Theses and Dissertations

May 2019

Machine Intelligence for Advanced Medical Data Analysis: Manifold Learning Approach

Fereshteh S Bashiri

University of Wisconsin-Milwaukee

Follow this and additional works at: https://dc.uwm.edu/etd

Part of the Artificial Intelligence and Robotics Commons, and the Electrical and Electronics Commons

Machine Intelligence for Advanced Medical Data Analysis: Manifold Learning Approach

by

Fereshteh Sadat Bashiri

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Engineering

at

The University of Wisconsin–Milwaukee

May 2019

ABSTRACT

MACHINE INTELLIGENCE FOR ADVANCED MEDICAL DATA ANALYSIS:
MANIFOLD LEARNING APPROACH

by
Fereshteh Sadat Bashiri

The University of Wisconsin–Milwaukee, 2019
Under the Supervision of Prof. Zeyun Yu and Prof. Roshan M. D'Souza

In the current work, linear and non-linear manifold learning techniques, specifically Principle Component Analysis (PCA) and Laplacian Eigenmaps, are studied in detail. Their applications in medical image and shape analysis are investigated.

In the first contribution, a manifold learning-based multi-modal image registration technique is developed, which results in a unified intensity system through intensity transformation between the reference and sensed images. The transformation eliminates intensity variations in multi-modal medical scans and hence facilitates employing well-studied monomodal registration techniques. The method can be used for registering multi-modal images with full and partial data.

Next, a manifold learning-based scale invariant global shape descriptor is introduced. The proposed descriptor benefits from the capability of Laplacian Eigenmap in dealing with high dimensional data by introducing an exponential weighting scheme. It eliminates the limitations tied to the well-known cotangent weighting scheme, namely dependency on triangular mesh representation and high intra-class quality of 3D models.

In the end, a novel descriptive model for diagnostic classification of pulmonary nodules is presented. The descriptive model benefits from structural differences between benign and malignant nodules for automatic and accurate prediction of a candidate nodule. It extracts concise and discriminative features automatically from the 3D surface structure of a nodule

using spectral features studied in the previous work combined with a point cloud-based deep learning network.

Extensive experiments have been conducted and have shown that the proposed algorithms based on manifold learning outperform several state-of-the-art methods. Advanced computational techniques with a combination of manifold learning and deep networks can play a vital role in effective healthcare delivery by providing a framework for several fundamental tasks in image and shape processing, namely, registration, classification, and detection of features of interest.

To dad and mom,

for their endless love and support.

TABLE OF CONTENTS

LIST OF FIGURES

List of Tables

Prerequisite Notations

X	Matrix X
x	Vector x
x	Scalar value x
V	set V

Acknowledgments

The important thing is to not stop questioning.
Curiosity has its own reason for existing.

Albert Einstein

Chapter 1

Manifold Learning for Medical Data Analysis

1.1 Introduction

There are diverse medical data sources available ranging from electronic health records (EHRs), exam reports and clinical notes to medical images and gene expression generated in structured and unstructured data formats. Their rapid expansion in size and complexity have turned into a big challenge in today's health informatics research area. A technical report published by EMC [1] indicates overall a 48% annual increase in medical data which exceeds 2,314 Exabytes by 2020 [1]. By utilizing such data sources, the healthcare community could deliver efficient diagnoses and treatment for a large number of patients worldwide.

Diagnostic methods developed in computational health informatics include several components, such as data acquisition, data processing, data visualization, data storage and transmission as illustrated in Figure 1.1 [2]. The aim of data processing is to turn data into high quality information and knowledge systematically. This has been widely explored in the literature under topics such as EHR based phenotyping [3–6], clinical and scientific text

---

[1]https://www.dellemc.com/en-us/index.htm

mining [7, 8], and medical image analysis [9–12] purposing to name a few.
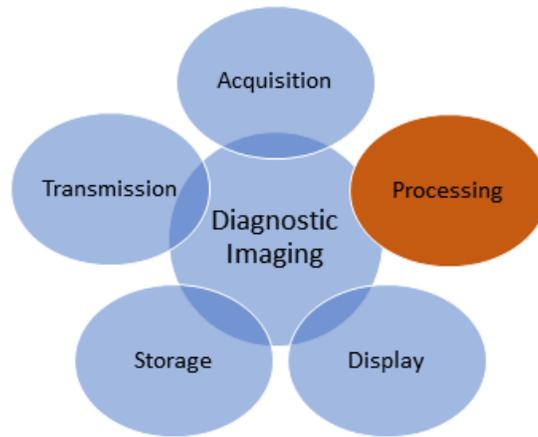
Figure 1.1: Components of modern diagnostic imaging.

As noted in the New England Journal of Medicine (NEJM), the evolution of medical imaging has been recognized as one of the top 10 medical developments in the past thousand years that have had an immeasurable impact on the clinical medicine [13]. IBM estimates that medical images recently accounted for almost 90% of all clinical data, making them the largest and fastest growing data in the healthcare community [14]. IBM Watson notes that 60 billion images were generated in the United States in 2015. With the extensive amount of medical images being generated every day, it is infeasible for physicians to effectively keep track of all images and extract information required to find the best solution for each patient [15].

Medical images play a vital role in early detection and diagnosis of a variety of diseases. With the massive number of medical images and the variety of image modalities, there exists a pressing need to design and develop efficient machine intelligence methods to harness this wealth of data, providing a "second opinion" and a precise interpretation of medical images [2, 16, 17].

In today's modern world, medical image processing often deals with high dimensional data. Working with such data is considerably time demanding and computationally expen-

sive [18]. However, the desired information may lie on or near a low dimensional smooth manifold. In order to uncover the underlying low dimensional manifold, several dimensionality reduction techniques have been developed.

Dimensionality reduction algorithms have shown a great potential in improving as well as enhancing image processing techniques. In brief, dimensionality reduction algorithms uncover or *learn* a low dimensional manifold embedded in a high dimensional space while respecting the intrinsic geometry of the data [19]. These algorithms hereupon are referred to as *manifold learning* techniques. The size of the low dimensional space must be chosen according to the intrinsic dimensionality of the data which is defined as the minimum number of parameters required for representing the data [20].

The extensive amount of medical images and capabilities of manifold learning for dealing with such high dimensional data **motivate** the **main contributions** of the current thesis. To make sure the current contribution is self-contained, Section 1.2 briefly explains the preliminary concepts of Manifold learning from a mathematical perspective. Our contributions are extended to both 2D and 3D medical image analysis, as well as 3D shape analysis, with the use of manifold learning, which are summarized in Section 1.3.

1.2 Manifold Learning

Manifold learning techniques, as advanced computational strategies, are categorized to *linear* and *nonlinear* methods with respect to the transformation that maps high dimensional space into the low dimensional space. Linear dimensionality reduction methods have been used by almost all scientific disciplines since the invention of the *Principal Component Analysis* (PCA) by Karl Pearson in 1901 [21]. However, real-world data is more likely to be nonlinear causing linear methods to be incapable of dealing with them adequately [20]. Accordingly, nonlinear manifold learning techniques have been introduced in order to cope with complex nonlinear relationships between data points [19]. Figure 1.2 shows the concept of manifold

learning with a synthetic dataset which is embedded in (a) 2D space as a low dimensional manifold, (b) 3D space with a linear transformation, and (c)-(d) 3D space with nonlinear transformation.

(a) Flat 2D-manifold   (b) Flat 3D-manifold   (c) The swiss roll   (d) The S-curve
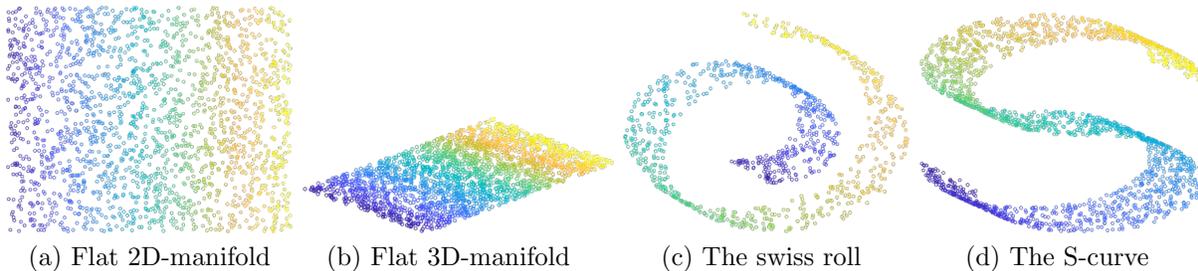
Figure 1.2: An example showing the concept of manifold learning. (a) A flat plane embedded in 2D space representing a low dimensional manifold. This manifold is embedded in 3D space via: (b) Linear transformation, (c) Nonlinear transformation on the surface of a swiss roll, (d) Nonlinear transformation on the surface of a s-curve shape.

Traditional linear methods include PCA and Multidimensional Scaling (MDS) [22, 23]. There are other linear methods such as Fishers Linear Discriminant Analysis (LDA) [24], Independent Component Analysis (ICA) [25] and Projection Pursuit (PP) [26], to name a few. Some popular nonlinear manifold learning techniques in chronological order are Multilayer Auto-encoders [27, 28], Kernel PCA [29], Isomap [30, 31], LLE [32, 33], Laplacian Eigenmap [34], Hessian LLE [35], Manifold Charting [36], Local Tangent Space Alignment [37], Maximum Variance Unfolding (MVU) [38], Diffusion maps [39], Riemannian Manifold Learning (RML) [40], Locally Multidimensional Scaling [41]. The major component of all types of algorithms is a set of eigenvectors associated with the top (or bottom) eigenvalues of a specific matrix adapted to a given application. Therefore, these algorithms are considered *spectral embedding methods* [42].

The other classification of manifold learning techniques is contingent on the properties of the high dimensional data preserved while mapping to the low dimensional space. Algorithms that tend to preserve global geometric characteristics of the manifold are called *global techniques* while, on the other hand, those that to preserve local geometric characteristics of the manifold are called *local techniques*. In general, linear manifold learning algorithms

are global, while a majority of nonlinear algorithms are local. Manifold learning techniques, their theory and performance on both real and synthetic data have been well studied in the literature. Interested readers are referred to [18–20, 40, 42–49] for further information.

From the application perspective, manifold learning has made a significant advance in many research studies, such as visualization [50], clustering [51], classification [52], shape analysis [53], registration [54], and segmentation [55]. Nonlinear manifold learning methods are engaged in medical data processing extensively, due to the nonlinear and high dimensional nature of the medical data.

Although several linear, as well as nonlinear manifold learning methods, are introduced in the data/image processing community, the common characteristics of all types of algorithms are flexibility and simplicity. These properties are briefly discussed in [18]. Two widely used methods, indeed, are *PCA* and *Laplacian Eigenmaps* which are employed in the presented dissertation. Therefore, the rest of this section is dedicated to explaining the structure and developing the foundation of PCA and Laplacian Eigenmaps as linear and non-linear methods, respectively.

1.2.1   Linear Methods

Linear manifold learning methods are based upon the assumption that the high dimensional manifold is projected onto a low dimensional manifold by using a linear transformation [42]. Two advantages of linear methods are that they are easy to compute and they provide an explicit mapping from the high dimensional manifold to the low dimensional embedding. The latter property states that new incoming data points can be projected to the same low dimensional space without going through the embedding procedure from the beginning. On the other hand, they are not capable of revealing nonlinear relationships between data points [18].

Although various linear manifold learning algorithms exist, indeed PCA is the most popular one [20]. In this section, we intend to describe and develop the necessary background

for PCA as it has been employed in the following chapters.

Consider a set $x_1, ..., x_n \in \mathcal{M}$ of n sampled data points from the manifold $\mathcal{M}$ with $x_i \in \mathbb{R}^l$ and we are looking for a set of corresponding points $y_1, ..., y_n$ in $\mathbb{R}^m (m \ll l)$ that preserves the relevant information of $\mathbf{X} = [x_1, x_2, ..., x_n]$. PCA is a dimensionality reduction technique that provides the directions (vectors) along which the data has maximum variance by considering the relative importance of these directions [43]. PCA is based upon the assumption that the variance of a dataset is a measure of the amount of information stored in the dataset. The higher the variance, the more information stored. Hence, PCA looks for a transformation $\mathbf{M}$ that preserves maximum variance [42].

Given the matrix $\mathbf{X} \in \mathbb{R}^{l \times n}$ whose columns are l-dimensional data points, we define matrix $\mathbf{X}_c \in \mathbb{R}^{l \times n}$ as the matrix of *centered* data points. Then the projections are given by $\mathbf{Y} = \mathbf{M}\mathbf{X}_c$, which is an $(m \times n)$ matrix. PCA optimizes the objective function:

$$\max_M \text{ var}(\mathbf{M}\mathbf{X}_c) \tag{1.1}$$

where $\mathbf{M}$ is an orthogonal $(m \times l)$ matrix. The covariance matrix of $\mathbf{Y}$ is:

$$\begin{aligned} \Sigma_y &= E[\mathbf{Y}\mathbf{Y}^T] = \frac{1}{n}\mathbf{Y}\mathbf{Y}^T = \frac{1}{n}(\mathbf{M}\mathbf{X}_c)(\mathbf{M}\mathbf{X}_c)^T \\ &= \frac{1}{n}\mathbf{M}\mathbf{X}_c\mathbf{X}_c^T\mathbf{M}^T = \mathbf{M}\frac{\mathbf{X}_c\mathbf{X}_c^T}{n}\mathbf{M}^T \\ &= \mathbf{M}\mathbf{V}\mathbf{M}^T \end{aligned} \tag{1.2}$$

where $\mathbf{V} = \frac{1}{n}\mathbf{X}_c\mathbf{X}_c^T = cov(\mathbf{X}_c)$ is the covariance matrix of centered data points. From equation (1.2) it is clear that $\mathbf{M}$ attempts to maximize the cost function $trace(\mathbf{M}\mathbf{V}\mathbf{M}^T)$. By adding a constraint $\|\mathbf{M}\|^2 = \mathbf{M}\mathbf{M}^T = \mathbf{I}$, we make sure that the transformation matrix is orthonormal.

Using the *Lagrange multiplier technique*, the maximization problem simplifies to solving

the eigenvalue problem of the covariance matrix $\mathbf{V}$:

$$\mathbf{MV} = \lambda\mathbf{M} \tag{1.3}$$

According to linear algebra, any matrix in the form of $\mathbf{AA}^T$ is symmetric. Therefore, $\mathbf{V}$ is symmetric which yields to the conclusion that, according to the *spectral theorem*, $\mathbf{V}$ is orthogonally diagonalizable and has only *real non-negative eigenvalues*. The orthogonally diagonalizable matrix $\mathbf{V}$ has *orthogonal non-zero real eigenvectors*.

It is proven that an $(l \times l)$ matrix $\mathbf{V}$ has l eigenvalues and l eigenvectors of size $(l \times 1)$. Let $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_l \geq 0$ (in decreasing order) with corresponding orthonormal eigenvectors $\vec{\mathbf{m}}_1, \vec{\mathbf{m}}_2, ..., \vec{\mathbf{m}}_l$. The eigenvectors of matrix $\mathbf{V}$ are *principal components* of data points, along which data points have maximum variance. Principal components of a sample synthetic dataset are illustrated in Figure 1.3. These principal components make a set of normal basis vectors that projects high dimensional data points into a low dimensional space while preserving most of the information. The desired linear transformation matrix $\mathbf{M}$ is made up of the first m principle components associated with the first m largest eigenvalues.
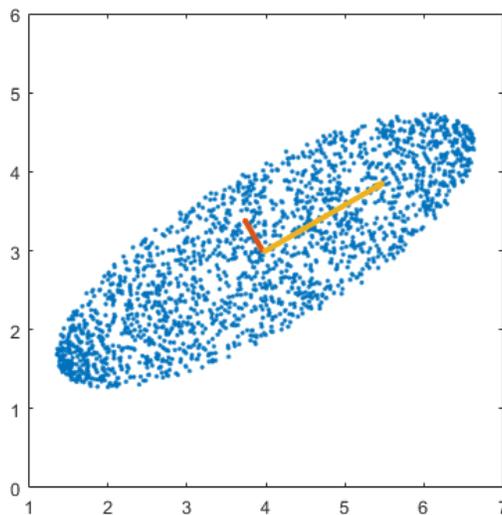
Figure 1.3: Principal components of a sample dataset.

1.2.2 Nonlinear Methods

Nonlinear dimensionality reduction is a problem of finding a meaningful low dimensional representation for a complex nonlinear high-dimensional data when linear methods fail to do so. In the literature, the term *manifold learning* usually refers to *nonlinear manifold learning* techniques. Hence, from now on in this thesis, manifold learning indicates nonlinear manifold learning.

Across different manifold learning methods, Laplacian Eigenmaps is relatively computationally efficient with well-established computational structure [54]. It is a spectral analysis technique which is based upon graph Laplacian and Laplace-Beltrami (LB) operator on the manifold and in close connection to the heat flow [56]. Accordingly, we have employed Laplacian Eigenmaps in the design and development of our algorithms. In the following, an introduction into manifold learning and the mathematical foundations of Laplacian Eigenmaps algorithm is provided.

Manifold learning, more specifically Laplacian Eigenmaps, deals with the following problem: consider a set $x_1, ..., x_n \in \mathcal{M}$ of n points on manifold $\mathcal{M}$ embedded in $\mathbb{R}^l$, and we are looking for a set of corresponding points $y_1, ..., y_n$ in $\mathbb{R}^m (m \ll l)$ as a structural representation which also benefits from *locality preserving* and *structural equivalence* properties. Locality preserving means structural representations of two similar patches of the same dataset are similar after being mapped to the new coordinates. On the other hand, structural equivalence is the case when structural representations of two similar patches of different datasets are similar. It is shown that Laplacian Eigenmap satisfies both properties [54].

Graph Laplacian is constructed over an *undirected weighted graph* $G = (V, E)$ with a set of points $x_i \in V$ and a set E of edges that connects nearby points with. The graph G is called the *adjacency graph* as it specifies connectivities between adjacent points. The weight w_{ij} between two neighbor points x_i and x_j determines the power of the connection.

With a given set of points in the high dimensional space, there are two ways to construct the adjacency graph. (1) Two points i and j from a set V are considered as neighbors if they

are in less than ϵ distance from each other. That can be formulated as $\|x_i - x_j\|^2 < \epsilon$ where the norm is the usual Euclidean norm in $\mathbb{R}^l$. This method which is called ϵ-*neighborhood* makes a symmetric graph, however, tends to construct a graph with several connected components. Besides, choosing the parameter ϵ would be challenging. (2) Alternatively, the graph is constructed by finding k-nearest neighbors of each point in the high dimensional space and specifying an edge between the point and each of its' neighbors. This method, which is called k-*nearest neighbor* (parameter $k \in \mathbb{N}$), tends to construct a directed graph, which is when node i belongs to the set of k-nearest neighbors of node j but, node j is not among k-nearest neighbors of node i [34].

In the mathematical theory behind the Laplacian Eigenmap method, it is assumed that the graph is undirected/symmetric. Therefore, when adjacency graph is constructed using k-nearest neighbors, an extra step is required in which the graph is forced to become symmetric by either adding more edges wherever there is a one-way edge (called *symmetric kNN graph*) or deleting one-way edges (called *mutual kNN graph*)[57]. Either way, after this step, the number of neighbors of each point might not be the fixed value k.

