and summing junctions is the weighting matrix W, where:

$$W = \begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,R} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,R} \\ \vdots & \vdots & \ddots & \vdots \\ w_{S,1} & w_{S,2} & \cdots & w_{S,R} \end{bmatrix}. \quad \textbf{(22)}$$

The indices of Wtherefore indicate which input and which neuron is being represented for the individual weights w. It follows that the weighted inputs are summed with the biases, which are all a function of the transfer function. Finally, the output activation a is expressed as:

$$a = f(Wp + b) \quad \textbf{(23)}$$

The layer of S neurons in Figure D.3 represents the first two layers (input and hidden) derived from neuron in Figure D.2. The output from Figure D.2 could be represented in Figure D.3 by simply appending another layer of S neurons to the hidden layer. In which case, the outputs from the hidden layer serve as the inputs to the output layer (Figure D.3).

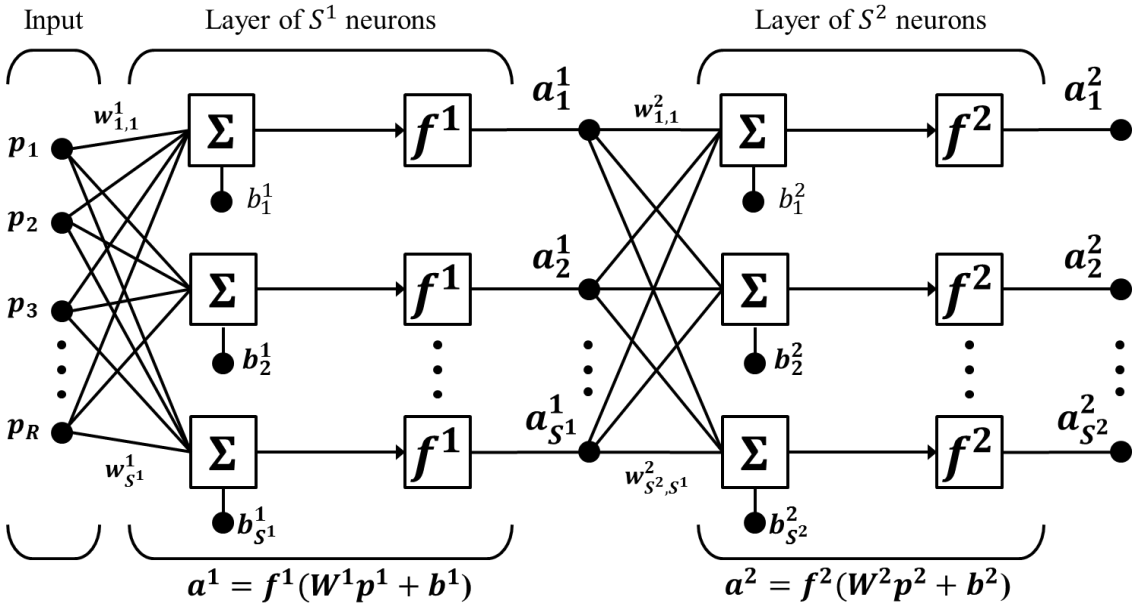

Figure D.3: Two layers of multiple neurons used to create the input-hidden-output network structure.

Note that the size of S^1 and S^2 are independent of each other, as well that specification of the transfer functions f^1 and f^2 need not correspond to each other.

D.2.2 Abbreviated Notation

Describing the input/hidden/output structure in diagram form (Section D.2.1) is extremely useful in neural network design, but can quickly become an overly cumbersome task and actually understanding structural details even more perilous. Therefore, an abbreviated notation was developed by [90] who simplified the design by removing the interconnecting nodes arrows, which are not structurally significant, and using vector notation with the inputs and outputs to more closely resemble Equation (23) (Figure D.4)

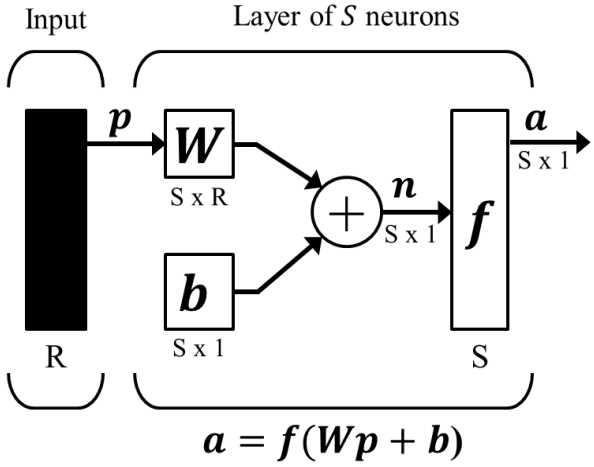

Figure D.4: The abbreviated notation for a 2-layer neural network greatly simplifies the visual appearance while emphasizing the structurally pertinent elements of the network.

The appeal of this abbreviated notation is that it contains only the elements necessary for network design – this is especially true when utilizing the Neural Network Toolbox for MATLAB, as the user is provided with significant control over how these specific elements are combined and manipulated to construct a network. The need for an abbreviated notation will

become more apparent in later chapters where network architectures, too complex to coherently present with the bulkier versions, will contain additional elements such as feedback loops, delays, additional weights, and networks built within other networks.

D.2.3 Backpropagation

Training a neural network generally involves optimization of some performance function. Most often optimization is accomplished by iteratively adjusting the weights and biases of the network until sufficient performance has been obtained. The most common index of performance is the mean squared error (MSE) between the network outputs a and target outputs t, shown by:

$$mse = \frac{1}{N}\sum_{i=1}^{N}(t_i - a_i)^2, \tag{24}$$

where N is the number of output elements.