In [34] the problem of finding a low dimensional data set that preserves local information is simplified to the optimization problem of minimizing the following function:

$$\sum_{ij}(y_i - y_j)^2 w_{ij}. \tag{1.4}$$

Indeed, the minimization problem ensures that two neighboring points x_i and x_j stay close after being mapped to y_i and y_j. It is shown that for any $\mathbf{y}$:

$$\sum_{ij}(y_i - y_j)^2 w_{ij} = 2\mathbf{y}^T \mathbf{L} \mathbf{y} \tag{1.5}$$

where $\mathbf{L} = \mathbf{D} - \mathbf{W}$, so called *Laplacian matrix*, $\mathbf{W}$ is a symmetric *weight matrix* with entries w_{ij}, and $\mathbf{D}$ is a diagonal matrix, so called *degree matrix*, with entries $d_{ii} = \sum_j w_{ij}$. The entries of the degree matrix $\mathbf{D}$ are column (or row, since $\mathbf{W}$ is symmetric) sums of $\mathbf{W}$. If

the set V consists n points, then $\mathbf{W}$ and therefore $\mathbf{D}$ and $\mathbf{L}$ are all sparse $n \times n$ symmetric matrices.

By adding the orthogonality constraint $\mathbf{y}^T \mathbf{D1} = 0$ in order to eliminate the trivial solution and the constraint $\mathbf{y}^T \mathbf{Dy} = 1$ for removing an arbitrary scaling factor in the embedding, the minimization problem of equation (1.4) simplifies to:

$$\underset{\substack{\mathbf{y}^T\mathbf{Dy}=1 \\ \mathbf{y}^T\mathbf{D1}=0}}{\arg\min} \ \mathbf{y}^T \mathbf{Ly}. \tag{1.6}$$

The solution vector $\mathbf{y}$ to equation (1.6) is obtained by the minimum eigenvalue solution to the generalized eigenvalue problem [34]:

$$\mathbf{Ly} = \lambda \mathbf{Dy}. \tag{1.7}$$

The equation (1.7) is called *eigen-decomposition* or *spectral decomposition* of LB operator, whereby the matrix is represented in terms of its eigenvalues and eigenvectors. The set of eigenvalues is called the *spectrum* of $\mathbf{L}$ [58]. The most important property of the eigenvalues and eigenvectors of LB operator is that they are isometric invariant. In other words, if the manifold is not stretched (e.g., bent into extra dimension), the spectral values will not change. Hence, two manifolds with different orientation should have similar spectral representations provided by the underlying topology of two mesh graphs [59].

Regarding the weighting scheme of the graph Laplacian, in [34], it is discussed that graph Laplacian is comparable to the LB operator on manifolds and hence, desirable embedding can be obtained by finding the eigenfunctions of LB operator. For the best approximation, it is recommended to use the following *heat kernel weighting scheme*:

$$w_{ij} = \begin{cases} e^{\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)} & \text{if } (i,j) \in E \\ 0 & \text{otherwise} \end{cases} \tag{1.8}$$

where $\sigma \in \mathbb{R}$. In the simplest case, so called *simple minded weighting scheme*, ($\sigma = \inf$) and therefore, $w_{ij} = 1$ if two nodes i and j are connected with an edge, and $w_{ij} = 0$ otherwise.

From equations (1.5) and (1.8) it is clear that $\mathbf{L}$ is real symmetric and positive semi-definite. All eigenvalues of a real symmetric matrix are real, and eigenvectors corresponding to distinct eigenvalues are orthogonal. As $\mathbf{L}$ is positive semidefinite, all eigenvalues are non-negative ($\lambda_i \geq 0$ for all i) [60]. Every row sum and column sum of $\mathbf{L}$ is zero. In consequence, eigenfunction $\mathbf{1}$ and eigenvalue $\lambda = 0$ are trivial solutions to equation (1.7). The multiplicity of eigenvalue zero is associated with the number of *connected components* of the graph [56]. For a graph with multiple connected components, $\mathbf{L}$ is a block diagonal matrix, where each block represents the Laplacian matrix of a connected component, possibly after reordering vertices [61].

The Laplacian Eigenmaps algorithm suggests to leave out all eigenvectors corresponding to eigenvalues equal to zero and use the next m eigenvectors with the smallest non-zero eigenvalues for embedding the manifold in low m-dimensional Euclidean space [34]. The embedded manifold will be in the form of an $n \times m$ matrix $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_m]$ where the ith row demonstrates embedding coordinates of the ith point.

There are many different algorithms available for finding m-smallest eigenvalues of the generalized eigenvalue problem. Among them are the Householder method [62, 63], the Q.R. method [63], subspace iteration [64, 65], etc. Many of these algorithms are inefficient when applied to very large structural systems [66]. The most efficient algorithms are implemented in ARPACK software. ARPACK which stands for ARnoldi PACKage is very powerful in approximating a few eigenvalues and corresponding eigenvectors. As the matrix $\mathbf{L}$ is symmetric, the Arnoldi process is reduced to a variant of the Lanczos process called the Implicitly Restarted Lanczos Method (IRLM) [67]. Computational approximations performed by eigen-solvers make the smallest eigenvalue not to be exactly equal to zero. Knowing the number of connected components in $\mathbf{L}$ helps with discarding as many eigenvectors as should be discarded.

To conclude, Laplacian Eigenmap algorithm is a local nonlinear dimensionality reduction as it considers the small neighborhood of each point. While preserving local information, it extracts structural information from the high dimensional manifold and embeds it in a low dimensional space. The embedded manifold is called *structural representation* [54].

1.3 Summary of the Thesis Work

The current chapter introduces the importance of medical images among other types of clinical information (e.g., EHRs, clinical notes), and highlights the high dimensionality of images available to the healthcare community. There are several computational approaches to deal with high dimensional space, in which many of them try to reduce the dimensionality while keeping most valuable information. Manifold learning is the area of studying all these algorithms for dimensionality reduction. This chapter briefly covered the learning strategy of manifold learning along with technical details of PCA and Laplacian Eigenmaps. The following chapters explore applications of manifold learning in medical image and shape analysis.

Chapter 2 introduces a multi-modal image registration algorithm in which a manifold learning-based intensity transformation maps reference and sensed images to a unified intensity system. The proposed method facilitates recovering strong scales, rotations, and translations while having application in registering multi-modal images with either full or partial data.

Chapter 3 introduces a global shape descriptor that benefits from the capability of manifold learning in extracting relevant features from a smooth manifold along with unique properties of Laplacian Eigenmap, namely, locality preservation and structural equivalence. It also presents a straightforward and efficient approach for scale normalization than can only be used for applications in which the scale does not characterize the object. The proposed descriptor intended to eliminate the limitations tied to the existing spectral descriptors,

namely dependency on triangular mesh representation and high intra-class quality of 3D models. It has a wide variety of applications such as shape classification and retrieval for a broad range of objects from molecules to planets.

Chapter 4 introduces a descriptive model for accurate and reliable diagnosis of pulmonary nodules. The model benefits from the proposed shape descriptor developed in Chapter 3 combined with discriminating features extracted automatically using a point cloud-based deep learning network. The innovative combination of spectral features with a PC-based deep network presents high sensitivity on detecting malignant nodules aiming to improve patients survival rate, while, showing high specificity that reduces unnecessary financial and emotional costs due to misclassification of benign nodules.

Chapter 5 summarizes the main findings of my dissertation and points to future areas of research.

Chapter 2

2D Image Analysis: Manifold Learning Approach

2.1 Introduction

Medical image analysis is an active research area that develops computational and mathematical methods in order to extract clinically relevant information. Medical images might have been acquired from different modalities based on the type of desired information such as Ultrasound, Computed Tomography (CT), X-ray, and Magnetic Resonance Imaging (MRI) for anatomical information alongside functional imaging modalities like Positron Emission Tomography (PET), Single Photon Emission Computed Tomography (SPECT), and Magnetic Resonance Spectroscopy (MRS). Medical image analysis focuses on the development and implementation of algorithms and strategies to solve different types of problems such as visualization, image segmentation, image fusion, image registration, and others.

Medical image analysis has a wide range of applications in research and in the clinical setting such as intra-operative studies, surgery planning, image-guided surgery, and monitoring growth/healing process of tumors, to name a few. In some applications, the information we are looking for can be achieved by comparing or fusing two or more images. As misalign-

ment between images is inevitable, image registration plays a critical role in the accuracy of information obtained [68].

Medical image registration is one of the key image processing techniques in medical image analysis that has attracted great attention in the past few decades. Image registration is the process of estimating the spatial transformation such that it overlays two or more images. Images representing the same Region-of-Interest (ROI) are acquired at different times, possibly with different sensors or from different viewpoints [69].

Generally speaking, image registration involves (at least) two images: *reference* and *sensed* [70]. Image registration aims to align the sensed image with the reference image [71]. Any registration technique is an ill-posed inverse problem [72] containing three components: 1) a transformation model, which maps the sensed image to the reference image; 2) a (dis)similarity measure, which determines the relative merit between reference and sensed image; and 3) an optimization method, which seeks parameters that lead to the optimal transformation [68, 72]. From a mathematical perspective, image registration is an optimization problem:

$$\arg\min_{\theta} \mathcal{M}(R, S \circ \mathcal{T}(\theta)) + \mathcal{R}(\mathcal{T}(\theta))$$

where $\mathcal{M}$ quantifies the level of alignment, R and S represent reference and sensed image respectively, $\mathcal{T}(\theta)$ is transformation parameterized by θ, and $\mathcal{R}$ is the regularization term which accounts for the ill-posedness of the problem [72].

Medical image registration has been studied extensively, and numerous techniques have been developed over the past three decades. Several surveys are available for interested readers [68, 71–77]. Despite all the efforts that have been used to resolve the problem of image registration, there is no unique technique that works in all circumstances. Considering the importance of image registration in the accuracy of further analysis, and due to the progress made in the past decade in the field of computational techniques, medical image registration is still a field of interest.

The existing challenges in the developing a single generic algorithm for the problem

of medical image registration can be summarized as: 1) the nature of transformation can vary widely and be highly non-rigid; 2) images acquired from different modalities present different structures of the underlying organ; 3) missing/partial data causes the loss of one-to-one correspondence between the images; 4) there is no unique similarity measure for every problem; 5) some optimization algorithms may trap in local extremum points; and, 6) each imaging modality introduces its unique challenges [78, 79].

The most widely-used taxonomy proposed by Maintz and Viergever [75] categorizes registration methods according to nine different criteria including dimensionality, nature of registration basis, nature of transformation, domain of transformation, interaction, optimization procedure, modalities involved, subject, and object were incorporated. Concerning modalities involved in acquiring images, registration methods are categorized to *mono-modal* and *multi-modal registration*. Unlike mono-modal registration in which images are captured using the same imaging technique, in multi-modal medical image registration, as is the focus of the current work, images from different imaging modalities are aligned and registered in one coordinate system, using rigid or deformable techniques [77]. Multi-modal image registration is usually necessary as it provides additional information due to the intrinsic differences between different imaging modalities and the data that they extract from the organs. For example, CT and MRI provide high-resolution anatomical representations with limited physiological information, while SPECT and PET provide physiological data with reduced resolution in comparison to the previous ones.

In a complementary taxonomy [80], based on the amount of overlap between images, registration methods are categorized in two groups; methods that work in the presence of: 1) full image data; 2) both full and partial image data. Image registration with partial overlap is the situation in which, only a portion of one image is captured in the other image. In such a situation, local convergence and lack of one-to-one correspondence are two factors that make it more challenging. Moreover, when it comes to multi-modal registration, the significant variations in the intensity levels of the images demand extra attention [79, 81, 82].

Generally speaking, registration methods proposed in the literature for the problem of multi-modal image correspondence can be divided into two major categories: 1) information theoretic based approaches; and 2) multi-modal to mono-modal reduction based approaches [72]. In the first class, methods based on Mutual Information (MI) and its variants (e.g., Normalized MI (NMI) and Regional MI (RMI)) [83–86], Correlation Ratio (CR) [87], Kullback-Leibler Divergence (KLD) [88], Jensen-Shannon Divergence (JSD) [89] and Correlation Coefficient (CCoef) [90] are great examples of the use of information theoretic approaches for the problem of multi-modal image registration.

In the second class, the problem of multi-modal registration is reduced to a mono-modal registration problem. The reduction can be made by simulating the results of one modality from the other. Examples of this can be found in the works of Roche et al. [91] and Wein et al. [92] for multi-modal registration of ultrasound-MRI and ultrasound-CT data, respectively. In another approach, which attracted more attention in recent years, a proper mapping to bring the two sensor data into a third common modality for registration is desirable. Several researchers have devised approaches that try to incorporate intensity gradients as a means for a more robust and accurate registration. In [93], a combination of morphological operators (opening and closing) for edge information extraction with cross-correlation is used for rigid registration. In [94], the common modality is created by use of normalized intensity gradients as a representative of the anatomical structure of the organs in the images. In another work, Butz and Thiran [95] employed local edge variance in combination with mutual information for multi-modal image registration. The usefulness of Gabor filter banks for describing edges and texture has also been investigated for deformable image registration [96]. The use of neighborhood descriptors is further investigated in the work of Heinrich et al. [97]. Authors inspired by the idea of using non-local means for image denoising, devised a Modality Independent Neighborhood Descriptor (MIND) for cross-modality registration.

In recent years, research in multi-modal to mono-modal reduction class took a new direction by employing Manifold Learning (ML) techniques [19, 54, 98–102]. Manifold learning,

e.g., Laplacian Eigenmap in [54, 98], is used to generate a more general structural representation of the input images, which is less affected by the differences in the intensities. The representations should satisfy two conditions: 1) locality preservation: if two patches of the image are close according to some norm, their structural representation should also be close; and 2) structural equivalence: two patches from different images are equivalent if and only if their structural representations are the same. In [99], deformable brain image registration is performed in two steps. First, the manifolds are learned for individual brain regions separately, and region-specific deformations are computed. Then, the whole brain images are registered using the global manifold while preserving the region-specific deformations. In [101], T1 and T2 MRI images are registered using manifold-based feature point matching. In the proposed technique, after transforming the images to feature domain images, Scale Invariant Feature Transform (SIFT) is used for feature extraction followed by Best-Bin-First (BBF) method for feature matching [103]. In a similar approach, a higher accuracy was obtained compared to using conventional MI for registration of T1, T2, and PD MRI images [102]. The proposed method transformed input images into a feature domain using diffusion maps.

In general, multi-modal to mono-modal transformation can be interpreted as intensity transformation in which two images from different modalities use comparable relative intensity levels to make the necessary contrast in the image. In the current contribution, we are inspired by the work done by Wachinger et al. [54]. For two primary reasons, namely using large patch sizes and inefficient manifold alignment, the method proposed in [54] results in a vague structural representation of input images. Poor representation affects the performance of the method for registering images with delicate structures. Also, it fails in registering images with partial data. To address these issues, we propose an intensity transformation method that allows us to deal with the problem of multi-modal image registration with full or partial data by employing advanced mono-modal registration methods. Below, we summarize the **main contributions** of the current work:

18

1. The current work, facilitates multi-modal image registration through an efficient intensity transformation by improving the technique presented in [54]. We elaborate on details of the algorithm, the choice of parameters, and aspects of the result that would be affected. Over and above that, we introduce a simple, efficient, and straightforward manifold alignment technique.

2. To evaluate the performance of the proposed method, extensive experiments, using human brain MRI as well as CT scans, were carried out on real-world (patient) data in addition to simulated data containing five degrees of freedom in 2D space; translation, rotation, and scale.

3. We claim and experimentally prove that image registration with partial data can also benefit from the proposed method. Therefore, the proposed method can be utilized in the preprocessing step of any multi-modal image registration.

The organization of this chapter is as follows. In Section 2.2, we introduce the proposed multi-modal to mono-modal transformation method along with technical background and discussion on the choice of parameters. The effectiveness of the proposed method on multi-modal image registration with full as well as partial data is examined by performing several qualitative and quantitative experiments using widely used datasets in Section 2.3. Finally, in Section 2.4, we conclude the work and experimental results in more details.

2.2 Proposed Method

In the current contribution, we propose to bring reference and sensed images into a same co-ordinate system that represents the internal structure of both images. As internal structures across images from different modalities are similar, the process of finding similarities between images and eventually, finding a transformation that aligns one image with the other one will be less challenging and more promising.

The Laplacian Eigenmaps guarantees the preservation of locality, meaning that if two patches of an image are close according to some norm, their structural representation should also be close. In other words, after embedding in low dimensional space, the embedded manifold, called the *structural representation* of the image, is similar to the original image while it highlights different features.
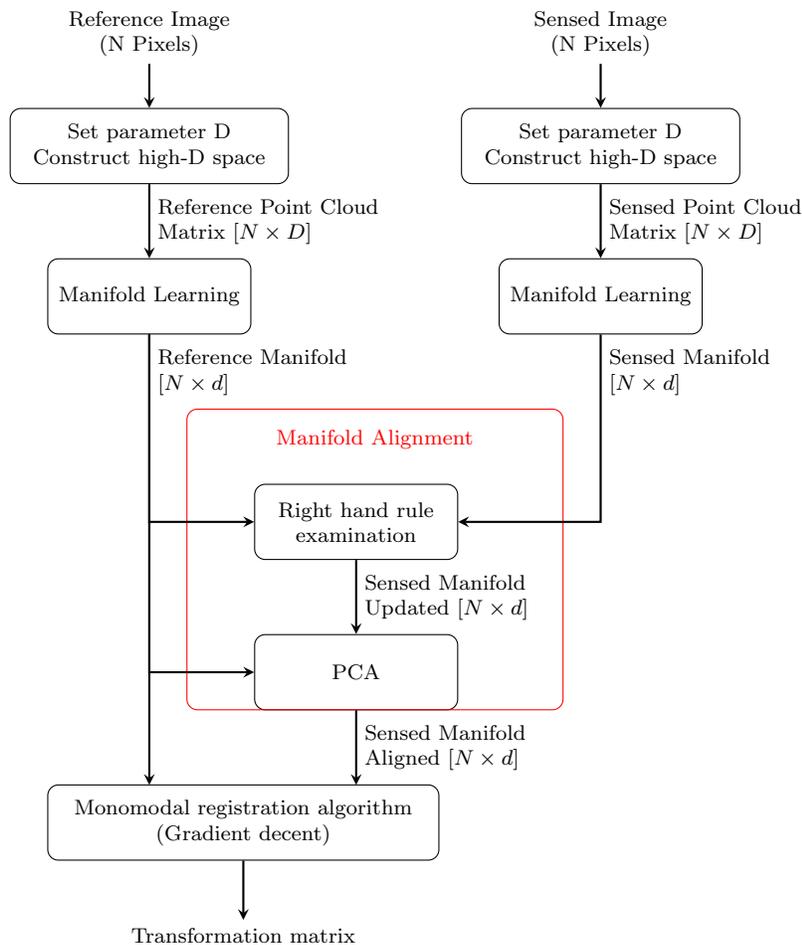
Figure 2.1: Pipeline of the proposed technique.

The pipeline of the proposed algorithm is illustrated in Figure 2.1. The idea is to construct a high dimensional manifold by extracting more information than just the intensity of each pixel. After that, the structural representation of each input image is computed using the Laplacian Eigenmap technique. Laplacian Eigenmap guarantees locality preservation and structural equivalence. Meaning that each embedded manifold is similar to the original

manifold, and as they represent internal structure of the same organ, they are structurally equivalent. However, it does not guarantee alignment of embedded manifolds. Consequently, a manifold alignment step is required [98]. In the proposed algorithm, rigid transformation using principal components is suggested. Sometimes the mapping between two embeddings requires a reflection. A simple step in the pipeline, which from now on we call it the *right-hand rule* step, handles that issue.

In summary, manifold learning computes the structural representation of the reference and sensed images, and manifold alignment combined with right-hand rule aligns embedded manifolds. The result is a transformation that justifies the intensity differences between images captured from different modalities. The *intensity transformation*, or *intensity registration*, leads us to direct application of L1 and L2 norm in image registration which has been studied well in the community and improves the accuracy of the registration process. In the following, the proposed idea is explained in details.