Backpropagation is by far the most widely utilized training algorithm. It begins by first forward propagating errors of the inputs using random weight values. The error sensitivities are then backward propagated through each neuron while updating the weights and biases until first layer is reached. New errors values are again determined by forward propagation, and the cycle continues so on. This forward-backward propagation is iteratively executed until a satisfactory accuracy of error has been obtained.

Backpropagation involves determining the sensitivity s of the error $\hat{F}$ due to changes in the network input n at layer m. The sensitivity is defined as:

$$s_i^m \equiv \frac{\partial \hat{F}}{\partial n_i^m}. \tag{25}$$

Using the notation defined in Equation (23) the first iteration of forward propagation is defined as:

$$a^{m+1} = f^{m+1}(W^{m+1}a^m + b^{m+1}) \text{ for } m = 0, 1, \dots, M-1, \tag{26}$$

where M is total number of layers and m is the current layer, the external inputs are passed to first neuron (i.e. $a^0 = p$), and the outputs of the last layer are the final outputs of the network (i.e. $a = a^M$). The error sensitivities are then backward propagated from last-to-first neuron. The sensitivity in the last neuron is calculated by:

$$s^M = -2\dot{F}^M(n^M)(t-a) \tag{27}$$

While the sensitivities at all other neurons are found by:

$$s^m = \dot{F}^m(n^m)(W^{m+1})^T s^{m+1} \text{ for } m = M-1, \dots, 2, 1. \tag{28}$$

Finally the weights and bias are updated using an approximation of the mean squared error. One of the more frequently utilized methods for finding the approximation is to determine

convergence by the steepest gradient [90]. This so called *gradient descent* method updates the weights and biases by:

$$W^m(k+1) = W^m(k) - \alpha s^m (a^{m-1})^T \quad \textbf{(29)}$$

and

$$b^m(k+1) = m^m(k) - \alpha s^m \quad \textbf{(30)}$$

where α is the learning rate.

D.3 Recurrent Neural Networks

The philosophy behind recurrent neural networks is vast, and too broad to abridge in a presentation here; however, the brief introduction given here provides enough information to proceed forward. The signature characteristic of recurrent neural networks is a feedback element. Feedback can be introduced in a variety of ways and is generally used to send the activation energies back to one of the neuron layers. A majority of recurrent neural networks employ one of two (or both) philosophies [91]: 1) nonlinear autoregressive network with exogenous outputs or 2) layer-recurrent networks.

The first introduces feedback from the output layer to the input layer [92]. One noted network notorious for using this architecture is nonlinear autoregressive network with exogenous inputs

(NARX) [93, 94]. This network is noted for giving more emphasis on the sequence of the output nodes (Figure D.5).

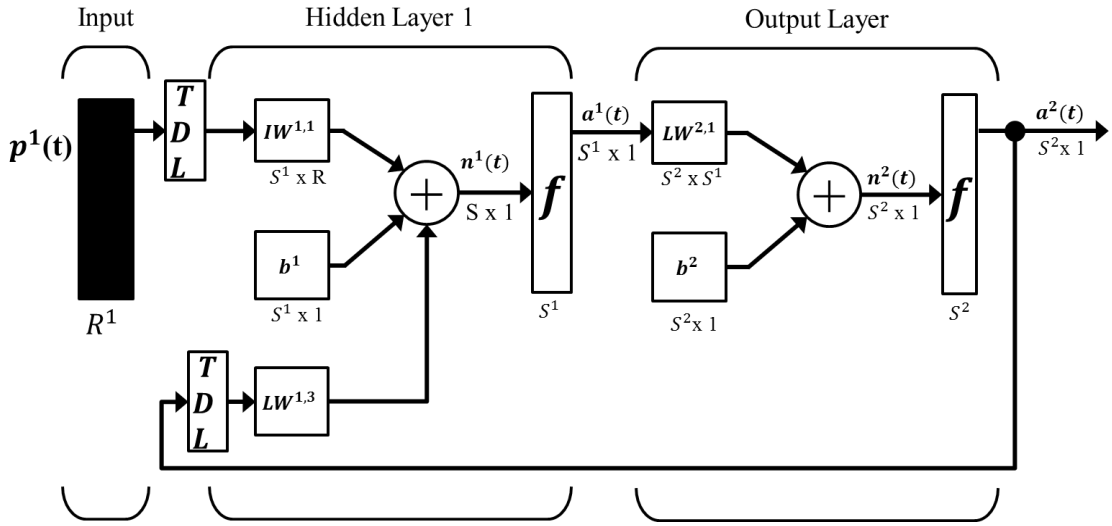

Figure D.5: Example of a NARX recurrent neural network

This network structure introduces several new network elements. Specifically are the: input weights (IW), layer weights (LW), and tapped delay line (TDL). The first two elements (IW and LW) are self-evident, while the TDL indicates how long the feedback loop is delayed to the network. For example, the TDL may have delay of 10 inputs, so, the feedback would not be initiated until the eleventh sequential input, whereupon memory is used with a first-in-first-out (FIFO) structure.

The second is layer-recurrent network (LRN). In this network the feedback is sent backwards prior to reaching the output layer [95]. This network gives more emphasis on the sequence of the input nodes (Figure D.6).

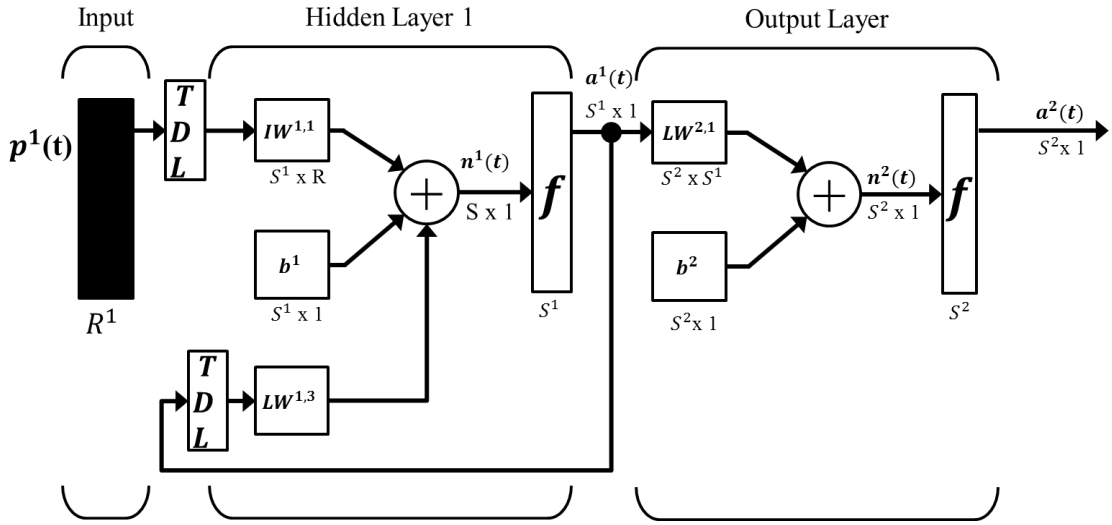

Figure D.6: Example of a layer-recurrent neural network.

Though both network structures will be explored, this network appears to be more relevant for the lumbar data (and other similar biomedical data that is multi-featured and temporal) presented in this proposal.

D.4 Architecture

The proposed research involves extracting *in vivo* lumbar image data and generalizing pathological classifications of a subject population based on characteristics of the paraspinal muscle and adipose tissues by dynamic machine learning. The overall process is elucidated by a

system that contains two primary modules (Figure D.7).

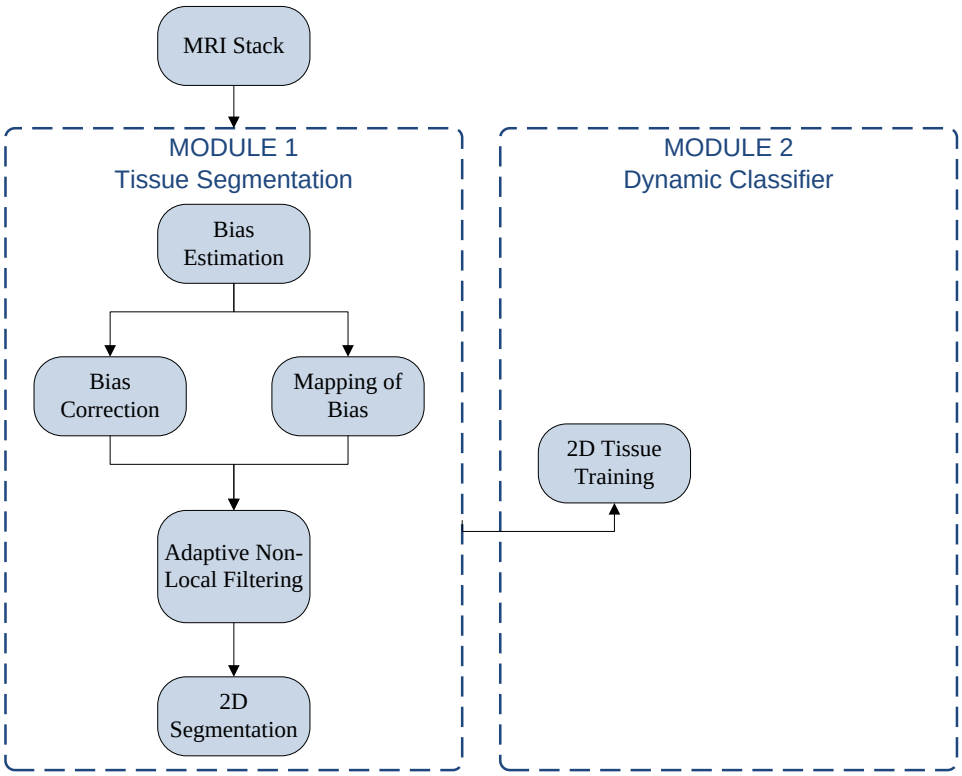

Figure D.7: System architecture with two primary modules.

MODULE 1 contains the imaging processing tools necessary to accurately extract the paraspinal muscle and adipose tissues. This is performed by estimating the global intensity inhomogeneity of a stack of lumbar MR images. The bias estimation reveals a mapping to inhomogeneity of the signal-to-noise ratio (SNR), which is turn is used to parameterized the kernel of the non-local means (NLM). This adaptive kernel enhances the filtering to promote more accurate and

homogeneous segmentations. This step is vital as the inter-group population statistics are strikingly different and, thus, require low-error segmentation results.

MODULE 2 contains the system for dynamic classification of muscle and adipose data. This data is unique from the perspective of pattern classification because the same six muscles are targeted throughout a stack of 50 axial MR slices. The goal is to observe patterns exhibited by the individual muscle as they traverse the superior-inferior (z) axis. In a sense the data (2D segmentations) moving equidistant along the z-axis is analogous to 2D data in the temporal domain. However, each MR stack contains the features of each paraspinal muscle and adipose boundary. Dynamic neural networks, therefore, inherently provide an excellent platform for constructing a machine capable of accurately classifying the complex nature of a multi-featured, temporally dependent, 2D population dataset.

A.2 MODULE 2: Dynamic Classification

We employed the scaled conjugate gradient (SCG) algorithm to train the neural network (NN). A full explanation of the SCG algorithm exceeds the scope and purpose of this paper. Therefore, for a detailed explanation the reader is encouraged to review [96], where SCG was originally introduced. In short, SCG is a variant of the conjugate gradient algorithm (CG), which uses a line searching method to determine the steepest gradient of the global error function. SCG instead uses a Lagrange Multiplier λ to scale the iterative outputs of the global error function.