2.2.1 Constructing High-Dimensional Space

Unlike problems in which manifold learning has application in population studies, where each point represents an image in the high dimensional space, in the image registration problem only two images, reference and sensed, are present. The point cloud in the high dimensional space is constructed by taking patches from each image. In the current problem, each dimension of the high dimensional space is dedicated to one image pixel of the patches. Patches consist up of an $s \times s$ neighborhood for each image pixel. Considering each patch is moved one pixel from the previous patch, for an image of size $n_r \times n_c$, the high dimensional manifold includes $n_r \times n_c = N$ points in an $s^2 = D$ dimension space. To be consistent with the description of manifold learning in section 1.2.2, the reader can clearly see that $n = N$ and $l = D$. Such a set of points can be represented in matrix format using the coordinates of points (Figure 2.2). For a gray-scale image of size N pixels with patches of size D pixels, the matrix size is $[N \times D]$.
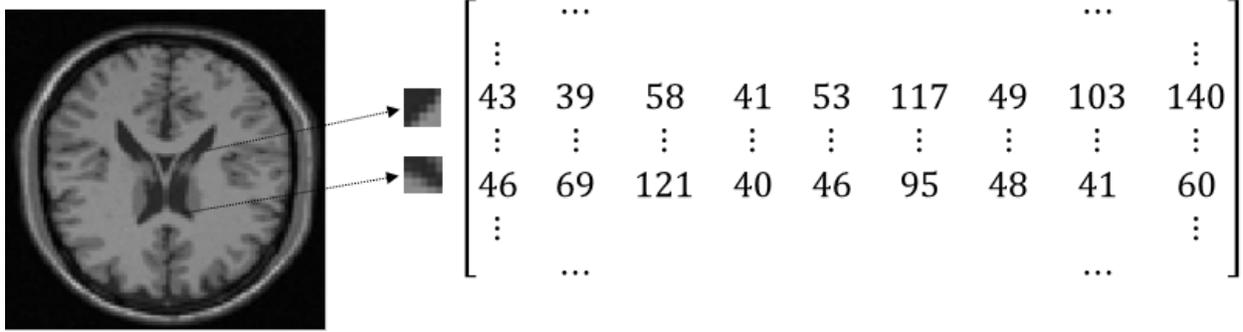
Figure 2.2: Matrix representation of point cloud in high dimensional space; For an image of size 200×200 and patch size of 3×3, the matrix is $40,000 \times 9$.

The choice of D depends on the application and size of the image. The embedded manifold, which resembles the original image, contains details of the original image when D is chosen relatively small. However, choosing a large value for D results in a blurred version of the image. In Figure 2.3 (a), a slice of a T2-MRI image is shown. Laplacian Eigenmap with several D values is used to study the image. Figures 2.3 ((b)–(d)) shows the effect of the selected value for parameter D on the sharpness of the rendered image after embedding.
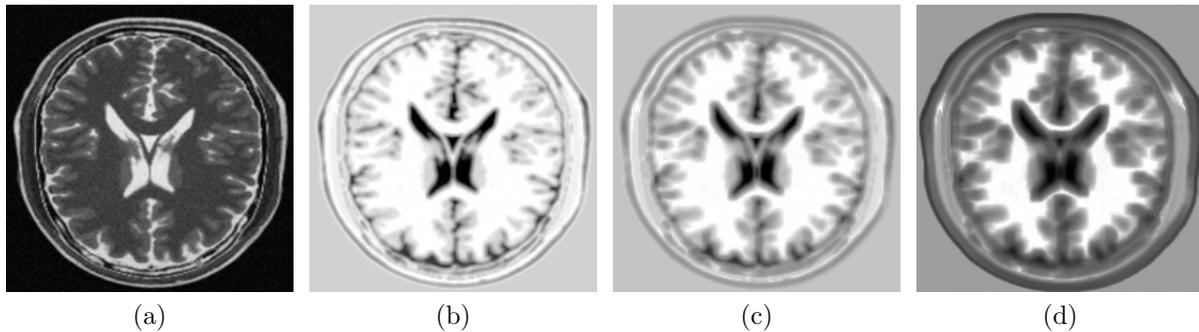
Figure 2.3: (a) A slice of a T2-MRI and the image representation of its' embedded manifold with: (b) $D = 9$; (c) $D = 25$; and (d) $D = 49$.

2.2.2 Dimensionality Reduction using Laplacian Eigenmaps

The Laplacian Eigenmap algorithm consists of several steps and determining multiple parameters. The registration accuracy depends on the correct choice of parameters. The implemented algorithm according to the steps introduced in [34] is summarized in Figure 2.4

and explained in more details below.

1. Constructing the adjacency graph

   - Compute the distance between every two data points;
   - Put an edge between every two data points by considering:
     – k-nearest neighbor with choice of k;
     – Symmetric kNN graph.

2. Choosing the weights

   - Assign weight to each edge between every two neighbor points using:
     – Heat kernel with choice of bandwidth.

3. Eigenmaps

   - Construct Laplacian Matrix;
     – Real symmetric sparse $[N \times N]$ matrix.
   - Compute number of connected components;
     – Number of connected components determines the number of eigenvalues equal to zero.
   - Solve generalized eigenvalue problem $\mathbf{Lv} = \lambda \mathbf{Dv}$;
     – Leaving out all eigenvectors corresponding to eigenvalues equal to zero.
   - Take the first d-eigenvectors corresponding to the smallest eigenvalues greater than zero as basis for embedding manifold in the d-dimensional space.

Figure 2.4: Summary of dimensionality reduction with Laplacian Eigenmaps

The first step in dimensionality reduction with Laplacian Eigenmaps requires constructing the *adjacency graph*. The point cloud acquired in the $\mathbb{R}^D$ space makes up graph nodes. It is worth noting that in this context, we use terms point and node interchangeably when we are referring to the high-dimensional space and the graph respectively. We form the set of edges E that specifies connections of points using the k-nearest neighbor followed by symmetric kNN graph technique. In other words, two points i and j are neighbors if one of them is in k-nearest neighbors of the other one.

In general, to extract the structure of the image both approaches for computing the set of edges (k-nearest neighbor and ϵ-neighborhood) work. Since choosing the parameter k is considerably less challenging than determining the value of ϵ, we chose the k-nearest neighbor approach. Now, a question arises as to what the appropriate value for the number of neighbors 'k' is? Strictly speaking, what are the effects of choosing either a small or large value for the parameter k in the results? The parameter k determines the sparsity of the

Laplacian matrix that we will later compute from the adjacency graph. A large value of k leads to a less sparse Laplacian matrix which is more memory demanding and time invasive. Besides, it is clear that finding more number of neighbors for each point takes more time. On the other hand, picking a small value for k may result in a graph consisting of several connected components. Also, the information contained in the resulting graph may be so small that it fails to reveal the internal structure of the image and the eigenvalue problem fails to converge. Therefore, there is a tradeoff between memory/speed and assuring to have a single connected component with a Laplacian matrix that will converge while solving the eigenvalue problem. There is no rule for choosing a value for the parameter k. Tentatively, $k = 20$ works for most medical head scans in size around 200×200 pixels. Symmetrizing the adjacency graph using symmetric kNN method has the advantage of keeping as many edges as possible and not breaking the graph into smaller connected components. Even with these considerations, the graph may still consist of multi-connected components. In this case, we might increase k as far as the processing time and required memory allow us.

The second step is assigning a weight to each and every edge of the graph. For better capturing local features it is more desirable to specify higher weights to edges between closer nodes. Therefore, we prefer the exponential weighting scheme offered by the heat kernel. For a set $\{x_i | 1 \leq i \leq N\}$ of N points in $\mathbb{R}^D$, edge weights are computed as follows:

$$w_{ij} = e^{\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)}; \qquad \text{where } \sigma \in \mathbb{R} . \tag{2.1}$$

Selecting a value for parameter σ is relatively challenging and data-dependent [104]. When σ is relatively large compared to $\|x_i - x_j\|$, the weight of connections is close to one which results in an un-weighted adjacency graph. On the contrary, a relatively small σ leads to non-significant edges which may later cause failure in the convergence of the Eigenvector problem. The right choice for σ lies in between these two extremes [104].

In [105], it is suggested to construct the σ-dependent weight matrix $\mathbf{W}(\sigma^2)$ and compute

$S(\sigma^2) = \sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij}(\sigma^2)$ for several values of σ. Then plot S-curve versus (σ^2) in semi-logarithmic scale. The S-curve has two asymptotes at $\sigma = 0$ and $\sigma = inf$ and can be used for choosing a proper value for standard deviation. Initially, we chose σ^2 at a point around the upper two-thirds of the total range of the function $S(\sigma^2)$. This idea adds extra processing time, for each image to compute the logarithmic curve. In order to reduce the processing time, we offer to select the variance equal to the maximum squared distance of all edges. In other words:

$$\sigma^2 = \max_{all\, x_i, x_j \in V}(\|x_i - x_j\|^2). \tag{2.2}$$

In equation (2.2), the lower limit of edge weight is $e^{-1/2} \approx 0.606$ and obviously, the upper limit is 1. Empirically, we found out that the results did not change much and were still reliable.

The last step is computing the *Laplacian matrix* and solving the generalized eigenvalue problem. Knowing the weight matrix $\mathbf{W}$ of the adjacency graph, the Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{W}$ is computed with no difficulty. At this stage, it is beneficial to compute the number of connected components that make the Laplacian graph. Because, it is equal to the number of trivial solutions $\lambda = 0$, which have to be left out from the eigen-decomposition of LB operator. As explained in section 1.2.2, the Laplacian matrix is a real, sparse, symmetric and positive semi-definite matrix [34]. Therefore, the generalized eigenvalue problem of equation (1.7) has real and positive eigenvalues. Besides, the eigenvectors corresponding to distinct eigenvalues are orthogonal [60]. The first d-eigenvectors corresponding to the smallest non-zero eigenvalues are orthogonal bases of new low dimensional space, and the resulting manifold in the d-dimensional space is called *the embedded manifold.*

Reordering each eigenvector to the size of the original image $(n_r \times n_c)$ resembles the original image while highlighting one of the layers of the image structure. Since each rendered image is associated with one eigenvalue of the spectrum of the image, in the following it is shortly referred to as the *feature image* [101]. As the corresponding eigenvalue increases the feature image highlights smaller details of the image. Whereas the input to the Laplacian

Eigenmap is an $[N \times D]$ matrix, the output is an $[N \times d]$ matrix. In the current work, we only need the first $d = 3$ eigenvectors.

2.2.3   Manifold Alignment

The Lanczos algorithm starts with a random initial vector. Therefore, the embedding in low dimensional space is arbitrary. In other words, along with preserving the locality, manifolds of different modalities would be embedded in different parts of the low dimensional space. Consequently, a manifold alignment step is required [98]. Several linear and non-linear manifold alignment techniques have been developed over the past decade [106–109]. A linear manifold alignment technique, namely Principal Component Analysis (PCA), provides the principal directions of reference and sensed embedded manifolds. Rigid registration of embedded manifolds is fulfilled by finding the translation and rotation transformation using the principal directions.

Rigid registration of point clouds (embedded manifolds) in the low dimensional space is adequate for some applications, e.g., population studies, in which the manifold conversion over the entire population is investigated. Nevertheless, for image registration, we need to investigate whether intensity inversion of feature images is needed, caused by the fact that inverse of a principal direction is a valid principal direction and cannot be compensated using rigid registration. In the current contribution, we suggest using PCA-based alignment combined with computing the covariance between feature images of the reference and sensed. In this work, we refer to this step as *right-hand rule* examination.

As Figure 2.5 shows, in the right-hand rule examination step, we aim to inspect whether the principal directions of both manifolds follow the same right- or left-hand rule; which is finding whether the direction of the cross product of the first two vectors is the same as the direction of the third vector. If principal directions of both manifolds follow the same rule, PCA is capable of aligning both manifolds without the need for inverting intensities of a feature image; otherwise, reversing one of the vectors compels them following the same rule.
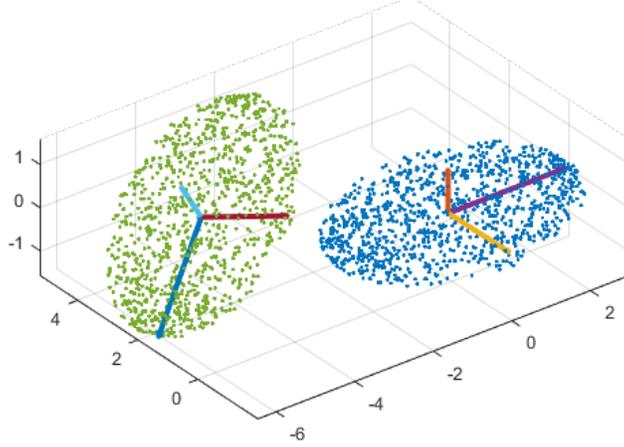
Figure 2.5: Right-hand rule inspection of two embedded manifolds using 3D synthetic data. The first three principal directions of the reference (blue) and the sensed (green) embeddings follow the left- and right-hand rule, respectively.

Right-hand rule compliance test requires computing the cross product of two vectors which is relatively easy in 3-dimensional space. However, it becomes more elusive in higher dimensional spaces. Therefore, we offer a more general test for any N-dimensional manifolds. The sign of the covariance matrix between two feature images indicates the direction of their association. If the covariance is negative, sensed feature image increases/decreases in the opposite direction to the reference feature image. Aligning every two corresponding feature images will eventually result in both embeddings follow the same right- or left-hand rule.

After compensating the reflection error, we employ PCA for the rigid transformation of embedded manifolds. Rigid registration using PCA is obtained by translating the center of both manifolds to the origin followed by rotating the principal directions of one manifold (*sensed*) to get aligned with the principal directions of the other manifold (*reference*) [110]. As discussed in details in section 1.2.1, PCA is based on the eigenvectors of the covariance matrix $\mathbf{M} = \mathbf{V}\mathbf{V}^T$. The largest eigenvector that is associated with the largest eigenvalue indicates the direction of the largest variance [110].

Let points $p_1, ..., p_N$ be the embedded manifold from the first three eigenvectors of the Laplacian matrix centered at location $\mathbf{c} = (c_x, c_y, c_z)^T$. The $\mathbf{V}_c$, matrix representation of a manifold centered at the origin, and the corresponding covariance matrix are defined as

27

follows:

$$\mathbf{V}_c = \begin{pmatrix} p_{1_x} - c_x, & p_{2_x} - c_x, & ..., & p_{N_x} - c_x \\ p_{1_y} - c_y, & p_{2_y} - c_y, & ..., & p_{N_y} - c_y \\ p_{1_z} - c_z, & p_{2_z} - c_z, & ..., & p_{N_z} - c_z \end{pmatrix} \qquad \text{and} \qquad \mathbf{M} = \mathbf{V}_c \mathbf{V}_c^T.$$

We solve for the eigenfunctions of the covariance matrix:

$$\mathbf{M}_s \mathbf{A} = \Lambda_s \mathbf{A}$$

$$\mathbf{M}_r \mathbf{B} = \Lambda_r \mathbf{B}$$

where $\mathbf{M}_r$, $\mathbf{M}_s$, $\mathbf{B}$, and $\mathbf{A}$ are covariance matrices and principal components of the reference and sensed centered manifolds, respectively. Then the rotation matrix $\mathbf{R}$ that aligns sensed along with reference satisfies the equation:

$$\mathbf{RA} = \mathbf{B} \ .$$

Since matrix $\mathbf{A}$ is orthonormal, the rotation matrix simplifies to:

$$\mathbf{R} = \mathbf{BA}^{-1} = \mathbf{BA}^T \ .$$

The PCA-based rigid registration of the sensed manifold to align with the reference can be formulated as follow:

$$\mathbf{V}'_s = \mathbf{R}\mathbf{V}_{sc} + \mathbf{c}_r \tag{2.3}$$

where $\mathbf{c}_r$ is the centroid coordination of the reference.

The objective of manifold alignment is to form a direct mapping between two embeddings that facilitates finding correspondences, and transferring knowledge from one domain to another. Here, coefficients from each dimension translates into an intensity of a feature

(a) T1-weighted MRI   (b) T2-weighted MRI

(c) Feature Image 1 of   (d) Feature Image 1 of   (e) Feature Image 1 of
T1-MRI   T2-MRI   T2-MRI after alignment
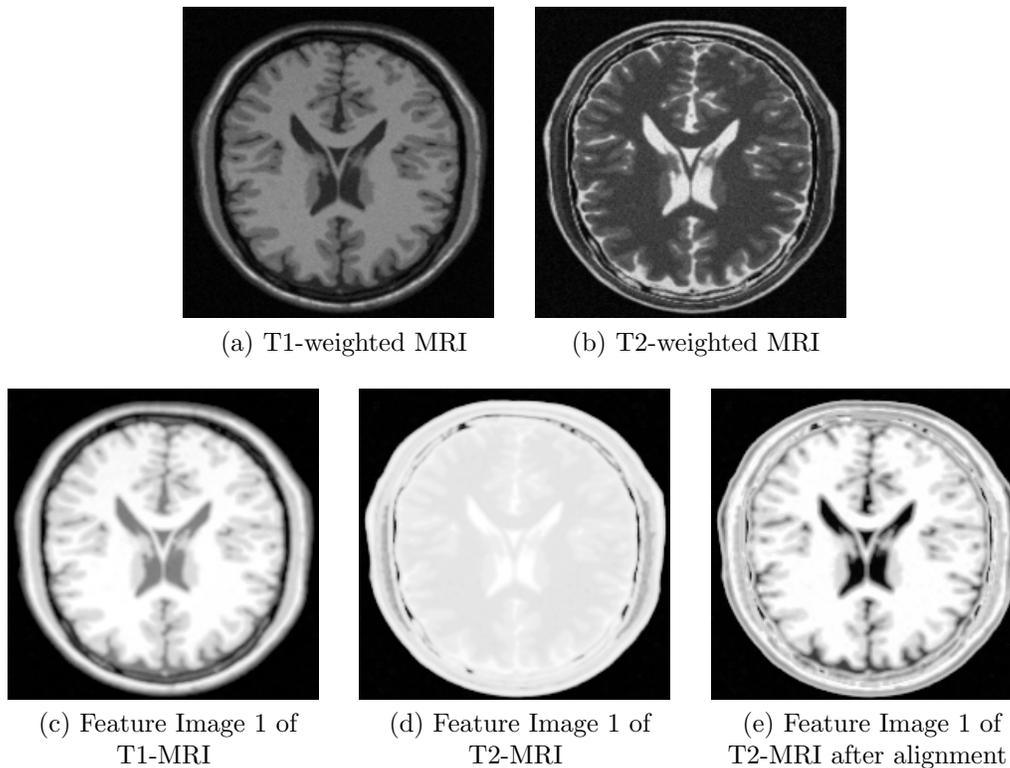
Figure 2.6: Intensity transformation of human brain T1- and T2-weighted MRI scans. Despite intensity variations between (a) and (b) as well as between (c) and (d), after manifold alignment, feature images (c) and (e) are comparable.

image, corresponding to the dimension. Hence, the mapping obtained from manifold alignment results in intensity transformation. That is to say intensity variations among different modalities will be removed, and feature images become comparable in terms of relative intensity. As shown in Figure 2.6, despite differences between T1- and T2- weighted MRIs (Figure 2.6 (a) and (b), respectively) as well as between their feature images (Figure 2.6 (c) and (d), respectively), multi-modal scans are transferred to mono-modal intensity systems with comparable intensity mapping (compare Figures 2.6 (c) and (e)). The final result of this step is the ability to perform mono-modal registration on feature images.

2.2.4 Registration

As discussed and shown earlier in this section, manifold learning of multi-modal medical scans followed by manifold alignment leads to multi-modal to mono-modal transformation which facilitates utilizing mono-modal registration algorithms to find registration parameters. In this work, we use intensity based, and Fourier-Mellin based [111] image registration for the problem of image registration with full and partial data, respectively.

Intensity-based registration is an iterative process that requires a metric, an optimizer, and a transformation type. We use a basic mono-modal intensity-based image registration with gradient descent optimizer that is available in the MATLAB software. The algorithm was used in its default configuration. It utilizes regular step gradient descent optimizer in order to adjust the transformation parameters so that the optimization follows the gradient of mean squares in the direction of the extrema. However, the optimizer settings were adjusted to improve the registration results.