Scaling the error function by λ leads to quicker convergence than the line searching method without yielding performance of the network [96]. This was an added benefit for our purposes as the computation time was of concern – given the large feature set and the complexity of the data. An additional benefit of using SCG is that the algorithm is largely self-contained and requires the user to specify only two scalar values, λ and σ, where λ was defined above and σ wad defined as the spread of the radial basis kernel function. Minimizing the number of user defined network parameters is desirable because it allows the network to train without a specific dependence on heuristic information, perhaps allowing the network to learn more consistently across a broad dataset.

[1] suggests selecting the scalar values to be on the order of $0 < \lambda \leq 10^{-6}$, $0 < \sigma \leq 10^{-4}$. However, using a 64-bit, double precision, number, the lower bound of those intervals are more realistically 10^{-16}, where round off errors begin to have an effect. In fact, the training sessions for our network failed to converge at $< 10^{-16}$. To determine appropriate values of λ and σ, the network was trained for 20 iterations with the values of λ and σ suggested by [1], while monitoring the MSE of the training session (Fig 1). The results showed that $\sigma \in [10^{-13}, 10^{-8}]$ and $\lambda \in [10^{-13}, 10^{-10}]$ returned the most consistent and lowest MSE of the network training (fig. 1).

D.2 Number of Hidden Layer Neurons

Determining the optimal number of hidden layer neurons requires a delicate balance of the generalizability of the network and the specificity of the network to the input data. For example, if too few neurons are supplied to the network, then there is small specificity and, thus, the

predictive power is lost; if too many neurons are supplied to the network, then the network overfits the data and generalization is lost. Determining the correct balance of the hidden layer for generalizability and specificity is ultimately a reflection on the quality, quantity, and non-linearity of the input data [ref or prove somehow]. That is, more separable and linear datasets require less complex network structure to perform accurate classifications.

[2] proposed a solution for determining the optimal number of neurons by performing singular valued decomposition (SVD) on the hidden layer's output activation matrix to determine the impact that each neuron has on the outcome of the output layer. By eliminating non-influential hidden neurons, over-fitting is avoided and minimization of the objective function is accomplished without sacrificing accuracy of the network. The results of the current study were obtained using criterion 6) from [ref], with $\lambda = 0.5$ and $\gamma = 0.95$ (note that [2] use of λ differs from our earlier use of λ as an SCG training parameter). [I can explain this in further detail should we decide it's needed, NVB]

To determine the optimal number of hidden neurons, the network was initially trained with 60 hidden neurons, k_0, and allowed to converge for 5000 epochs. Using criteria 6) the hidden layer was reduced to the specified number of neurons determined by the γ-threshold and re-trained with the updated number of neurons, k_r. This process was performed for 20 iterations. Over these iterations the mean value of hidden layer neurons was reduced to 28.1, with a standard deviation of 1.0. The mean MSE values for k_0 was 6.37×10^{-3} and for k_r was 4.50x10, with a p-value = 0.0006.

Appendix E

Muscle Volume Data

Psoas R					
	L5	L4	L3	L2	L1
Healthy 1	45.08470	52.57579	29.81140	19.59676	3.38142
Healthy 2	25.01432	26.41000	16.93490	7.62492	2.45937
Healthy 3	42.34377	42.05525	27.31032	14.54418	2.10394
Healthy 4	14.68236	24.13051	19.72624	9.87311	3.12592
Healthy 5	43.14502	56.57249	45.64349	28.07941	6.79934
Healthy 6	49.44639	61.73977	41.61638	15.08298	1.78239
Healthy 7	11.40175	24.36689	21.14538	14.53984	5.79908
Healthy 8	38.42007	41.56511	25.23940	14.91874	4.70061
Healthy 9	27.70225	34.07055	22.97209	11.43564	3.90632
Healthy 10	15.75214	24.25566	22.68183	18.01337	6.77848
LBP 1	28.42877	43.58562	33.75422	20.91856	4.98566
LBP 2	33.92107	31.44084	29.64715	16.45085	3.71078
LBP 3	18.30537	34.56068	32.17865	16.23880	4.26870
LBP 4	16.62465	17.17301	12.10393	7.78569	1.75980
LBP 5	35.27590	45.03430	43.64645	24.11313	7.40332
LBP 6	13.51351	16.71416	10.96810	7.91170	2.31337
LBP 7	10.75258	19.27000	16.54209	10.62309	3.06509
Degen 1	23.84026	34.56416	37.67270	26.30310	7.09133
Degen 2	28.23671	21.73371	19.26913	8.69470	2.72790
Degen 3	36.15276	53.99319	44.73448	25.19595	5.58616
Degen 4	46.64810	47.72744	38.09853	20.24853	6.93577
Degen 5	34.97869	50.73691	37.02787	22.08567	7.25037
Degen 6	10.16511	19.18309	16.73763	10.13296	2.89302
Degen 7	12.74267	23.11374	18.02380	10.24680	2.49500
Degen 8	37.48586	60.16247	47.51627	26.96705	4.56852
Degen 9	54.10964	62.48367	37.19907	22.80871	7.71964

Psoas L					
	L5	L4	L3	L2	L1
Healthy 1	40.74126	52.50453	32.24296	22.48630	5.19944
Healthy 2	24.98999	25.47318	16.96097	8.87112	2.19084
Healthy 3	43.49784	47.05742	30.32674	16.82279	4.07752
Healthy 4	13.35534	20.72476	15.90075	7.88824	3.21978
Healthy 5	45.06559	53.56997	42.97208	27.89605	6.37351
Healthy 6	52.31247	56.82190	42.47499	18.80506	3.09551
Healthy 7	10.96375	24.20351	21.64160	15.47405	5.69305
Healthy 8	40.77081	45.48272	27.80567	17.98730	6.23272
Healthy 9	25.35238	27.50498	20.11731	11.10975	4.81620
Healthy 10	19.35951	23.83678	18.96062	14.47553	6.79586
LBP 1	28.76074	43.12590	34.47378	21.74588	6.60641
LBP 2	35.94071	30.39887	26.20838	15.44190	2.86608
LBP 3	17.97340	34.92220	34.77360	19.75405	8.28799
LBP 4	17.78482	15.88597	10.52837	7.39115	2.37073
LBP 5	33.51958	46.77237	41.54425	19.61153	6.17711
LBP 6	16.41522	18.78942	11.55209	9.16398	3.48657
LBP 7	11.78412	19.75579	16.85929	10.65177	2.53845
Degen 1	25.14555	36.49776	37.06437	27.88736	8.45832
Degen 2	24.12530	26.47344	24.24523	13.95237	4.63457
Degen 3	36.43172	50.25373	41.41564	23.81158	5.14469
Degen 4	46.86362	48.61386	42.43936	21.70330	8.52176
Degen 5	33.99581	48.61560	39.13007	23.36750	8.97019
Degen 6	10.85252	21.00720	18.34708	11.74415	3.48223
Degen 7	12.40114	23.00685	18.37055	11.56860	3.33101
Degen 8	35.81383	56.84450	45.82425	28.51828	6.69766
Degen 9	50.01301	61.25485	36.92011	21.91447	7.71443

Erector R					
	L5	L4	L3	L2	L1
Healthy 1	22.23167	42.48107	41.15753	48.63906	45.38452
Healthy 2	10.76735	33.06159	31.36002	34.86571	28.83026
Healthy 3	31.95531	47.92471	53.43353	55.28284	50.86814
Healthy 4	3.38750	27.19387	44.55546	48.53999	39.85659
Healthy 5	12.93038	49.30127	67.54580	67.85431	55.19246
Healthy 6	9.91657	50.32064	57.76828	48.10895	62.01787
Healthy 7	1.00721	16.01372	26.76630	31.70677	30.88553
Healthy 8	18.05335	45.09861	48.42180	51.70589	48.20976
Healthy 9	9.25871	29.28737	31.32265	30.40843	25.40105
Healthy 10	8.95976	25.17510	32.23862	48.69381	44.84311
LBP 1	6.69331	29.23697	51.22270	51.58857	44.00102
LBP 2	10.66394	35.24374	55.07253	53.80982	42.36202
LBP 3	6.56122	45.17508	67.35983	52.98424	48.08983
LBP 4	15.11688	26.92186	29.83573	26.20317	23.54218
LBP 5	5.95463	27.06003	42.34724	40.55355	40.11208
LBP 6	4.51899	16.36394	18.43746	24.02362	22.01354
LBP 7	3.14591	21.35568	32.36724	36.33699	24.69539
Degen 1	7.77613	29.66193	42.03960	43.98624	35.85989
Degen 2	13.27365	26.49342	43.03899	40.29979	34.77881
Degen 3	11.26531	61.70154	74.45637	74.43291	60.85510
Degen 4	15.97201	39.69147	59.41684	62.00309	57.51539
Degen 5	7.64404	32.91386	43.51696	52.28206	46.07454
Degen 6	6.57599	19.93481	24.34951	32.05178	26.09019
Degen 7	5.58008	29.36906	27.74136	33.70729	25.80254
Degen 8	10.67697	43.54825	48.42702	45.89465	34.44162
Degen 9	22.83218	43.95843	48.11677	51.18099	51.23139

Mult R					
	L5	L4	L3	L2	L1
Healthy 1	18.03510	14.98739	10.44320	7.94386	5.17771
Healthy 2	19.82184	21.32440	11.80846	5.91031	4.71017
Healthy 3	21.09584	21.93359	16.89405	12.81654	8.55739
Healthy 4	8.63561	16.75935	9.74015	6.21621	4.93352
Healthy 5	26.58902	27.70312	16.25010	10.41192	8.24107
Healthy 6	14.17050	16.01285	12.59667	6.68288	7.00530
Healthy 7	7.00443	16.13191	10.73781	7.64056	5.14122
Healthy 8	21.58163	18.88154	15.83905	11.51125	7.32858
Healthy 9	17.79090	17.29903	11.04892	8.07508	6.09629
Healthy 10	20.62396	18.13851	11.53384	9.19961	8.91718
LBP 1	12.85130	20.40496	11.95446	7.26948	5.30373
LBP 2	17.51715	14.84487	11.82845	7.70748	6.51777
LBP 3	14.85008	22.68270	17.17041	9.28738	8.78247
LBP 4	16.63856	14.55461	9.27696	5.98939	3.91587
LBP 5	14.31389	16.64725	13.01989	7.80046	5.38715
LBP 6	11.51646	9.03797	4.73016	3.83245	3.16068
LBP 7	9.55852	14.51898	11.95446	6.42739	2.39159
Degen 1	17.14173	20.73172	17.06091	12.95037	6.89754
Degen 2	22.80871	14.72668	9.39688	4.95524	3.38229
Degen 3	28.24540	30.00693	17.67792	11.42869	8.51916
Degen 4	23.50568	23.62474	15.27156	9.98348	8.06726
Degen 5	20.47361	23.57520	17.05917	9.49161	6.52733
Degen 6	10.26331	16.96879	9.73494	6.12583	3.01903
Degen 7	13.77856	19.03014	10.74650	5.43321	3.40575
Degen 8	16.06934	19.65498	12.19779	8.10810	4.51464
Degen 9	25.87988	20.65524	9.82879	9.02928	8.11071