As illustrated in Figure 2.1, the registration algorithm is supplied with the structural representation of each modality instead of the original raw images. The optimizer estimates rigid transformation matrix which then will be used to align the original sensed image with the reference.

2.3 Experimental Validation

In this section, we evaluate the general performance of our proposed intensity transformation method using Laplacian Eigenmap and the proposed manifold alignment technique in improving and facilitating the multi-modal medical scans registration, in situations either full or partial data is available. To do so, we performed three sets of experiments.

The purpose of the first set of experiments is to visually assess the effectiveness of the proposed multi-modal to mono-modal transformation method via intensity transformation over different modalities. The goal of intensity transformation is to redraw a sensed feature

image using comparable intensity levels with a reference feature image.

We expect to achieve a higher accuracy by taking advantage of well-developed mono-modal registration techniques after transferring multi-modal images into a mono-modal intensity space rather than using complicated multi-modal registration techniques. That brings the motivation of our second set of experiments. Brain images captured from same patients with different modalities are registered, and the accuracy of alignment is assessed using distance error and MI metrics, as a measure of similarity.

As we assert that partial to full image registration can also benefit from the proposed intensity transformation, specifically in the case of multi-modal registration, we have designed another set of experiments to investigate the claim.

The remainder of this section is organized as follows. First, in Section 2.3.1 and 2.3.2, we present our experimental setup and datasets which are used in this work, respectively. Then, in Section 2.3.3, we qualitatively visualize how well the multi-modal to mono-modal transformation is done using different modalities. After that, in Section 2.3.4, the competence of the proposed pipeline is evaluated against mutual information based multi-modal registration techniques. Finally, in Section 2.3.5, extensive experiments are carried out to study the effect of modality transformation in improving image registration with partial data.

2.3.1 Experimental Setup

The present experimental studies were implemented using MATLAB R2013a environment running on a personal computer with Xeon(R) E3-1245 CPU @ 3.50GHz and 32GB memory. We took advantage of the elastix toolbox [112, 113] for NMI-based registration.

2.3.2 Dataset

To validate the effectiveness of the proposed idea, we exploited two standard, widely used, and publicly available datasets of simulated and real human brain scans.

We performed experiments on T1-, T2-, and PD-weighted MR images from the BrainWeb

database[1] [114–117]. The BrainWeb dataset contains simulated brain MRI data in a variety of slice thicknesses, noise levels, and levels of intensity non-uniformity produced using an MRI simulator. The images in our experiments contained 3% noise and 20% intensity non-uniformity.

We also made use of the CT, T1-, T2-, and PD-weighted MR images from the Vanderbilt Database which was collected in Retrospective Image Registration Evaluation (RIRE) project[2] [118–120]. Moreover, RIRE database provides rectified images generated using [121] for correcting the static field inhomogeneity. The size of the 3D scans of the human brain are not necessarily equal; however, the difference in their resolution is negligible.

In the current work, all images were resized to 200×200 to be cost-effective regarding processing time and memory usage.

2.3.3 Multi-modal to Mono-modal Transformation

We aim to evaluate the performance of intensity transformation using manifold learning followed by the proposed manifold alignment technique. To inspect the general performance of the method aside from patient-dependent similarities in images, we employed images from different datasets.

The manifold of each image was studied independently. First, each high dimensional manifold was built up using patches of size 3×3 which makes the manifold containing 40,000 points in a 9-dimensional space. Then, to construct the Laplacian graph we used $k = 20$ nearest neighbors of each point. Each embedded manifold was generated in a three-dimensional space. Manifold alignment was performed by keeping the T1-weighted MRI as the reference and aligning the rest along with it. Figure 2.7 represents the result of multi-modal to mono-modal transformation. It is clear that after manifold alignment, all modalities are using similar intensity mapping as of T1-MRI, as expected.

---

[1] http://www.bic.mni.mcgill.ca/brainweb/
[2] http://www.insight-journal.org/rire/index.php
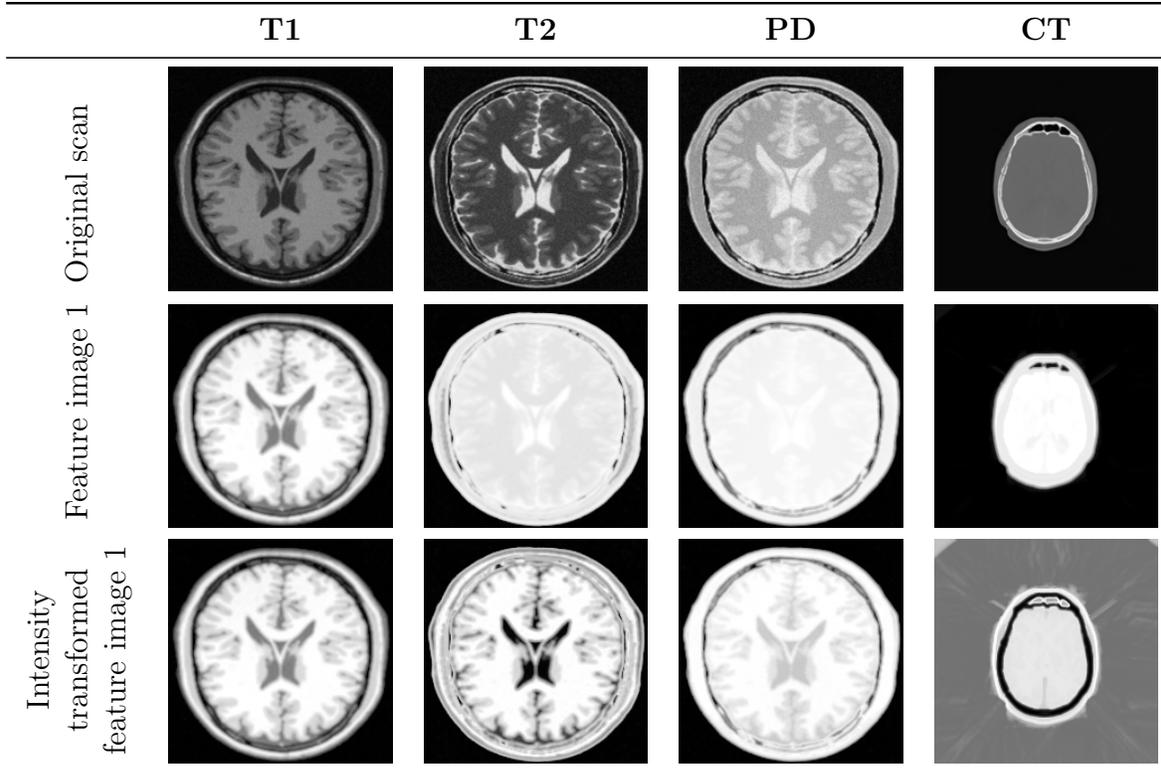
Figure 2.7: Intensity transformation of multi-modal medical scans. (*Top row*) The original MRI and CT scans of the human brain; (*Middle row*) The first feature image of each modality after manifold learning; (*Bottom row*) The same feature images after manifold alignment. Intensity mapping of feature images obtained from T2-, PD-MRI, and CT are matched with the one obtained from T1-MRI.

2.3.4   Multi-modal Registration with Full Data

To evaluate the efficacy of the proposed transformation, two sets of experiments were carried out exploiting RIRE dataset as it contains images in different modalities from same patients. The CT image of each patient was set as the reference image and scans from other modalities were aligned with the reference. The manifold of all images was studied independently using the same set of parameters including $D = 9$, $k = 10$, and $d = 3$. Then, the proposed manifold alignment technique was conducted using the the reference and sensed embeddings. Next, rigid-body registration parameters were estimated utilizing the first feature image of each modality. In the end, the registration parameters were applied on the original set of images.

In the first set of experiments, we examined the performance of registering CT scans with

multiple MR modalities using the proposed method. As the ground truth information for the RIRE testing dataset was not available, we measured the similarity metric MI between image pairs before and after registration. MI quantifies the degree of statistical dependency between the reference and the sensed images. We desire to maximize MI as indication of The current dataset contains CT, T1-, T2-, and PD-MRI scans of 10 patients. Only 6 of them have the rectified version of MRIs. Table 2.1 reports mean and standard deviation of MI measured before and after registration for each pair of the CT-MR scan. The results are statistically significant with ($p < 0.5\%$). Figure 2.8 presents an example pairwise display of CT and PD-MRI scan before and after registration with their MI values. It demonstrates the improvement of MI value resulting from the proposed method, even without the MI value being computed and optimized during registration.

Table 2.1: Mean and standard deviation of mutual information computed for each pair of the CT-MR scan before and after registration using the proposed idea. The higher MI indicates better alignment.

Modality Pair	Before Registration	After Registration
CT - PD	0.8752 ± 0.15	1.2084 ± 0.22
CT - T1	0.8685 ± 0.15	1.1863 ± 0.23
CT - T2	0.8151 ± 0.10	1.0800 ± 0.20
CT - PD Rectified	0.7846 ± 0.08	1.1132 ± 0.09
CT - T1 Rectified	0.7759 ± 0.06	1.0958 ± 0.08
CT - T2 Rectified	0.7736 ± 0.06	1.0450 ± 0.06

(a) (b)
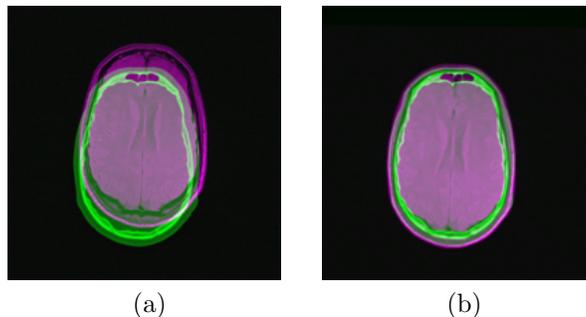
Figure 2.8: Pairwise display of a sample CT and PD-MRI scans from RIRE database. (a) Unregistered (MI=0.8276), (b) Registered using the proposed transformation followed by mono-modal registration (MI=1.1715).

To further evaluate the competence of our method quantitatively, we designed the second set of experiments. We utilized RIRE training dataset, as it contains the ground truth aligning information. We created a non-aligned slice of CT (sensed image) using randomly generated rotation angle in the range of $[-\pi/4, +\pi/4]$ and translation along each axis, constrained to keeping the brain region in the frame. Then, the distorted image was registered using three different methods: 1) the proposed method in [54, 98] using Laplacian images; 2) multi-modal registration with Mattes MI metric [122] and One plus One Evolutionary [123] optimizer; and 3) the proposed method. Then, the registration error was computed using the Mean Absolute Error (MAE) of 5 random relocated points. The experiment was repeated 30 times for each pair of modalities. Table 2.2 reports mean and standard deviation of the distance error (in millimeter). Bold-faced numbers show the best result obtained for each modality pair. All the results presented except those with asterisks are statistically significant with ($p < 0.1\%$). Registration error of CT - T1 rectified scans is statistically significant with ($p < 0.3\%$). Registration error of CT - T2 rectified scans using proposed method is higher than using MI; nonetheless, it is not statistically significant.

Table 2.2: Mean and standard deviation of distance error (in millimeter) for each pair of CT-MR modality scans. Images are aligned using Laplacian images, MI-based registration, and the proposed transformation followed by mono-modal registration.

Modality Pair	Laplacian Images (ref. [54, 98])	MI-based Reg. w/o Intensity Trans.	Mono-modal Reg. w/ Proposed Trans.
CT - PD	3.5622 ± 0.05	2.0239 ± 0.14	$\mathbf{1.3740 \pm 0.09}$
CT - T1	2.9912 ± 0.11	1.3873 ± 0.18	$\mathbf{1.0042 \pm 0.11}$
CT - T2	1.9467 ± 0.15	1.7335 ± 0.12	$\mathbf{1.4137 \pm 0.09}$
CT - PD Rect.	2.3318 ± 0.12	1.3260 ± 0.21	$\mathbf{1.1462 \pm 0.11}$
CT - T1 Rect.	2.0075 ± 0.07	0.9727 ± 0.39	$\mathbf{0.7424 \pm 0.22}$ *
CT - T2 Rect.	1.9447 ± 0.33	$\mathbf{0.9294 \pm 0.23}$	0.9922 ± 0.16 **

2.3.5 Multi-modal Registration with Partial Data

We assessed the role of intensity transformation in the success of partial image registration. In the problem of co-registration of multi-modal medical scans with partial data, while images are acquired from different modalities, the sensed (partial data) only covers part of the reference (full data) and we seek to locate the small one within the space of the larger one.

The experiments were carried out as follows. First, synthetic examples of partial data were generated by rotating, translating, and cropping the sensed image with random parameters. For some experiments, extra scaling was added to the template image. Next, both images were studied with the proposed technique (Figure 2.1) and then a feature image of each was passed on to the FMT algorithm. Once they are available, the FMT algorithm computes the transformation parameters (translation, rotation, and scale) accordingly. This step requires fine-tuning the FMT algorithm. Finally, the co-registration was performed on the original full and partial sensed images using the estimated transformation parameters.

The performance of the registration is investigated qualitatively with a pairwise display of original and registered images in Figure 2.9. It's worth mentioning that none of the following results were feasible by using FMT without benefiting from manifold learning and multi-modal to mono-modal transformation. Hence, comparison of registration error is not reported.

In a more clinically related scenario, a T1-MRI of a patient from a past scan (reference) is registered with a newly-captured T2-MRI (sensed) that has covered part of the patient's brain. Both images were acquired in the axial direction; however, they are reviewed and registered in the sagittal direction. While taking the new scan (T2), the patient was positioned differently from T1 on the imaging bed. The reference and sensed images, as well as registration result from multiple methods, are presented in Figure 2.10. Images do not represent a real patient and are synthetically generated using the BrainWeb database.

Full-Partial modality pair	T1 - T2	T2 - T1	CT - T1	PD - T2
Before registration				
Registered – MI-based method				
Registered – NMI-based method				
Registered – Laplacian Images (ref. [54, 98])				
Registered – Proposed method				
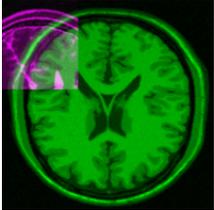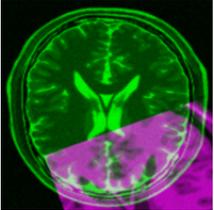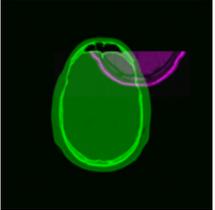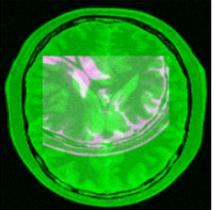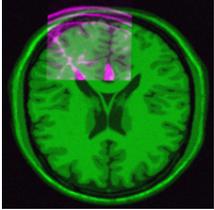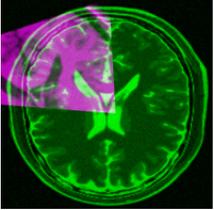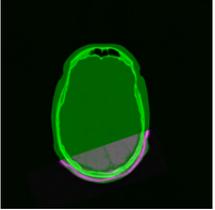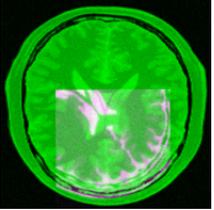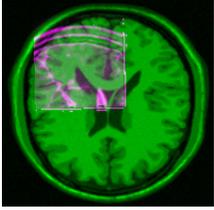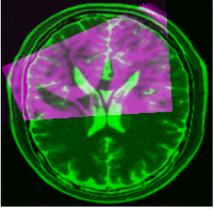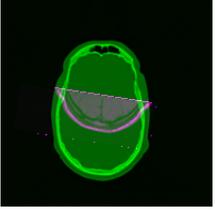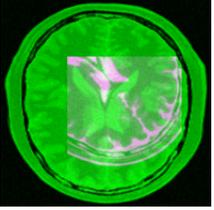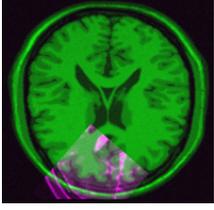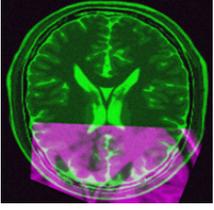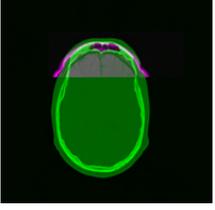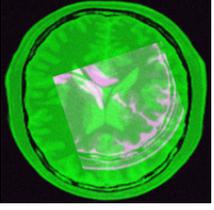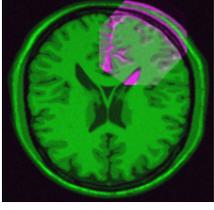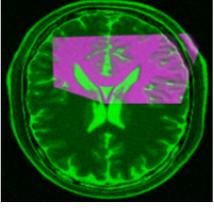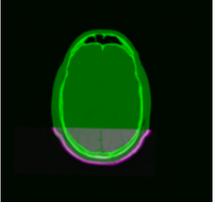
Figure 2.9: Pairwise display of multi-modal image registration with partial data.

2.4   Discussion and Conclusion

Medical image registration has hugely contributed to our understanding and interpretation of medical images. Its' application broadly lies in a variety of medium and high-level image analysis tasks, such as segmentation, object localization, tumor detection, and diagnostics analysis of X-ray, MRI, and CT images. While computational image registration methods

Input images	Registration result	
Reference (T1)	MI-based	NMI-based
Sensed (T2)	Laplacian Images	Proposed method
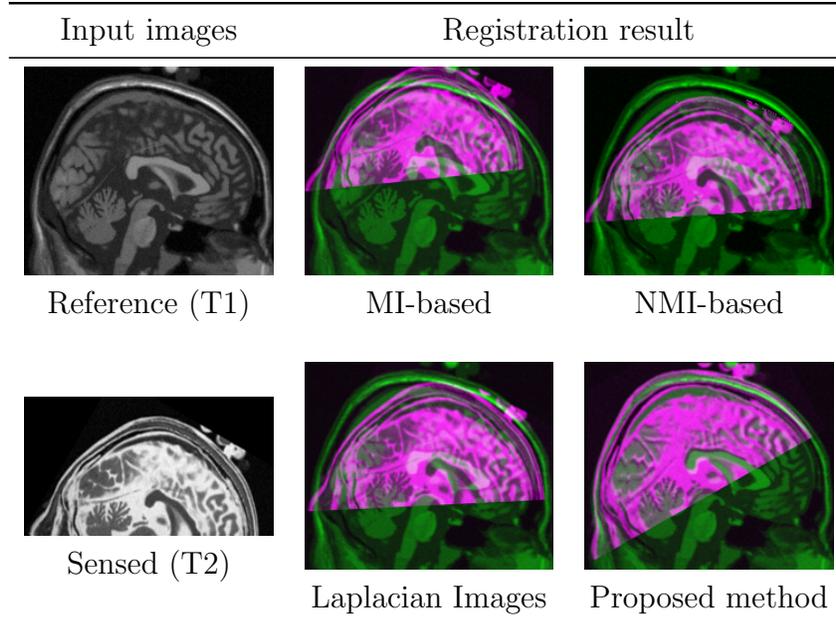
Figure 2.10: Registration of a synthetic example with partial sensed data.

have been widely available in the medical image analysis community, it is almost impossible to design and develop a general registration method which is optimized for all use cases. The present work discusses a large body of recent progress in medical image registration with a great emphasis on utilizing different computational algorithms to develop a multi-modal to mono-modal transformation technique, providing a significant capability in the multi-modal registration of medical images. Advancement in the proposed method is applicable for images with either full or partial overlap.