Mult L					
	L5	L4	L3	L2	L1
Healthy 1	20.73780	16.57425	10.51707	7.36334	4.96306
Healthy 2	21.95619	18.98930	10.74910	4.80316	3.47962
Healthy 3	21.10888	21.96227	17.19648	12.39419	8.82940
Healthy 4	9.79143	15.98157	9.91570	6.78022	5.59398
Healthy 5	27.79524	24.95784	15.85034	10.35022	7.98122
Healthy 6	14.13747	15.73824	12.72094	6.61423	6.35613
Healthy 7	7.48327	13.81680	10.00260	6.34831	4.36256
Healthy 8	23.90457	18.02119	15.46623	9.90527	7.11480
Healthy 9	15.99113	15.80081	10.40062	6.75936	4.55288
Healthy 10	20.72737	18.22107	11.99617	9.47249	9.07621
LBP 1	12.86347	15.05170	11.61640	8.26714	4.71365
LBP 2	18.65820	15.24462	12.83131	8.54349	5.87294
LBP 3	13.09029	21.59901	17.72225	10.12514	8.25410
LBP 4	13.87242	12.55496	9.17528	5.18293	3.72034
LBP 5	13.59433	15.72781	13.80289	7.36334	4.62153
LBP 6	10.41539	7.04788	3.79160	3.62301	3.29191
LBP 7	10.86295	16.04674	14.42599	8.12201	3.62301
Degen 1	15.84252	18.93021	17.04266	11.05500	6.52298
Degen 2	20.42929	16.35004	11.42347	5.95724	4.43121
Degen 3	25.40626	26.84538	15.50794	10.60571	8.15764
Degen 4	24.35646	22.91300	16.05283	9.54462	7.43547
Degen 5	18.82853	22.33683	16.92708	10.01824	6.67159
Degen 6	9.93916	14.98218	8.88328	5.57400	3.22325
Degen 7	13.56825	17.19735	9.74884	5.36630	3.17719
Degen 8	16.13278	18.59823	13.75510	9.31606	5.02476
Degen 9	24.72233	23.34838	10.91596	7.90127	8.14895

Erector L					
	L5	L4	L3	L2	L1
Healthy 1	23.42833	45.88596	44.02709	51.56771	47.89864
Healthy 2	15.48100	33.13981	29.68539	28.48004	25.09514
Healthy 3	33.23193	49.83920	54.26259	52.89386	50.76385
Healthy 4	3.42226	27.12260	48.38096	52.22209	42.47934
Healthy 5	14.94741	55.96851	70.30586	68.49305	58.72422
Healthy 6	12.07091	58.15153	65.73821	53.85849	63.45525
Healthy 7	0.19293	17.84565	28.59649	30.95505	30.65350
Healthy 8	24.68757	60.08165	56.86709	53.56736	42.92341
Healthy 9	14.21047	28.35577	31.18448	28.00381	25.22029
Healthy 10	9.83922	22.42807	30.97070	44.65714	42.42545
LBP 1	7.33292	31.91968	44.53634	48.23757	40.76212
LBP 2	9.61501	36.68113	61.12363	58.85197	46.92358
LBP 3	5.09081	39.28302	66.75585	54.27563	45.79818
LBP 4	11.46780	25.18118	30.12947	28.53826	25.92073
LBP 5	3.68558	29.75578	46.41694	40.29023	41.07671
LBP 6	9.44294	21.71025	19.61414	24.90222	19.73841
LBP 7	3.40575	24.03144	34.36254	38.88327	24.95002
Degen 1	7.81089	27.99772	44.95869	47.12607	35.50445
Degen 2	7.55105	26.24922	37.82130	39.60717	36.62030
Degen 3	13.83418	67.66660	80.91070	79.46376	61.65808
Degen 4	15.54444	40.55964	60.16160	61.74933	58.45308
Degen 5	7.40679	31.80323	38.01336	49.16222	46.47777
Degen 6	6.63769	23.46570	24.65194	33.23193	24.45032
Degen 7	6.38828	32.89387	30.97939	34.27042	27.34855
Degen 8	13.56043	45.64436	53.19455	49.66365	36.67505
Degen 9	19.80532	42.41329	45.79818	51.65288	57.03742

Appendix F

Adipose Volume Data

Psoas R					
	L5	L4	L3	L2	L1
Healthy 1	7.07097	8.24585	4.67554	3.07350	0.53033
Healthy 2	4.20074	4.43512	2.84393	1.28048	0.41301
Healthy 3	5.92441	5.88404	3.82105	2.03491	0.29437
Healthy 4	2.05424	3.37615	2.75994	1.38137	0.43735
Healthy 5	5.27790	6.92048	5.58354	3.43494	0.83176
Healthy 6	6.24087	7.79248	5.25261	1.90370	0.22496
Healthy 7	2.36411	5.05238	4.38441	3.01478	1.20242
Healthy 8	5.49372	5.94343	3.60900	2.13324	0.67214
Healthy 9	4.17013	5.12878	3.45808	1.72145	0.58803
Healthy 10	2.20392	3.39366	3.17347	2.52029	0.94839
LBP 1	3.97753	6.09816	4.72263	2.92676	0.69755
LBP 2	4.59487	4.25891	4.01594	2.22840	0.50265
LBP 3	2.77610	5.24129	4.88004	2.46269	0.64737
LBP 4	2.67712	2.76542	1.94913	1.25375	0.28339
LBP 5	5.53258	7.06306	6.84540	3.78184	1.16112
LBP 6	2.80197	3.46562	2.27419	1.64046	0.47967
LBP 7	1.96733	3.52571	3.02660	1.94364	0.56080
Degen 1	3.61549	5.24182	5.71324	3.98899	1.07543
Degen 2	4.11199	3.16499	2.80608	1.26617	0.39725
Degen 3	4.89717	7.31380	6.05964	3.41299	0.75669
Degen 4	6.79317	6.95035	5.54813	2.94871	1.01003
Degen 5	4.92950	7.15028	5.21828	3.11250	1.02178
Degen 6	2.27210	4.28779	3.74118	2.26491	0.64665
Degen 7	2.41639	4.38305	3.41785	1.94310	0.47313
Degen 8	5.45891	8.76121	6.91959	3.92710	0.66529
Degen 9	7.09867	8.19726	4.88017	2.99229	1.01274

Psoas L					
	L5	L4	L3	L2	L1
Healthy 1	6.38975	8.23467	5.05690	3.52669	0.81547
Healthy 2	4.19665	4.27779	2.84831	1.48976	0.36791
Healthy 3	6.08588	6.58391	4.24308	2.35371	0.57049
Healthy 4	1.86858	2.89965	2.22471	1.10366	0.45049
Healthy 5	5.51285	6.55318	5.25675	3.41251	0.77967
Healthy 6	6.60261	7.17177	5.36097	2.37348	0.39070
Healthy 7	2.27329	5.01850	4.48730	3.20848	1.18043
Healthy 8	5.82985	6.50361	3.97596	2.57202	0.89122
Healthy 9	3.81640	4.14044	3.02834	1.67240	0.72500
Healthy 10	2.70863	3.33506	2.65282	2.02530	0.95082
LBP 1	4.02398	6.03384	4.82330	3.04251	0.92432
LBP 2	4.86845	4.11776	3.55013	2.09173	0.38823
LBP 3	2.72575	5.29612	5.27358	2.99579	1.25691
LBP 4	2.86394	2.55816	1.69541	1.19022	0.38177
LBP 5	5.25712	7.33566	6.51569	3.07582	0.96880
LBP 6	3.40363	3.89591	2.39528	1.90012	0.72293
LBP 7	2.15607	3.61459	3.08464	1.94889	0.46444
Degen 1	3.81344	5.53506	5.62099	4.22925	1.28274
Degen 2	3.51327	3.85521	3.53073	2.03182	0.67491
Degen 3	4.93496	6.80726	5.61007	3.22546	0.69689
Degen 4	6.82455	7.07943	6.18027	3.16056	1.24099
Degen 5	4.79098	6.85133	5.51454	3.29315	1.26416
Degen 6	2.42575	4.69551	4.10093	2.62504	0.77834
Degen 7	2.35162	4.36278	3.48360	2.19375	0.63166
Degen 8	5.21542	8.27802	6.67319	4.15300	0.97535
Degen 9	6.56123	8.03606	4.84357	2.87497	1.01206

Erector R					
	L5	L4	L3	L2	L1
Healthy 1	3.48676	6.66262	6.45504	7.62842	7.11799
Healthy 2	1.80820	5.55214	5.26639	5.85512	4.84156
Healthy 3	4.47094	6.70525	7.47600	7.73475	7.11707
Healthy 4	0.47395	3.80475	6.23385	6.79134	5.57642
Healthy 5	1.58177	6.03099	8.26284	8.30058	6.75166
Healthy 6	1.25162	6.35121	7.29122	6.07206	7.82758
Healthy 7	0.20884	3.32038	5.54989	6.57427	6.40399
Healthy 8	2.58146	6.44869	6.92388	7.39347	6.89356
Healthy 9	1.39375	4.40875	4.71513	4.57751	3.82372
Healthy 10	1.25358	3.52230	4.51058	6.81286	6.27410
LBP 1	0.93648	4.09061	7.16668	7.21787	6.15628
LBP 2	1.44451	4.77404	7.46001	7.28896	5.73827
LBP 3	0.99504	6.85101	10.21544	8.03531	7.29305
LBP 4	2.43431	4.33530	4.80453	4.21957	3.79106
LBP 5	0.93391	4.24403	6.64163	6.36031	6.29107
LBP 6	0.93699	3.39300	3.82294	4.98120	4.56442
LBP 7	0.57559	3.90731	5.92203	6.64835	4.51836
Degen 1	1.17929	4.49837	6.37550	6.67072	5.43832
Degen 2	1.93299	3.85813	6.26759	5.86869	5.06469
Degen 3	1.52597	8.35796	10.08570	10.08252	8.24330
Degen 4	2.32594	5.78010	8.65262	9.02925	8.37572
Degen 5	1.07726	4.63850	6.13278	7.36803	6.49322
Degen 6	1.46986	4.45581	5.44258	7.16419	5.83166
Degen 7	1.05815	5.56924	5.26058	6.39190	4.89293
Degen 8	1.55484	6.34175	7.05222	6.68344	5.01559
Degen 9	2.99536	5.76693	6.31246	6.71446	6.72107

Mult R					
	L5	L4	L3	L2	L1
Healthy 1	2.82858	2.35058	1.63788	1.24589	0.81206
Healthy 2	3.32875	3.58108	1.98303	0.99254	0.79100
Healthy 3	2.95157	3.06878	2.36368	1.79319	1.19728
Healthy 4	1.20823	2.34484	1.36277	0.86972	0.69026
Healthy 5	3.25262	3.38891	1.98786	1.27368	1.00812
Healthy 6	1.78853	2.02106	1.58989	0.84348	0.88417
Healthy 7	1.45234	3.34489	2.22644	1.58424	1.06601
Healthy 8	3.08598	2.69989	2.26484	1.64600	1.04792
Healthy 9	2.67814	2.60409	1.66324	1.21557	0.91770
Healthy 10	2.88554	2.53780	1.61373	1.28714	1.24762
LBP 1	1.79805	2.85490	1.67257	1.01709	0.74206
LBP 2	2.37284	2.01085	1.60226	1.04404	0.88288
LBP 3	2.25208	3.43994	2.60397	1.40848	1.33190
LBP 4	2.67935	2.34377	1.49389	0.96449	0.63058
LBP 5	2.24495	2.61091	2.04201	1.22340	0.84491
LBP 6	2.38789	1.87399	0.98078	0.79464	0.65536
LBP 7	1.74886	2.65644	2.18723	1.17598	0.43757
Degen 1	2.59962	3.14406	2.58737	1.96398	1.04604
Degen 2	3.32154	2.14458	1.36843	0.72161	0.49255
Degen 3	3.82606	4.06467	2.39461	1.54811	1.15399
Degen 4	3.42303	3.44037	2.22393	1.45385	1.17480
Degen 5	2.88532	3.32242	2.40412	1.33764	0.91989
Degen 6	2.29405	3.79285	2.17595	1.36924	0.67481
Degen 7	2.61282	3.60868	2.03785	1.03030	0.64583
Degen 8	2.34011	2.86227	1.77631	1.18075	0.65745
Degen 9	3.39520	2.70977	1.28944	1.18456	1.06405