There are a number of advantages to the current contribution that we discuss here. The presented multi-modal to mono-modal transformation can be used as a general preprocessing step, regardless of the overlapping of images. Moreover, it facilitates recovering strong scales, rotations, and translations. In the context of medical image registration strategies, the presented method is considered as a parametric one, however, configuring and assigning the exact value to the parameters is not essential. On the other hand, this gives some degree of freedom for further tuning the smoothness of the high-dimensional manifold, which will eventually affect the convergence rate of the eigenvalue problem positively. In support of

accurate structural representation, with the use of the proposed multi-modal to mono-modal transformation, multiple feature images (maps) are generated as the output of the structural representation from which we can select. Although we evaluated the performance of the proposed transformation by only considering the rigid transformation of human brain images (except registration with partial data that considered affine transformation), we believe that registering images of other organs as well as non-rigid registration can benefit from the same concept.

Two computational bottlenecks of the proposed method are: a) the construction of the adjacency graph; b) solving the eigenvalue problem. In our experiments and with the system specification mentioned in Section 2.3.1, for an image size 200×200 with $k = 10$, the average time to construct the adjacency graph was 1, 12, and 114 seconds with $D = 9$, 25, and 225, respectively. The computational time to find the eigenvectors depends significantly on the smoothness of the manifold and the sparsity of the Laplacian matrix (parameter k). On average, for $k = 10$ to 500, it can vary from 15 to over 300 seconds.

There is also a list of limitations to this study. First of all, the current implementation is computationally expensive, which makes it inefficient for 3D volume registration. Second, the proposed pipeline requires extra processing time for learning the structure of the input images. The solutions to these limitations are beyond the scope of this paper and require more advanced computational methods (e.g., super-pixels) as well as efficient computational platforms (e.g., parallel processing on top of high-performance clusters). Furthermore, since the inner structure must be present in both modalities, the proposed transformation applies to images of anatomical modalities.

At the expense of time for computing structural representation, we obtained higher accuracy in rigid-body registration of multi-modal images with full data. More importantly, a new door is opened to the registration of partially-overlapped multi-modal images. In the future, we plan to investigate the effectiveness of the proposed intensity transformation in registering multi-modal partially-overlapped images quantitatively. In the current implementation, we

took advantage of Laplacian Eigenmap as a method of nonlinear dimensionality reduction. However, preliminary experiments indicate some other methods of manifold learning, e.g., Isomap and LLE, are also capable of intensity transformation of multi-modal images.

Chapter 3

3D Shape Analysis: Manifold Learning Approach

3.1 Introduction

3D models are ubiquitous data available in the form of 3D surface meshes, point clouds, volumetric data, etc. in a wide variety of domains, such as material and mechanical engineering [124], genetics [125], molecular biology [126], entomology [127], and dentistry [128, 129], to name a few. Processing such large datasets (e.g., shape retrieval, recognition, matching) is computationally expensive and memory intensive. For example, to query against an extensive database of 3D models to find the closest match for a 3D model of interest, one needs to come up with a good distance measure besides an efficient algorithm to implement the searching and retrieval processes. Global shape descriptors play a fundamental role in dealing with many shape analysis problems such as shape matching [130, 131], classification [132], and retrieval [133, 134].

A global shape descriptor is used to represent an original 3D shape in the form of a low-dimension vector $\mathbf{y}$ to drastically lower the burden of shape analysis as mentioned earlier. Such a vector needs to be informative yet concise to capture as much information as possible

from the 3D shape including the geometric and topological features.

3D shape descriptors have been studied in the literature in depth. For comprehensive studies and classifications of shape descriptors, we refer the readers to the surveys conducted by Tangelder and Veltkamp [135], Zhang et al. [136] and Rostami et al. [137]. While a large number of successful non-spectral shape descriptors have been proposed in the literature, spectral descriptors have proved to be beneficial in many applications [138, 139].

Generally speaking, the spectral methods take advantage of the eigenvectors and eigenvalues (referred to as the spectrum) of the Laplace-Beltrami (LB) operator applied on the 3D shapes. As the spectrum of the LB operator is computed based on the gradient and divergence that depend on the Riemannian structure of the manifold, the computed descriptors possess the isometry invariant property [140]. Various well-known local techniques such as Heat Kernel Signature (HKS) [141] and Wave Kernel Signature (WKS) [142] as well as global techniques such as Global Point Signature (GPS) [143], Shape-DNA [140] and its variants such as compact Shape-DNA (cShape-DNA) [144] have been developed from the LB operator and had found successful applications in graph processing [145], computational biology [146], and point-to-point correspondence [147].

Different discretization schema (e. g., Taubin Discretization [148], Mayer Discretization [149]) of the LB operator on the triangular meshes were discussed in [150]. The Shape-DNA and many other shape descriptors, global and local, use the cotangent scheme to estimate the LB operator. The main disadvantage of this method is that it does not converge in general and is sensitive to the peculiarities and quality of the particular triangulation of the mesh [141, 151].

Inspired by the Shape-DNA approach for extracting fingerprint signature of 3D models and motivated by the idea of dimensionality reduction for compressing data into a vector, we introduced a novel Laplacian Eigenmaps based Scale Invariant global shape descriptor, called LESI descriptor [152]. The proposed descriptor benefits from the capability of Laplacian Eigenmap in dealing with high dimensional data by introducing an exponential

weighting scheme, and eliminates the limitations tied to the well-known cotangent weighting scheme, namely dependency on triangular mesh representation and high intra-class quality of 3D models. The extensive experiments performed on LESI descriptor prove that the present contribution provides a highly discriminative and robust shape descriptor under the presence of high level of noise, random scale variations and low sampling rate, in addition to the known isometric-invariance property of the LB operator. The proposed method consistently outperformed several widely used algorithms on multiple non-rigid shape retrieval benchmarks.

We made **two contributions** in the current work which is presented in this chapter. In our **first contribution**, inspired by the idea of Laplacian Eigenmaps [34], we learn the manifold of a 3D model and then, analogous to the approach taken by Shape-DNA, we make use of the spectrum of the embedded manifold. To the best of our knowledge, this is the first time that dimensionality reduction by Laplacian Eigenmaps is used to build the global shape descriptors. There are two advantages to this approach. First, it relies on the adjacency of the nodes disregarding the fine details of the mesh structure, which is achieved by considering the length of an edge connecting two adjacent nodes rather than the two angles facing the edge. Therefore, this idea can be used for the degenerate meshes or non-uniform sampled meshes. Second, as it does not depend on the mesh structure, it can be applied easily on any other mesh types like quadrilateral meshes.

In our **second contribution**, we present a simple and straightforward normalization technique (motivated by [140, 153, 154]) to make our proposed global shape descriptor scale invariant. To this end, we propose to subtract the first non-zero eigenvalue from the shape descriptor after taking the logarithm of the spectrum. One advantage of our approach over the idea of Bronstein et al. [153] is that we avoid taking direct derivative since the differential operator amplifies the noise. We owe it to the fact that we only make use of eigenvalues, and hence, normalization of eigenfunctions are not needed.

The remainder of this chapter is organized as follows. In Section 3.2, to make the current

work self-contained, we introduce the proposed LESI shape descriptor along with some technical background. In Section 3.3, the performance of the proposed method, as well as the robustness of the algorithm, are examined and compared with multiple well-known shape descriptors by performing several qualitative and quantitative experiments using widely used 3D model datasets. Section 3.4 discusses the results in more details and draws the conclusion.

3.2 Global Shape Descriptors

Global shape descriptors, also called *shape fingerprints* or *shape signatures*, aim to simplify a 3D shape into a d-dimensional vector through the intrinsic information latent in high dimensional space. From the data compression perspective, global descriptors can also be considered as a dimensionality reduction problem. Such descriptors are desired to fulfill several properties namely isometry, scale invariance, similarity, efficiency as well as compression [140]. Shape recognition, classification and retrieval are known to be their primary applications [137].

Shape-DNA, one of the first of its kind, is a spectral analysis method that fulfills the properties as mentioned above. It is a well-known and widely used global descriptor. The original Shape-DNA and its variants require normalization process in order to obtain uniform scaling factors for the geometric objects [155]. The cShape-DNA proposed in [144] normalizes the eigenvalues of the LB operator through multiplying the spectrum by the area of the underlying shape. This approach was initially suggested in [140].

In this section, we first briefly review Shape-DNA. After that, as theoretical background of the Laplacian Eigenmaps is thoroughly discussed in section 1.2.2, we elaborate our proposed LESI descriptor; a Laplacian Eigenmaps based scale invariant global shape descriptor.

3.2.1 Shape-DNA

Shape-DNA, introduced by Reuter et al. in 2006 [140], utilizes the first k smallest eigenvalues (referred to as the spectrum) of LB operator as a shape signature. It grabbed a great deal of attention for its unique features such as isometry and rotation invariance [140].

The Laplace-Beltrami (LB) operator Δ is a linear differential operator defined on a Riemannian manifold as the divergence of the gradient of a function f on a differentiable manifold as the following form [140, 144]:

$$\Delta f = div(grad(f)). \tag{3.1}$$

Let $\mathcal{M}$ be a compact Riemannian manifold. The heat diffusion equation that describes the distribution of heat over $\mathcal{M}$ is governed by:

$$\Delta u(x,t) = -\frac{\partial u(x,t)}{\partial t}, \tag{3.2}$$

where $u(x,t)$ is the temperature at the point $x \in \mathcal{M}$ at time t. Given initial heat distribution $u(x,0) = u_0(x)$, and Dirichlet boundary condition (if manifold has a boundary), the solution to (3.2) is called *heat kernel* and can be numerically computed. Explanation of heat kernel is essential as the basis of Shape-DNA computations [140].

It is shown that the heat kernel $k_t(x,y)$ can be expressed as [141]:

$$k_t(x,y) = \sum_{i=0}^{\infty} e^{-\lambda_i t} \mathbf{\Phi}_i(x) \mathbf{\Phi}_i(y) \tag{3.3}$$

where λ_i and $\mathbf{\Phi}_i$ are respectively eigenvalues and associated orthonormal real eigenfunctions of the LB operator, which are the solutions to the Laplacian Eigenvalue problem that can be stated as [155]:

$$\Delta f = -\lambda f. \tag{3.4}$$

The spectrum of the LB operator is isometric invariant, independent of the shape location, informative (contains a considerable amount of geometrical and topological information), and can become scale invariant without difficulty. These properties are motivations for defining the global shape descriptor Shape-DNA over the sequence of non-zero eigenvalues of LB operator [140].

The Laplacian Eigenvalue problem on a triangle surface mesh is discretized using linear Finite Element Method resulting in the *generalized eigenvalue problem* of the form [151]:

$$\mathbf{A}_{cot} f = -\lambda \mathbf{B} f, \qquad \text{where}$$

$$\mathbf{A}_{cot}(i,j) := \begin{cases} \frac{cot\alpha_{ij} + cot\beta_{ij}}{2} & (i,j) \text{ edge} \\ -\sum_{k \in N(i)} \mathbf{A}_{cot}(i,k) & i = j \end{cases} \tag{3.5}$$

$$\mathbf{B}(i,j) := \begin{cases} \frac{|t_1| + |t_2|}{12} & (i,j) \text{ edge,} \\ \frac{\sum_{k \in N(i)} |t_k|}{6} & i = j \end{cases}$$

where $N(i)$ stands for the index set of the 1-ring neighborhood of the vertex i, $|t_i|$ represents the area of the triangle t_i. In particular, t_1 and t_2 are the two triangles sharing the edge (i,j) and α_{ij} and β_{ij} are the two angles facing the edge (i,j) [151]. Since $\mathbf{A}$ and $\mathbf{B}$ are sparse positive semi-definite symmetric matrices, the eigenvalues are non-negative $0 \leq \lambda_0 \leq \lambda_1 \leq ... \leq +\infty$ and eigenfunctions are orthonormal. The generalized eigenvalue problem can be solved with direct solvers or the Lanczos method [155].

3.2.2 LESI

From the compression perspective, a global shape descriptor is similar to the problem of dimensionality reduction in terms of squeezing the latent information of a 3D model into a vector. Owing to the fact that Laplacian Eigenmap has two properties of structural equivalence and locality preservation, we propose a global shape descriptor using the spectrum of

graph Laplacian.

Theoretical background of Laplacian Eigenmaps, a nonlinear dimensionality reduction method, is extensively covered in chapter 1. Briefly speaking, Laplacian Eigenmaps is a simple-core algorithm which tends to capture the intrinsic geometry of a manifold and seeks for an optimal embedding that has a minimum distance from the manifold. The optimal embedding is obtained by employing the LB operator over the graph Laplacian.

The Laplacian Eigenmap, similar to Shape-DNA, deals with sparse, symmetric, and positive semi-definite matrices and solves generalized eigenvalue problem. Hence, it is computationally efficient, and its' spectrum has all the properties of the Shape-DNA descriptor, above all is isometric invariant and informative. Moreover, the locality preservation property makes it to a certain degree insensitive to outliers and noise.

Laplacian Eigenmap is limited neither to 2D/3D manifolds nor a specific type of, e.g., triangulated, mesh representation. It applies to an any-dimension smooth manifold. Therefore, unlike other spectral methods that utilize the cotangent weighting scheme, the proposed descriptor is capable of handling any connectivity variations as well as defective models. Besides, its' convergence error rate is relatively low. Motivated by the mentioned benefits, LESI descriptor is proffered as follows.

The proposed descriptors' pipeline, illustrated in Figure 3.1, complies with Laplacian Eigenmaps steps followed by a normalization step to generate a scale-invariant descriptor. The major difference with Laplacian Eigenmaps algorithm is that the neighborhood connectivity is already defined in the 3D mesh model and hence, we can skip the first step which is constructing a graph Laplacian.
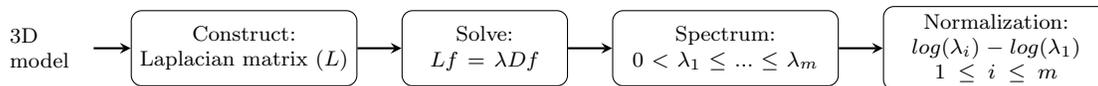
Figure 3.1: The pipeline of the proposed anisotropic Laplacian Eigenmaps based scale-invariant (LESI) global shape descriptor.

For a given surface mesh, we first compute the weight matrix $\mathbf{W}$ for a 1-ring neighbor

of every point as stated in the equation (1.8). This way, closer points are assigned with higher weight and so, will have more influence on the process of embedding. In the current work, the parameter σ is suggested to be chosen as the maximum distance among all 1-ring neighborhood points:

$$\sigma = maxdist(x_i, x_j) : (i, j) \in V, w_{ij} \neq 0. \tag{3.6}$$

Therefore, all non-zero weights would be in the range ($e^{-1/2} \approx 0.6 \leq w_{ij} \leq 1$). As the connectivity of adjacent points on the surface mesh is defined in both ways, there is no need to make the weigh matrix symmetric forcefully. From there, the Laplacian and degree matrices ($\mathbf{L}$ and $\mathbf{D}$, respectively) can be computed without difficulty.

Next, we form the generalized eigenvalue problem (equation (1.7)). After solving the equation, we take the sequence of smallest non-zero eigenvalues as the shape descriptor. Since in most cases, a single 3D model is made up of one connected component, we only need to leave out one eigenvalue and its corresponding eigenfunction.

For shape retrieval and comparison, one of the requirements of a well-designed shape descriptor is its insensitivity to scale variations. According to Weyl's law [156], when an object is scaled with factor α, the eigenvalues are scaled with factor α^2. In order to achieve a scale invariant shape descriptor, we propose to eliminate the parameter of scaling by taking a logarithm of the descriptor vector and then, computing the difference of the new vector from its smallest element. Further details of the proposed *normalization technique* are shown in Figure 3.2.

In Figure 3.2, two Teddy Bear models are shown. One model is in the original size, whereas the other model is scaled with a factor of 0.7. After constructing the Laplacian matrix and solving the generalized eigenvalue problem, it is clear from Figure 3.2(b) that the spectrum of scaled model is almost $(0.7)^2$ of the spectrum of the original model. Taking logarithm of the spectrum, takes away the scaling factor from multiplicand and leaves it as augend. Therefore, subtracting one term (e.g., the first element) removes the scaling factor from all other terms. The result is a normalized and scale-free spectrum.

| (a) Teddy Bear models | (b) Spectrum of LB operator | (c) Logarithm of spectrum | (d) Normalized Spectrum |
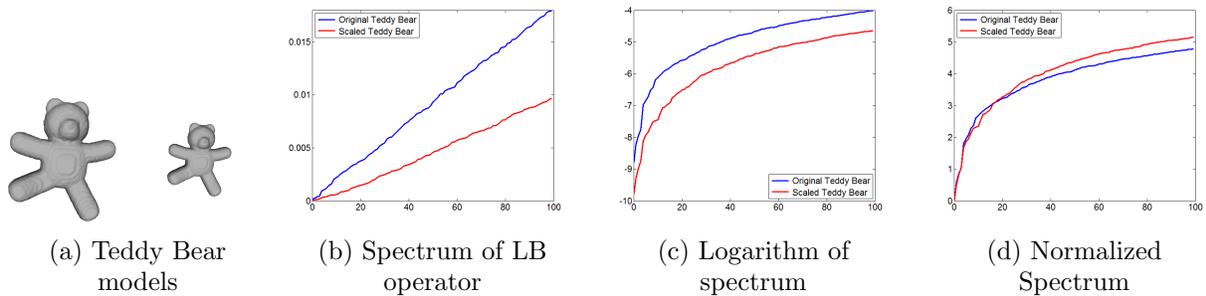
Figure 3.2: An example showing the proposed method for normalizing the spectrum. (a) The Teddy Bear model and its down scaled version (scale factor 0.7). (b) The spectrum of original (blue) and scaled (red) Teddy Bear models. Please note that the original spectrum is approximately multiplied by half. (c) The Logarithm of the spectrum shown in (b). (d) The normalized spectrum of original and scaled Teddy Bear models after subtracting first element of logarithm of the spectrum .

3.3   Experimental Validation

In this section, we evaluate the general performance of our proposed shape descriptor and compare it with several widely-used spectral-based global shape descriptors. In section 3.3.1, we first present two datasets used in our experiments. Then, in section 3.3.2, we qualitatively visualize and measure the competence of the proposed method in discriminating different clusters compared with candidate methods from the literature. Next, in Section 3.3.3, we validate the effectiveness of the LESI descriptor to distinct multiple classes by measuring the accuracy of multi-class classification. Finally, in section 3.3.4, extensive experiments are carried out to study the robustness of the proposed shape descriptor with respect to noise, scale invariance, and down sampling. All the algorithms were implemented using MATLAB R2013a environment running on a personal computer with Intel(R) Core(TM) i3-4130 CPU @ 3.40GHz and 12GB memory.

3.3.1   Dataset

To validate the utility of the proposed shape descriptor, we utilized two standard, widely-used, and publicly available datasets of 3D polygon meshes. The high-resolution TOSCA

dataset [157] contains 80 three-dimensional non-rigid models, including 11 cats, 6 centaurs, 9 dogs, 4 gorillas, 8 horses, 12 women poses, 3 wolves and two men with 7 and 20 poses respectively. In our experiments, we use all models except the gorilla models, as they contain isolated points. The models in each class of the TOSCA dataset are almost identical in terms of scale, the number of vertices, quality of triangulation, and structure, which all represent the same object with different poses.

To assess the capability of our proposed shape descriptor in describing models with poor intra-class quality, we test our algorithm on the McGill dataset with articulating parts [158]. It contains 3D models of 30 ants, 30 crabs, 25 glasses, 20 hands, 29 humans, 25 octopuses, 20 pliers, 25 snakes, 31 spiders, and 20 Teddy bears. The classification of the McGill dataset models is more challenging due to scale and shape variations.