Mult L					
	L5	L4	L3	L2	L1
Healthy 1	3.25246	2.59946	1.64947	1.15485	0.77839
Healthy 2	3.68718	3.18894	1.80513	0.80661	0.58434
Healthy 3	2.95339	3.07279	2.40600	1.73410	1.23534
Healthy 4	1.36994	2.23602	1.38733	0.94864	0.78267
Healthy 5	3.40017	3.05308	1.93896	1.26614	0.97634
Healthy 6	1.78436	1.98640	1.60557	0.83481	0.80224
Healthy 7	1.55163	2.86486	2.07400	1.31630	0.90456
Healthy 8	3.41813	2.57687	2.21153	1.41636	1.01735
Healthy 9	2.40721	2.37856	1.56565	1.01751	0.68536
Healthy 10	2.90001	2.54935	1.67841	1.32532	1.26987
LBP 1	1.79976	2.10592	1.62528	1.15667	0.65950
LBP 2	2.52740	2.06500	1.73810	1.15728	0.79554
LBP 3	1.98520	3.27559	2.68766	1.53552	1.25177
LBP 4	2.23391	2.02176	1.47752	0.83462	0.59910
LBP 5	2.13210	2.46671	2.16481	1.15485	0.72483
LBP 6	2.15959	1.46135	0.78617	0.75122	0.68256
LBP 7	1.98752	2.93597	2.63943	1.48603	0.66288
Degen 1	2.40259	2.87085	2.58460	1.67654	0.98924
Degen 2	2.97503	2.38099	1.66355	0.86753	0.64530
Degen 3	3.44148	3.63642	2.10067	1.43663	1.10502
Degen 4	3.54693	3.33672	2.33771	1.38994	1.08280
Degen 5	2.65348	3.14790	2.38551	1.41186	0.94022
Degen 6	2.22159	3.34880	1.98558	1.24590	0.72046
Degen 7	2.57294	3.26113	1.84867	1.01761	0.60249
Degen 8	2.34935	2.70838	2.00310	1.35666	0.73174
Degen 9	3.24333	3.06309	1.43207	1.03657	1.06906

Erector L					
	L5	L4	L3	L2	L1
Healthy 1	3.67444	7.19663	6.90509	8.08775	7.51230
Healthy 2	2.59978	5.56528	4.98517	4.78275	4.21431
Healthy 3	4.64955	6.97311	7.59200	7.40050	7.10248
Healthy 4	0.47882	3.79478	6.76909	7.30651	5.94338
Healthy 5	1.82851	6.84659	8.60047	8.37871	7.18370
Healthy 6	1.52353	7.33959	8.29714	6.79774	8.00900
Healthy 7	0.04000	3.70023	5.92937	6.41841	6.35588
Healthy 8	3.53010	8.59113	8.13148	7.65964	6.13766
Healthy 9	2.13916	4.26851	4.69433	4.21553	3.79651
Healthy 10	1.37663	3.13796	4.33318	6.24808	5.93584
LBP 1	1.02597	4.46595	6.23118	6.74903	5.70312
LBP 2	1.30243	4.96875	8.27967	7.97196	6.35617
LBP 3	0.77205	5.95746	10.12384	8.23115	6.94551
LBP 4	1.84669	4.05500	4.85183	4.59560	4.17409
LBP 5	0.57804	4.66682	7.27991	6.31902	6.44236
LBP 6	1.95796	4.50154	4.06691	5.16338	4.09268
LBP 7	0.62313	4.39688	6.28709	7.11422	4.56494
Degen 1	1.18456	4.24599	6.81820	7.14689	5.38442
Degen 2	1.09963	3.82256	5.50776	5.76783	5.33286
Degen 3	1.87395	9.16597	10.95999	10.76399	8.35207
Degen 4	2.26367	5.90653	8.76108	8.99230	8.51228
Degen 5	1.04383	4.48198	5.35717	6.92836	6.55004
Degen 6	1.48365	5.24504	5.51018	7.42798	5.46512
Degen 7	1.21141	6.23765	5.87461	6.49869	5.18610
Degen 8	1.97475	6.64700	7.74650	7.23231	5.34083
Degen 9	2.59827	5.56422	6.00829	6.77637	7.48277

Vita

Nicholas V. Battaglia was born in Lakewood, CO, to the parents of Daniel and Michelle Battaglia. He is the youngest of three siblings, with an older brother, Andrew and older sister, Kelly. He attended South Lakewood Elementary and Creighton Middle before graduating from Lakewood High Schools in May 2000. He received his Bachelor Science from the University of Colorado, Denver in May 2005 with a major in Chemistry and minors in Mathematics and Physics. In June 2008 he was accepted to the graduate school at the University of Tennessee, Knoxville where he would begin his graduate career in biomedical engineering. In 2010 he accepted a Graduate Research Assistantship under the guidance of Dr. Mohamed Mahfouz in the department of Mechanical, Aerospace, and Biomedical Engineering. Nicholas graduated with a Doctor of Philosophy degree in Biomedical Engineering in May 2016. His research interests include medical image processing, computer vision, pattern recognition, and machine learning.