3.3.2 Retrieval Results

In this section, we evaluate the general performance of our proposed shape descriptor and compare it with several state-of-the-art spectral-based global shape descriptors including Shape-DNA [140], cShape-DNA [144], and GPS [143] algorithms. We chose these methods because they are widely used by researchers (e.g., [159–161]) to develop new descriptors or applications, or to evaluate the performance of their proposed descriptors. Moreover, cShape-DNA represents the normalized version of the original Shape-DNA. Even though there are multiple ways to convert a local point descriptor to a global shape fingerprint, in this article we focus only on algorithms that have been originally introduced as global fingerprints. To this end, we take advantage of the source code made available on Dr. Kokkinoss homepage[1] [153], as well as the shape descriptor package provided by Li et al. [162] available on a GitHub repository[2] to generate the Shape-DNA and GPS descriptors, respectively. We also compare the performance of shape retrieval using the code provided for evaluation by SHREC'11 [163].

The shape descriptors are compared using the TOSCA dataset to discriminate between

---

[1]http://vision.mas.ecp.fr/Personnel/iasonas/descriptors.html
[2]https://github.com/ChunyuanLI/spectral_descriptors

different classes of 3D objects. In this experiment, we use the first 33 non-zero eigenvalues ($d = 33$). Then, to visualize the locations of objects in the shape space, we project them onto a 2D plane using Principle Component Analysis (PCA). Figure 3.3 displays the effectiveness of our method compared with the fingerprints of interest.

(a) Shape-DNA	(b) cShape-DNA	(c) GPS	(d) LESI
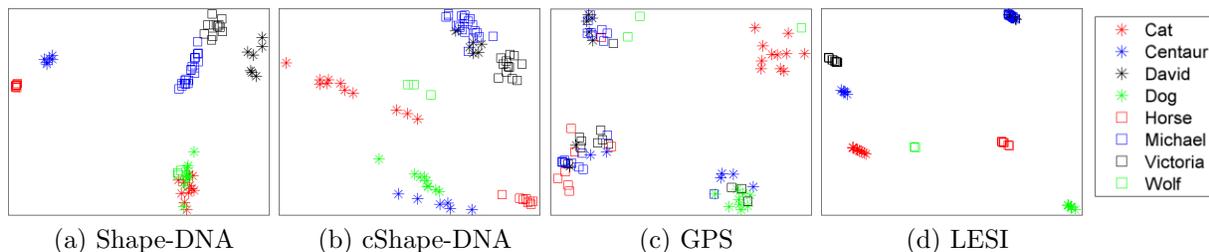
Figure 3.3: 2D PCA projection of shape descriptors computed using (a) Shape-DNA, (b) cShape-DNA, (c) GPS, and (d) LESI algorithms on TOSCA dataset.

Figure 3.3 reveals that LESI can differentiate models of various classes significantly better than the other methods for a refined and normalized dataset. Even though all human models (David, Michael, and Victoria) are very similar, it can distinguish the women from the men's group. However, it fails to discriminate models of Michael from David. Despite the large isometric deformations in each class, the proposed LESI method clusters all models of the same class together very tightly.

To demonstrate the power of our method in classifying objects with low intra-class similarity compared with other shape descriptors, the same experiment is carried out on the McGill dataset. Models of the same class with articulating parts are in different scales, shape, and structure. The 2D PCA projections of 33-dimension descriptors from all four algorithms are shown in Figure 3.4.

As illustrated in Figure 3.4, the original Shape-DNA is highly sensitive to scales. Multiple methods are presented in [140] to make the descriptor normalized to scale. cShape-DNA represents a normalized version of it by multiplying the descriptor with the surface area. Although cShape-DNA can separate models from each other, classes are not separated efficiently. LESI outperforms the other algorithms by providing distinct descriptors, which can

(a) Shape-DNA	(b) cShape-DNA	(c) GPS	(d) LESI

Figure 3.4: 2D PCA projection of shape descriptors computed using (a) Shape-DNA, (b) cShape-DNA, (c) GPS, and (d) LESI algorithms on McGill dataset.

Table 3.1: Shape retrieval performance using TOSCA and McGill datasets

Dataset	Method	NN	FT	ST	E	DCG
TOSCA	ShapeDNA	**1.0000**	0.8091	0.9391	0.4486	**0.9584**
	cShapeDNA	0.9474	0.7748	0.8984	0.4748	0.9241
	GPS	0.4868	0.4244	0.6320	0.3614	0.6787
	LESI	0.8684	**0.8456**	**0.9430**	**0.4860**	0.9244
McGill	ShapeDNA	0.7922	0.3452	0.4977	0.3411	0.7192
	cShapeDNA	0.7882	0.3943	0.5483	0.3852	0.7470
	GPS	0.3843	0.2508	0.4066	0.2588	0.6020
	LESI	**0.9647**	**0.7046**	**0.8739**	**0.6644**	**0.9251**

separate classes. Shape descriptors offered by LESI prove superior to the other algorithms in the shape retrieval and classification tasks, as described below and in the next section respectively.

To examine the superiority of LESI quantitatively, we computed multiple standard retrieval measures including Nearest Neighbor (NN), First Tier (FT), Second Tier (ST), e-Measure (E), and Discounted Cumulative Gain (DCG). These measures represent state-of-the-art quality metrics used when evaluating matching results for shape-based search engines [164]. Table (3.1) reports the results of shape retrieval. Boldface numbers indicate the highest value for each measure per each dataset. From Table (3.1), it is clear that the LESI descriptor outperforms all other methods concerning all measures in retrieving models from the McGill dataset. When retrieving models of the TOSCA dataset, LESI outperforms all methods concerning FT, ST, and E measures. Shape-DNA outperforms LESI by a higher value for NN and DCG measures, due to the poor discrimination between David and Michael

performed by the LESI descriptor. However, it does not diminish the validity of our claim that LESI performs well for meshes with non-uniform sampling or peculiarities.

3.3.3 Multi-class Classification Results

In this section, we corroborate the findings of Section 3.3.2 by training a linear multi-class SVM classifier to assess the accuracy of LESI compared to other shape descriptors. For this experiment, we utilized the McGill dataset. In addition to the shape descriptors evaluated in Section 3.3.2, we computed another normalized version of Shape-DNA by dividing the feature vector by its first element (similar to what LESI offers) as suggested in [140]. This way we can compare the effect of the exponential weighting scheme without the influence of the normalization method or compactness (offered by cShape-DNA). Using 10-fold cross-validation and repeating the experiment 3 times, we report the average accuracy for each method in Table (3.2).

The new LESI approach significantly outperforms all other methods when using a two-tailed paired t-test ($p < 0.05$). The t-test was performed on one set of 10 folds in order to avoid violating the independence assumption of the t-test. There is a significant improvement in accuracy when comparing the Shape-DNA (Normalized) to other variants of the Shape-DNA, which is due in part to the normalization method. However, the average accuracy of the LESI descriptor is noticeably higher (95%) when compared to 90% of the Shape-DNA (Normalized).

Table 3.2: Classification accuracy using McGill dataset

Method	Average accuracy
Shape-DNA	21.02%
Shape-DNA (Normalized)	90.60%
cShape-DNA	71.37%
GPS	50.11%
LESI	**95.69%**

Finally, Figure 3.5 shows the confusion matrix obtained from the linear multi-class SVM

Figure 3.5: Confusion matrix obtained from linear multi-class SVM for McGill dataset using LESI descriptors.

using LESI descriptor. The number of correct classifications made for each class (indicated by the green diagonal), confirms that our method captures the discriminative features of the shapes.

3.3.4 Robustness

In this section, we address the robustness of the LESI shape descriptor to shape variations, including noise, scale, and down-sampling by performing another set of experiments. First, we generate the disturbed version of every model in the TOSCA dataset. Then, we test the capability of every method mentioned above in discriminating between different classes. For this purpose, besides plotting the 2D PCA projection of shape descriptors, we also compute and plot the pairwise Euclidean distance matrix, in every case. The distance matrix represents the dissimilarity between each pair of models in the set. It is often used to compute other evaluating metrics such as nearest-neighbor, and first and second tier, to name a few. The dissimilarity of descriptors increases from blue to red, and the more separate classes differ in color, the better they are discriminating from each other.

Resistance to noise. Multiple noisy versions of the TOSCA dataset are generated following the idea articulated in [165]. To this end, the surface meshes of all models are disturbed by changing the position of each point along its normal vector that is chosen randomly from an interval $(-L, L)$ with the 0 mean, where L determines the noise level and is a fraction of the diagonal length of the model bounding box. In this experiment, three noise levels $L = 0.5\%$, $L = 1\%$, and $L = 2\%$ are tested, where the latter one represents a greater level of noise. Two-dimensional PCA projections of all descriptors with the presence of different levels of noise are plotted in Figure 3.6. Combining these with the results shown in Figure 3.3, where no noise is present, demonstrates that the LESI algorithm is highly noise-resistant while the performance of the Shape-DNA and cShape-DNA decreases as the level of noise increases. Moreover, GPS fails in separating different classes of models with the presence of noise. Figure 3.9 reflects the effect of noise on the discriminative power of the descriptors. The LESI algorithm shows consistent results as the level of noise increases from 0% (top row) to 2% (bottom row).

Scale invariance. In order to validate the insensitivity of the LESI descriptor to scale variations and compare the robustness of the proposed method with other descriptors, each model of the TOSCA dataset is scaled by a factor of 0.5, 0.875, 1.25, 1.625, or 2 randomly. Figure 3.7 shows that the LESI algorithm surpasses other methods in discerning different classes. Comparing the result of this experiment with the results shown in Figure 3.3 demonstrates the consistency of the LESI algorithm with the presence of scale variation. The distance matrices in Figure 3.10 show that the original Shape-DNA algorithm is very susceptible to scale variations. Even though the cShape-DNA has significantly improved scale sensitivity of the original Shape-DNA, it does not provide as accurate results as the LESI algorithm does.

Resistance to the sampling rate. To investigate the effect of sampling rates on the discriminative power of the shape descriptors, Bronstein et al. [134] propose to reduce the
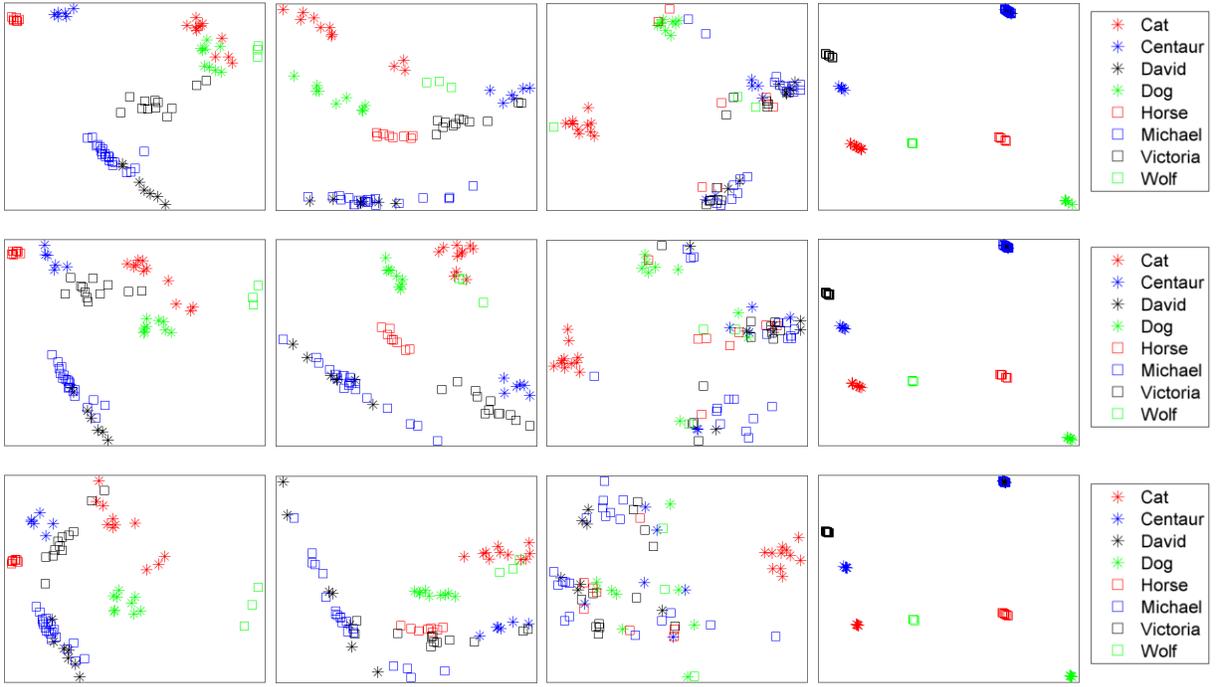
Figure 3.6: 2D PCA projection of shape descriptors computed by (from left to right) Shape-DNA, cShape-DNA, GPS, and LESI algorithms on perturbed TOSCA dataset with (from top to bottom) 0.5%, 1%, 2% noise level, respectively.
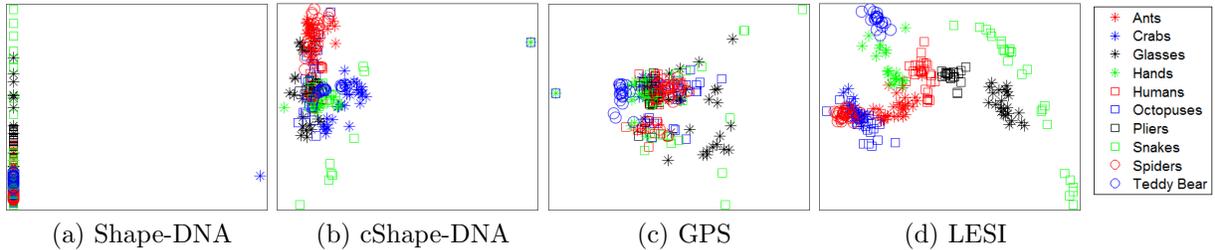

(a) Shape-DNA   (b) cShape-DNA   (c) GPS   (d) LESI
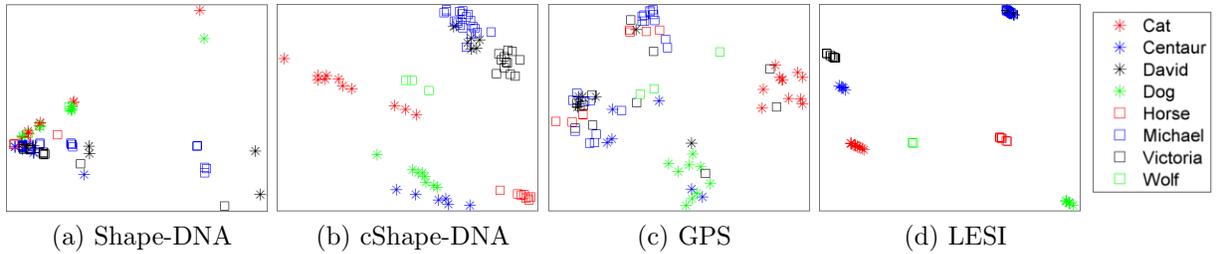
Figure 3.7: 2D PCA projection of shape descriptors computed by (a) Shape-DNA, (b) cShape-DNA, (c) GPS, and (d) LESI algorithms on scaled samples of TOSCA dataset by a randomly chosen factor of 0.5, 0.875, 1.25, 1.625, or 2.

number of vertices to 20% of its original size. Accordingly, the down-sampled version of the TOSCA dataset is generated, and shape descriptors associated with them are computed. The 2D PCA projections and distance matrices of descriptors are illustrated in Figures 3.8 and 3.11, respectively. Although the original Shape-DNA shows a more accurate result than cShape-DNA, the separation of cat, dog, and wolf models is challenging. Although the performance of the LESI method is slightly affected, it still outperforms cShape-DNA and GPS methods.

    (a) Shape-DNA   (b) cShape-DNA   (c) GPS   (d) LESI
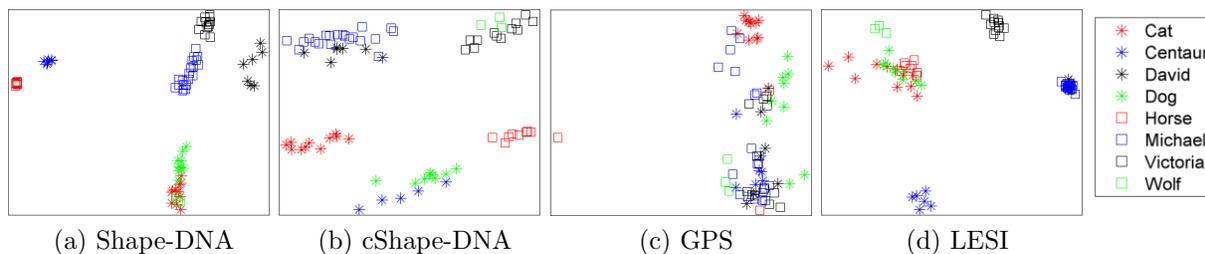
Figure 3.8: 2D PCA projection of shape descriptors computed by (a) Shape-DNA, (b) cShape-DNA, (c) GPS, and (d) LESI algorithms on down sampled TOSCA dataset by rate of 20%.

3.4   Discussion

In this chapter, motivated by the unique properties of Laplacian Eigenmap (i.e., locality preservation, structural equivalence, and dimensionality reduction) and inspired by the existing spectral-based shape descriptors, we investigated the application of manifold learning in deriving a shape fingerprint in order to address the limitations tied to popular cotangent-based shape descriptors. We proposed a global descriptor (LESI) with an easy-to-compute and efficient normalization technique that facilitates applications such as shape classification and retrieval. Our method applies fewer restrictions on the class of meshes as well as improving the quality of tessellations. Analogous to other spectral descriptors, LESI uses the spectrum of the LB operator, which is independent of the shape location, is informative (contains a considerable amount of geometrical and topological information), and above all

Noise Level	Shape-DNA	cShape-DNA	GPS	LESI
0%				
0.5%				
1%				
2%				
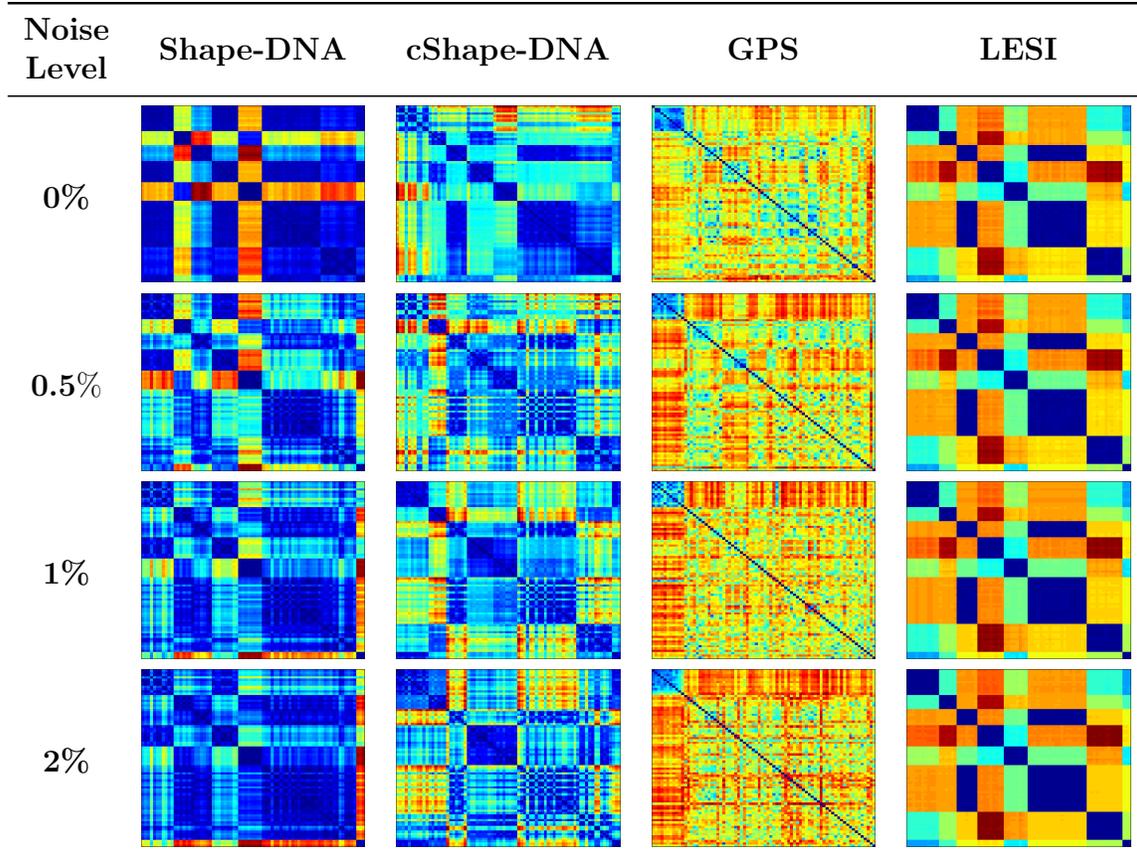
Figure 3.9: The Euclidean pairwise distance matrix of shape descriptors computed by (from left to right) Shape-DNA, cShape-DNA, GPS, and LESI algorithms on perturbed TOSCA dataset by (from top to bottom) 0%, 0.5%, 1%, 2% noise levels.

	Shape-DNA	cShape-DNA	GPS	LESI
Random Scale				
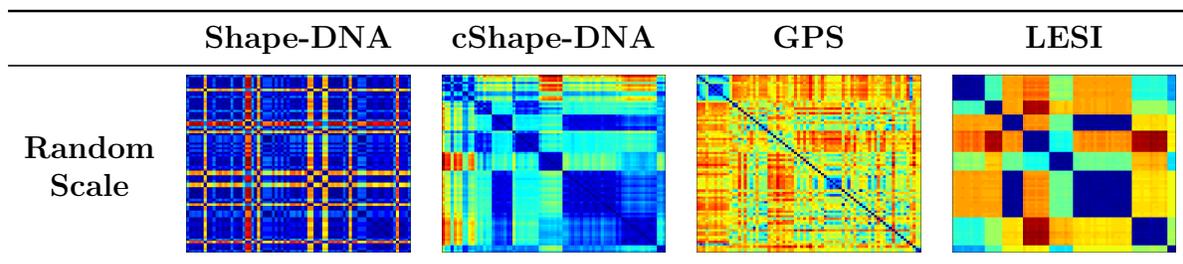
Figure 3.10: The Euclidean pairwise distance matrix of shape descriptors computed by (from left to right) Shape-DNA, cShape-DNA, GPS, and LESI algorithms on scaled version of the TOSCA dataset by a randomly chosen factor of 0.5, 0.875, 1.25, 1.625, or 2.

isometric invariant. We compared the discriminating power of LESI with three prominent descriptors from the literature, namely Shape-DNA, cShape-DNA, and GPS, and found it to be superior.

In the first set of experiments illustrated in Figures 3.3 and 3.4, our method substantially

	Shape-DNA	**cShape-DNA**	**GPS**	**LESI**
Down Sampled				
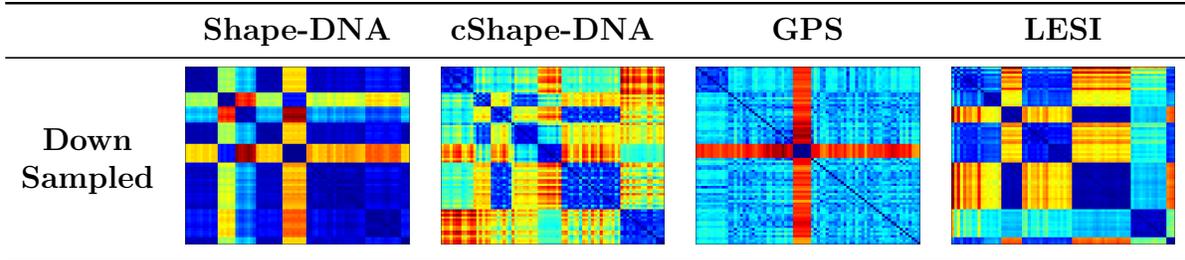
Figure 3.11: The Euclidean pairwise distance matrix of shape descriptors computed by (from left to right) Shape-DNA, cShape-DNA, GPS, and LESI algorithms on down sampled TOSCA dataset by rate of 20%.

outperforms the others. The superiority of LESI is more significant when the McGill dataset is used (Table 3.2 and Figure3.5). This dataset includes wide variations in mesh structure and scales, causing the failure of the other methods to generate acceptable results. However, LESI, due to utilizing a different method of discretization to form the LB operator, focuses on the vicinity rather than the quality of the triangulation. Therefore, our technique, unlike other methods, is not affected by the low quality of polygon meshes.

The second set of experiments evaluates the reliability of our method in the presence of noise, scale variations, as well as different sampling rates. LESI shows impressive robustness against the first two sets of perturbation. Despite the negative impact of down sampling in LESI descriptor, it continues to show better performance when compared to cShape-DNA and GPS. It should be noted that the result could also be improved by increasing the size of the output vector.

In addition to the discriminating power of the descriptor, degenerate and non-uniform meshes may also cause failure of an algorithm to converge. The cotangent weight-based algorithms were not able to compute the descriptors for 2 shapes from the McGill dataset. GPS also failed to compute descriptors for 6 models of the down sampled TOSCA dataset. However, our technique converges at all times despite the quality of the polygon mesh structure.

Moreover, LESI, unlike cotangent weight-based techniques, is not confined to the triangulated meshes as it disregards the mesh geometry [166]. LESI inherits this property from

the capability of manifold learning techniques in coping with high dimensional data. The discretization of the LB operator using cotangent weights on the quadrilateral meshes is not as straightforward as on triangular meshes. To compute the LB operator on a quadrilateral mesh, all rectangles need to be divided into triangles. It could be done easily, however, as for each quad there are two possible triangulations, the result is not unique.

In the original Laplacian Eigenmaps, the high dimension data requires a considerable amount of processing as the list of all connections need to be computed for the dataset. In fact, for each point in the high dimension space, a given number of nearest neighbors need to be extracted which could be challenging and unmanageable. While applying this technique to the 3D meshes, we skip this step as the neighbors are already defined and given in the mesh structure.

This work benefits from the Laplacian Eigenmap technique in a space in which the vicinities are given. LESI takes advantage of simple Laplacian computation, to form the LB operator, which provides concise and informative shape descriptors. Experimental results prove that LESI is more effective compared with the other powerful descriptors.

One limitation of LESI is its inability to separate models of different men (David and Michael). However, it was able to differentiate between the wolf and dog, as well as between women and men.

Although we investigated only the application of Laplacian Eigenmap in introducing a shape descriptor, there are some other spectral-based manifold learning methods, such as Isomap, LLE, and Diffusion map, which have not been examined. This can be considered future work.

Chapter 4

Lung Cancer Descriptive Analysis: A Combined Manifold and Deep Learning Approach

4.1 Introduction

Lung cancer is the second most common cancer disease among both men and women, and has been the leading cause of cancer mortality worldwide, surpassing colon, breast, and prostate cancers combined [167]. It is a heterogeneous and aggressive form of cancer, accounting for 1 in every 4 cancer deaths in the United States [167]. Lung cancer does not develop specific symptoms until it spreads. However, studies have shown that annual screening of people with higher risk of getting lung cancer, using low-dose CT scan, can save many lives [168].

An uncontrolled cell growth in tissues of the lung, called a pulmonary nodule, may result in lung cancer. Pulmonary nodules are easy to spot on an X-ray or CT scans, but hard to diagnose [169]. Diagnosis of a nodule involves determining whether it is cancerous or non-cancerous. Accurate diagnosis of lung nodules from pulmonary abnormalities is of vital importance for early detection and effective treatment of lung cancer, where any misclassifi-

cation of lung nodules may lead to additional medical costs, invasive surgeries, unnecessary lung biopsies, prolonged hospitalizations and/or physically aggressive treatment plans. It becomes more apparent when statistics state that the majority – at least 60 percent of lung nodules overall – are not malignant or cancerous, and thus, there exists a pressing need to build highly accurate descriptive models to analyze lung nodules reliably [169].

Medical imaging of the chest/lung using X-ray, CT, PET/CT, and MRI is widely used by radiologists and physicians to screen and analyze lung nodules. Given the harmful effects of lung nodule misclassification mentioned earlier, computer-aided diagnosis (CAD) systems are needed to improve patient survival rates as well as to minimize patient radiation exposures. Advanced computational methods including machine learning and computer vision techniques combined with various medical images have the potential to serve as an accurate and timely descriptive model that can assist physicians in their ability to diagnose and treat lung nodules [170–173].

4.1.1 Pulmonary Nodules: Background

A pulmonary nodule is an abnormal growth in the lung tissue that looks like a small, round or oval-shaped spot. Hence, it also is called "a spot on the lung". Pulmonary nodules are very common; about 0.2% of chest X-rays reveal the presence of pulmonary nodules. However, not every nodule is cancerous. Pulmonary nodules can be *benign/non-cancerous* or *malignant/cancerous*. A pulmonary nodule is defined as three centimeters in diameter or less. Otherwise, it is called a mass and more likely represents a tumor [169].

Early detection of pulmonary nodules is essential as it increases the survival rate of patients significantly. There are a few symptoms associated with the presence of lung nodules, and often the nodules are misdiagnosed with a chest cold or mild flue. Most pulmonary nodules, about 90%, have been discovered accidentally while the patient had an X-ray or CT scan taken for other purposes [169].

One of the notable differences between a benign and malignant nodule is the growth rate.

The growth rate of a benign nodule tends to be zero or very small. On the other hand, a malignant nodule tends to grow relatively quickly, usually doubling in size every four months. However, it is possible to grow as fast as doubling in size in 25 days [169].

There are three approaches to determine whether a nodule is benign or malignant. The first is to monitor the growth rate of the nodule. As discussed earlier, malignant nodules engulf surrounding tissues and grow fast. The second way is to evaluate the nodules' apparent features such as size, edge characteristics, and calcification. While benign nodules tend to have smoother and more regular shape with monochromatic appearance, malignant nodules are known to have irregular shapes, rough surface structure and vary in color over the entire surface. The last approach is using invasive methods such as a biopsy to retrieve cells by using a needle or performing a localized surgery [169, 174, 175].

Information gained from a nodule's growth rate and its morphological structure provide doctors with enough evidence to make reliable diagnoses. This fact motivates us to exploit computational algorithms to obtain quantitative measures that distinguish cancerous nodules from non-cancerous ones.

4.1.2 Literature Review

Lung nodule analysis using a CAD system consists of several components including candidate nodule region-of-interest (ROI) detection, false-positive reduction, and nodule classification [176]. A candidate ROI detection method needs to have high sensitivity in order to detect every suspicious region; meanwhile, a false-positive reduction algorithm removes non-nodule regions to reduce the volume of further processes. A nodule classification method, which is the focus of this research, aims to distinguish malignant nodules from benign ones. In recent years, numerous research articles introduced pulmonary nodule classification as a two-step classification process: 1) extract concise and distinct features; and 2) train a classification model [177]. Although for a specific problem statement some classification models provide a more robust and accurate out-of-sample prediction than others, still the

performance of classifiers relies on the extracted features.

Hand-crafted features including shape, intensity, and texture have been used to develop a nodule classification model [178–182]. In 2010, El-Baz et al. [178] utilized Markov-Gibbs Random Field (MGRF) to model the spatial distribution of image intensities (Hounsfield values) from segmented nodules aiming to develop an early diagnosing of malignant nodules. Later in 2012, authors extended their work to assess malignancy of lung nodules in 3D space [179]. In an alternative approach, El-Baz et al. [180] proposed to describe shape complexity of lung nodules with a shape index; the number of Spherical Harmonics (SHs) required to approximate the 3D surface of a lung nodule. Their approach is based on the hypothesis that malignant nodules tend to have more irregular surface structure when compared with benign nodules that have smoother edge characteristics. Besides the morphological features, the effectiveness of texture features has also been investigated [181, 182]. To assess the efficacy of texture features in differentiating malignant and benign pulmonary nodules, Han et al. [181] compared three widely-used texture features, namely the Haralick [183], Gabor [184], and Local Binary Pattern (LBP) [185]. Also, authors explored the sensitivity of these features to image thickness and proposed an extension of 2D Haralick features to 3D space. Later, in 2016, Dhara et al. [182] presented a combined shape and texture features extracted from 2D and 3D image slices/volumes for the purpose of lung nodule classification. Authors extracted 56 features from segmented nodules and examined the contribution of these features in nodule classification by computing the AUC and their p-values from two-tailed Student's t-test. The task of classification was completed by training an SVM classifier using the 49 most relevant features.

In recent years, Deep Convolutional Neural Networks (DCNN) has made a huge impact in the field of computer vision and consequently, has found applications in various domains such as image classification [186], visual tracking [187, 188], concrete crack detection [189], and biomedical image segmentation [190], to name a few. Since the introduction of DCNN, automatic feature extraction approach has gained researchers interest in the area of medi-

cal image analysis. Deep learning architectures derive higher level features from lower level features using a combination of linear and non-linear transformations in a hierarchical fashion. For the purpose of diagnostic classification of lung nodules, several works have been published in the literature [191–200]. Most early works on lung nodule classification were designed and evaluated at the 2D image slice level using 2D CNNs (e.g., [191, 192]). This approach does not benefit from the 3D context of pulmonary nodules to its full extent. To account for the volumetric information due to the 3D nature of a nodule, two different approaches were taken: 1) processing multi-view snapshots of a nodule with 2D CNNs to mimic a 3D volume (e.g., [193, 194]); and 2) utilizing 3D CNNs with 3D convolutional and 3D max-pooling layers (e.g., [195]). In another approach, to improve the classification performance of the CNN architecture, multi-scale CNNs were proposed aiming to extract features at multiple scales and to provide information from the surrounding tissue of a nodule with multiple viewing areas. That idea was evaluated in 2D-scale (e.g., [197]) and 3D-scale with one-view-one-network (e.g., [196, 198, 199]) and multi-view-one-network (e.g., [200]) implementation strategies. One drawback of 3D CNNs is that they are computationally intensive and require long training time.

A requirement of deep learning architecture in image recognition is a large-scale, well-annotated dataset that represents intra-class variations and characteristics. However, obtaining such a dataset in the medical imaging domain is very challenging for two primary reasons: rigorous regulations that limit access to patient data, and the need for domain-specific experts who wish to spend a vast amount of time on annotating a large-scale dataset [201]. As a result, deep-based classification models face the problem of over-fitting [202]. To address the over-fitting issue, a number of research articles propose combining deep features with shape/texture features for obtaining a more reliable prediction of future observations [177, 203]. Buty et al. [203] proposed combining appearance features extracted from a DCNN model with shape fingerprints obtained from the scale and rotation-invariant spherical harmonic analysis of 3D shapes aiming to characterize the degree of malignancy of lung

nodules. In another research, Xie et al. [177] proposed a classification approach by combining deep features extracted from a DCNN model, texture features computed from the Gray Level Co-occurrence Matrix (GLCM), and shape features obtained from the Fourier shape descriptor. A common point among all the deep learning-based classification models that have been introduced so far is that the input to their network is a series of 2D image slices or a 3D volume (a stack of 2D images). The former one ignores partial information related to the 3D context of lung nodules, and the latter one is extremely computation intensive.

Studies show that morphological features and edge characteristics of lung nodules provide useful information to distinguish between benign and malignant nodules [204]. While a considerable amount of research activities has been conducted regarding lung nodule classification using 2D/3D intensity-based images, little attention has been paid to developing a model based upon the surface structure of nodules. This motivates **our first aim**, which is to classify lung nodules based on the features that are extracted from their surface structure. **The second aim** focuses on a computational perspective. **Our main contribution** is in the development of a lung nodule descriptive model in which, we propose to use deep features computed using point sets sampled from 3D surfaces. To address the over-fitting issue caused by using a relatively small medical imaging dataset, and to account for the connectivity of sampled points, which is ignored in point-cloud based deep learning models, we propose combining deep features with shape fingerprints computed using surface spectral analysis. Our proposed classification approach uses discriminating features that are extracted using a point-cloud based deep learning model along with a spectral analysis of the 3D shape surface. To the best of our knowledge, deep learning networks that are based on the 3D surface structure of lung nodules have not been employed for automatic feature extraction, nor have they been combined with shape descriptors. A summary of the reviewed related works is presented in Figure 4.1.

The rest of this chapter is organized as follows. In Section 4.2, we propose a descriptive model that employs deep learning and manifold learning techniques with an objective to
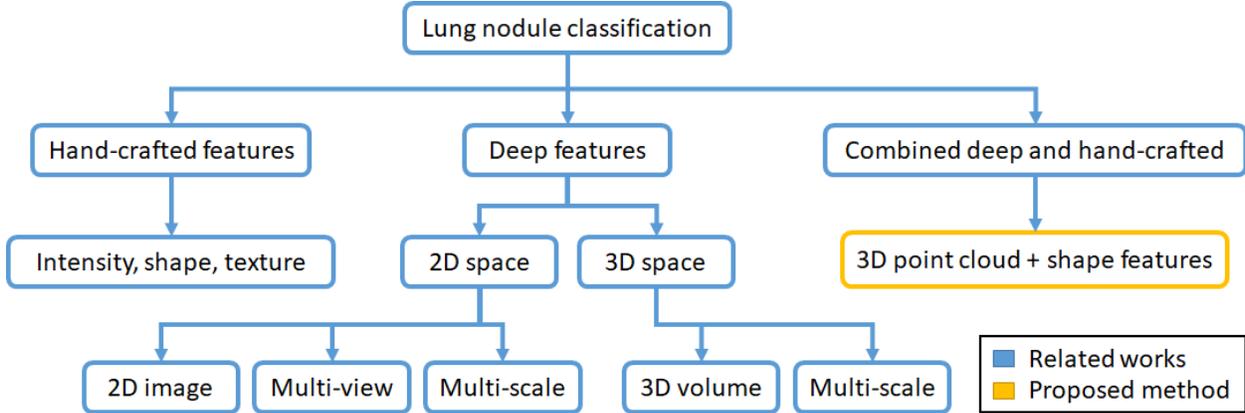
Figure 4.1: Summary of related works and the proposed classification approach.

extract discriminative features from the surface structure of lung nodules. Then, in Section 4.3, the evaluation experiments are presented, and the results are reported. Finally, the main findings, limitations, and future work are discussed in Section 4.4.

4.2 Proposed Method

In this section, we elaborate on a descriptive model that characterizes pulmonary nodules aiming to specify whether a nodule is benign or malignant. The objective here is to obtain a robust and accurate pipeline that has clinical relevance. Meaning that We desire to develop a classification model with high sensitivity in identifying cancerous nodules; meanwhile, we wish to obtain a high specificity to prevent complications caused by a false diagnosis.

Our descriptive model benefits from the power of deep learning models in automatic extraction of high-level features along with the spectral analysis of nodule surface structure, to make up for shortcomings associated with not having a large-enough dataset in medical imaging domain. Figure 4.2 presents the general pipeline of our proposed model. The pipeline starts with nodule localization and segmentation from a chest CT scan. Then, having the binary volume, a 3D surface polygonal model is generated. The feature extraction component is made up of two branches; extracting deep features from a 3D shape deep learning architecture, and extracting shape fingerprints from well-established and mathematically

powerful spectral analysis of 3D models. In the last step, we combine these features, and feed them into a classification model; in our case, a random forest classifier. Once a classifier is trained, it predicts the category of new instances. To assess the generalization capability of the trained classifier, we followed the k-fold cross-validation strategy.
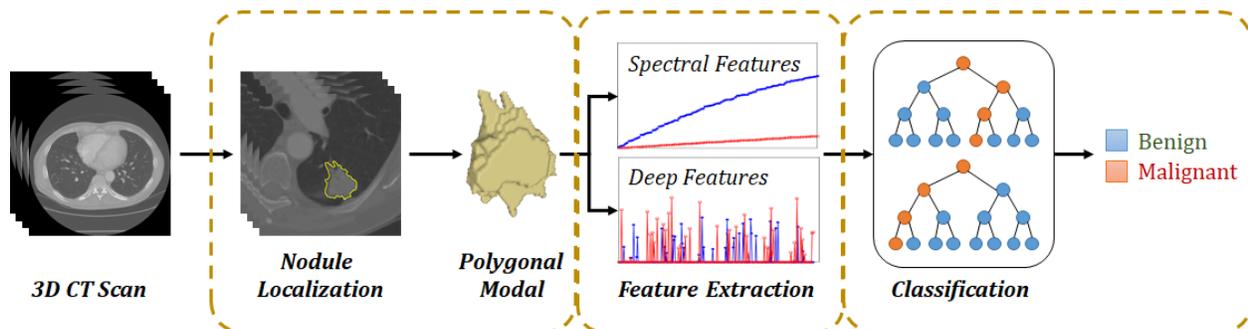
Figure 4.2: Pipeline of the proposed descriptive model for pulmonary nodule classification from lung CT scans.

Extensive research has been done on nodule localization and polygonal model generation. Therefore, the primary focus in the current contribution is on the feature extraction and classification components, which we will discuss in more details.

4.2.1 Dataset and Preprocessing

LIDC-IDRI Dataset. The Lung Image Database Consortium (LIDC) Image Database Resource Initiative (IDRI) image collection is a web-accessible resource that is provided and funded by the National Cancer Institute (NCI), the Foundation for the National Institutes of Health (FNIH), and the Food and Drug Administration (FDA) through The Cancer Imaging Archive (TCIA) repository. It consists of 1018 thoracic computed tomography (CT) scans from 1010 patients with marked-up annotated lesions. Each individual scan was reviewed by four experienced thoracic radiologists in a two-phase image annotation process, and is accompanied with an associated XML file that provides information about every possible nodule including its category (non-nodule, nodule$\geq$3 mm, and nodule$<$3 mm), boundary markings for nodules$\geq$3 mm, and malignancy rate evaluated by each physician on

a 5 point scale. The LIDC-IDRI dataset also contains nodule-specific details, i.e., a list of available cases, the associated size of each nodule, and the total number of nodules for each patient [205–208].

Preprocessing. We adopted the protocol reported in [194] which only considers nodules≥ 3 mm diagnosed by three or more radiologists. We computed a composite malignancy rate by taking the median of all malignancy levels. Then, binary classification ground truth labels were derived by discarding nodules with composite malignancy level equal to 3. Nodules with a composite level less (greater) than 3 are considered benign (malignant). Table 4.1 reports the number of nodules in each class.

Table 4.1: Description of ground truth information on the LIDC-IDRI dataset

Composite malignancy rate	1,2	3	4,5
Number of nodules	446	511	430
Classification label	Benign	-	Malignant

Before generating a polygonal surface mesh for each nodule, we extracted nodule volume by masking pixels that were inside or on the boundary region of the nodule on a slice-by-slice basis. For that purpose, we collected the boundary markings of each nodule drawn by physicians independently. A consensus of opinion among physicians is obtained by keeping pixels that are part of the nodule area detected by two or more physicians. Then, having the binary volume of a nodule, we generated a polygonal model using iso2mesh mesh generation toolbox [209–211]. These steps are shown in Figure 4.3. In the end, a fixed number of points were sampled on the mesh surface with probabilities according to the faces area to generate the point cloud representation of every nodule. The number of sampled points is 2048, determined by the input size of the PointNet network.
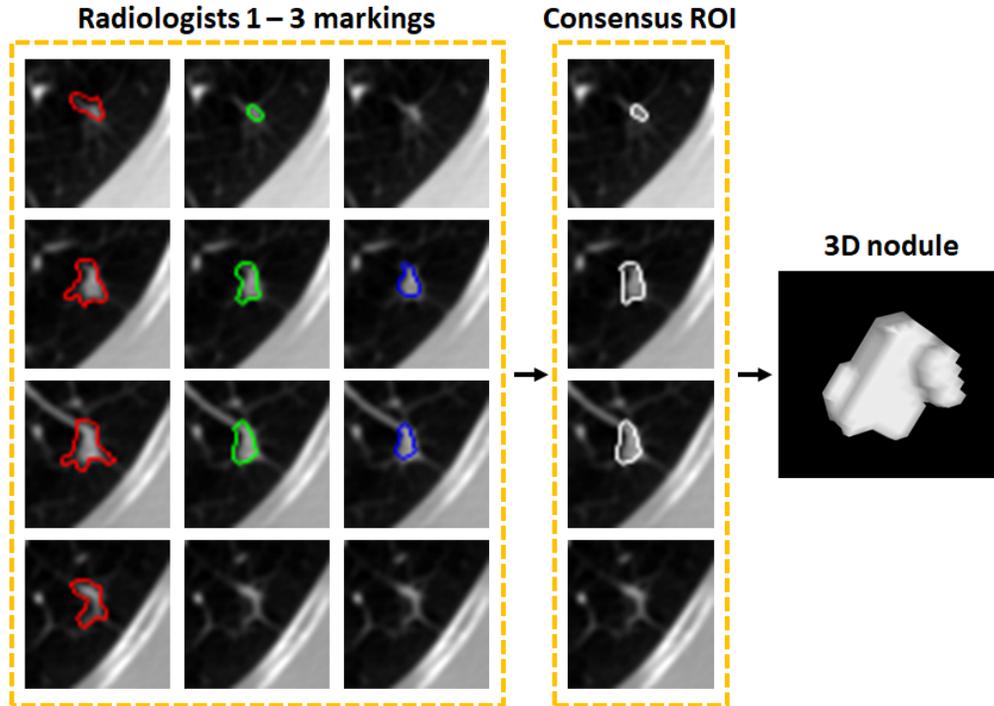
Figure 4.3: Consensual ROI extraction and 3D model generation of a lung nodule.

4.2.2 Deep Feature Extraction

Deep learning architectures that accept 3D voxel grids as input are computationally expensive. Moreover, shape representation using volumetric occupancy grids is unnecessarily voluminous [212], and more importantly, a significant portion of volume does not contain informative content useful for nodule classification.

To make use of the 3D context of lung nodules, meanwhile, avoiding the extensive computations related to volumetric data, we suggest sampling points from the reconstructed surface meshes. Sampled points form a point cloud (PC), which represents the geometry and morphological structure of a lung nodule in a compressed form. A general and unified deep network that learns abstract and useful features of 3D PCs is PointNet [212].

PointNet is a deep convolutional architecture on 3D geometric point sets. This simple yet efficient network is invariant to permutations of points ($N!$ permutations), as well as rigid transformations of the point cloud in whole. PointNet accepts unordered point sets as input, learns the general structure of the point cloud, and provides n scores for an n-class

classification task. It is robust to noise as it finds which points contribute the most to the global features. The number of parameters and FLOPS/sample (floating-point operations per sample – a representative of speed) utilized by PointNet is significantly less than other volumetric CNN networks [212].

In this descriptive model, we utilize PointNet as a deep feature extractor. PointNet contains a series of Multi Layer Perceptron (MLP) combined with input and feature transformations to extract a vector of abstract and efficient features. At the end of the network, a three-layer MLP provides the prediction scores. To utilize the network as an efficient deep feature extractor, we first train its network with training samples. Then, for every new test sample, we take features from the last MLP layer before the output layer as the extracted features.

While PointNet captures local structures from neighboring points, it ignores the connectivity of points that form the surface. The edges in the surface mesh representation form faces and boundaries. Also, PointNet looks for global points that contribute to the part segmentation and object category detection. In this process, it does not take advantage of the relative distance between neighboring points. To address the issue that comes with PC-based deep networks, we suggest incorporating spectral shape descriptors in the descriptive model.

4.2.3   Spectral Feature Extraction

The spectral analysis of geometric shapes aims to extract a compact and informative feature vector – called a shape fingerprint – from the surface structure of a 3D shape. It defines the Laplace-Beltrami (LB) operator on the Riemannian manifold to find the distribution of various functions, e.g., heat or wave, over the surface of the manifold. The LB-operator and its spectrum offer unique features that are useful for descriptive analysis of 3D geometric shapes. Among those features, rotation and isometric-invariance are the most important ones [140].

The spectrum of the LB operator consists of a series of real and non-negative eigenvalues, obtained from solving the Laplacian eigenvalue problem. While eigenvectors form a set of real and orthogonal basis that forms the algebra of a function over the manifold, the spectrum contains information that characterizes the geometry [213]. In the past two decades, the eigenvalues and eigenvectors of the LB-operator have been used to develop several local and global shape descriptors including Shape-DNA [140] and LESI [152].

Spectral shape descriptors are isometric invariant; however, they are able to efficiently characterize, and tell the difference between benign and malignant nodules due to the amount of surface deformation between two classes. When a geometric shape is represented with a 3D surface mesh, the spectral method uses one of the several discretizations of the Laplace operator to solve the Laplacian eigenvalue problem. One key difference between Shape-DNA and LESI descriptors is the method of discretization used to estimate the distribution of function on the sampled mesh points. The method of discretization used in the descriptor leads to forming completely different Laplacian matrices which eventually determines some properties of the method such as its ability to handle open surfaces, peculiarities of the triangulation, and accuracy of distribution for non-uniform meshes [141, 151, 214].

Shape-DNA uses widely-used cotangent-based discretization method that is capable of handling surface boundaries. However, organic structures such as lung nodules are closed surface meshes. Therefore, they can be analyzed with descriptors that generally have a problem dealing with manifolds with boundaries, e.g., LESI.

In some problems, such as shape retrieval of everyday objects, it is more desirable to have a scale-invariant shape fingerprint. Therefore, some shape descriptors provide an optional method for scale normalization that removes the parameter of object scale from the fingerprint. However, the scale of a nodule is a decisive factor in the problem of lung nodule classification, and so, scale-sensitive descriptors are more effective.

In the current project, we evaluated the performance of LESI descriptor in extracting useful and distinctive features with and without scale normalization. The one with better

performance was used in the implementation of the final descriptive model.

4.2.4 Classification with Random Forest

After extracting features from a PC-based deep network and a spectral shape descriptor, it is time for building a predictive model. Prediction is made by fitting the training data into a classification model that learns any relationship between features and class labels. In the current contribution, we used random forest as a classification model with Gini impurity criterion as a measure of the overall gain in class purity.

Random forest is an ensemble decision tree that improves the prediction accuracy and controls over-fitting of individual trees. Random forest classifier has a few substantial highlights. It decorrelates individual trees by randomly sampling from the set of features that contribute to fitting a tree. Sampling features at each split can be implemented with replacement (i.e., bootstrap). Also, it reduces the variance of predictions by aggregating decisions made by a number of trees. Not only is it highly efficient on a large dataset, but also it ranks the importance of features [215].

4.3 Experimental Validations

4.3.1 Experimental Setup

The present experimental study has been carried out on a high-performance computer with 64-bit MS Windows 10 operating system, Intel(R) Xeon(R) E3-1245 CPU @ 3.50GHz, and 32GB memory. The current system is equipped with an NVIDIA Quadro M5000 GPU. Data preprocessing and spectral analysis of lung nodules are implemented using MATLAB R2017b environment. Deep learning of point cloud sets is performed using Anaconda 5.1 and Python 3.6 with TensorFlow 1.10, CUDA 9.0, and cuDNN 7.1.4.

The dataset was split into training and testing sets as reported in Table 4.2. To find the best classification model, and to achieve a reliable estimate of the out-of-sample classification

prediction, 10-fold cross-validation was utilized. After the deep network, as well as the classification model, were trained with the training set, the performance of the whole model was evaluated with the actual out-of-sample data from the testing set.

Table 4.2: Train/test split of the LIDC-IDRI dataset

	Benign	Malignant	Total
Training set	348	336	684
Testing set	98	94	192
Total	446	430	876

4.3.2   Experimental Results

We evaluated the performance of the proposed descriptive model (Figure 4.2) on the dataset presented in Table 4.2 using 10-fold cross-validation. More specifically, the training set is split randomly into 10 folds; using 9 folds for training and the last fold for testing, and repeating the experiment such that each fold is used once for testing. Several classification metrics including accuracy, sensitivity, specificity, and the area under the receiver operating characteristic curve (AUC) are recorded to assess the predictive capability of the model. Optimization of hyper-parameters for random forest classifier was obtained using cross-validated grid-search over a set of parameters. The k-Fold cross-validation provides an estimate of prediction accuracy for out-of-sample data. After tuning the parameters of the random forest classifier, we examined the performance of the model on the testing set (actual out-of-sample data). The results of the proposed model have presented in Table 4.3. In this table, we compare the performance of our model with multiple state-of-the-art lung nodule classification algorithms.

As Table 4.3 shows, our proposed method outperforms methods that use hand-crafted features (as in [181, 182]), multi-scale deep learning features (as in [196]), and a combination of shape, texture and image-based deep learning features (as in [177]) in accuracy and sensitivity. Even though our specificity and AUC are lower than some other methods, our

Table 4.3: Classification metrics of five studies.

	Han et al. [181]	Dhara et al. [182]	Shen et al. [196]	Xie et al. [177]	Proposed
Accuracy	85.59	88.38	87.14	86.79	**88.54**
Sensitivity	70.62	84.58	77.00	60.26	**87.23**
Specificity	93.02	90.03	93.00	**95.42**	89.80
AUC	89.25	**95.76**	93.00	-	92.30

method is the one that obtains both sensitivity and specificity over 87%. Higher specificity obtained by some methods is the result of low sensitivity; meaning misclassification of malignant nodules. High specificity is important for reducing unnecessary procedures; meanwhile, high sensitivity and accurate labeling is essential for making informed decisions.

In another experiment, we investigated the effect of scale normalization on the discriminating power of the shape fingerprint. We hypothesized that the addition of scale normalization would reduce the accuracy of prediction, simply because the nodule size is a key feature in classification. Comparing the performance of the spectral feature with and without scale normalization, reported in Table 4.4, proves the hypothesis. It also suggests that eliminating the parameter of scale from the shape fingerprint has a devastating impact on the feature vector by shifting a majority of predictions toward the malignant class. Therefore, while the sensitivity of predicting malignant nodules is increased, the specificity is significantly reduced; meaning many benign nodules are classified as malignant.

Table 4.4: The impact of scale normalization on the performance of the spectral feature using Laplacian Eigenmaps (LE).

	LE	LE + Norm
Accuracy	**84.90**	70.31
Sensitivity	86.17	**92.55**
Specificity	**83.67**	48.98
AUC	**90.52**	78.57

We also trained a separate classifier for each set of feature vectors to investigate the contribution of each set in the overall prediction. Table 4.5 compares the performance of each

set of features (deep from PointNet and spectral from LESI) with the combined descriptive model. Combining features improved false-positive reduction significantly (p-value < 0.05). It is good to note that due to the small size of the testing data, the increase in the sensitivity of deep features compared with the combined feature set, is not significant.

Table 4.5: Contribution of deep (Pnet) and spectral (LE) features in the overall classification.

	Pnet	LE	Pnet + LE
Accuracy	83.85	84.90	**88.54**
Sensitivity	**88.30**	86.17	87.23
Specificity	79.59	83.67	**89.80**
AUC	91.23	90.52	**92.30**

4.4 Conclusion and Discussion

Diagnostic classification of lung nodules is of high importance. In this contribution, we developed a descriptive model that provides high sensitivity and specificity. Having a high sensitivity for a classification model results in a high rate of malignant nodule detection, and an increase in the overall survival rate; meanwhile, a high specificity reduces false-positive detections and unnecessary invasive procedures such as biopsies. Comparing our model with four other studies shows that the proposed descriptive model achieves higher accuracy with sensitivity and specificity both more than 87%.

Our proposed descriptive model is based on an analysis of the surface structure of lung nodules. It combines deep features that are obtained from a PC-based deep network with shape fingerprints obtained from a spectral analysis of the 3D surface mesh. Addition of spectral features to the model addresses two problems: 1) the over-fitting problem of deep networks caused by relatively small dataset in medical imaging domain; and 2) ignoring the connection of neighboring points that form the face of the 3D shape. Experimental results show that combining deep and spectral features significantly (p-value < 0.05) improves the specificity and overall accuracy of prediction. By comparing the performance of spectral

features with and without scale normalization, we found out that spectral features without built-in scale normalization are more appropriate for this study as the nodule scale is a determining factor in class prediction.

In the present work, the number of sampled points for all nodules is constant, which is determined by the input size of PointNet. A question that arises here is whether that sampling rate provides enough information from large nodules. In our observations, iso2mesh generates non-uniform sampled meshes in which, more samples are taken at irregular spots and fewer samples on smooth surfaces. Therefore, for LIDC-IDRI dataset, we only up-sampled point clouds.

The proposed model benefits from PC-based deep learning of 3D shapes, which speeds up the analysis. Because it only consumes sampled points from the surface of the nodule and ignores voluminous data from the internal structure of the nodule that does not contain informative content useful for nodule classification. Our model mimics what physicians do as they review a patient's scan slice-by-slice and envision the nodules' structure in their mind. Another advantage of this approach is that it can simply be generalized to diagnostic classification of abnormalities at other organs. Although the method proposed by Dhara et al. presents higher overall AUC and better specificity, their method uses hand-crafted features which makes it less adaptable to new tasks.

One limitation of the proposed descriptive model is related to data preparation for analysis. Unlike image-based deep learning networks that accept an image patch as input directly from CT scans, the proposed model with its current implementation relies on manual segmentation of nodules, which is a tedious and clinically impractical task.

As a future work, we would like to equip our descriptive model with an end-to-end data preparation package using semantic segmentation of biomedical images. Not only it will remove the ambiguity of disagreement between radiologists, but also it makes it suitable for use in clinics. Also, PointNet would benefit from the calcification information if we map the intensity inside the nodule to the surface color. It can help with classification of pulmonary

nodules with Ground Glass Opacity (GGO). Moreover, visualizing the network provides an understanding of the network performance, which can help with improving the model in the future. Two other possible future works can be summarized as: 1) using transfer learning that has the potential to compensate the limitations of deep learning tied to relatively small dataset in the medical image domain; and 2) using a graph-based deep learning instead of a PC-based network as it considers edge-connectivity into the network and eliminates the need for spectral analysis.

Chapter 5

Conclusions and Outlook

5.1 Summary

In this dissertation, I demonstrate the innovative use of machine intelligence to a number of medical data processing problems using advanced computing techniques including manifold and deep learning methods. I begin by emphasizing the importance of medical imaging in the healthcare community, which has been recognized as one of the top 10 medical developments in the past thousand years. Then, I provide technical background information on Principal Component Analysis (PCA) and Laplacian Eigenmap as linear and non-linear manifold learning techniques. Finally, the application of manifold learning in different problems including image segmentation, shape classification, and descriptive analysis of pulmonary nodules are discussed in Chapters 2–4, respectively.

Chapter 2 discusses the challenges related to registering multi-modal images. As intensity variations between images make some similarity metrics inefficient in this kind of problem, missing and partial data eliminates one-to-one correspondence between input images, and causes the failure of many mutual information-based registration techniques. One approach that has recently been investigated is reducing the problem of co-registration of multi-modal images to mono-modal scans. This research introduces an intensity transformation which

relies on learning a manifold generated from image patches by using Laplacian Eigenmap as a manifold learning technique. It also presents a simple and efficient manifold alignment that completes the intensity transformation in the manifold embedded domain. As a result, using RIRE dataset, mean absolute error of 1.00, 1.41, and 1.37 mm are obtained for registering CT images with T1-, T2-, and PD-MRIs, respectively. We also empirically investigate the efficacy of the proposed transformation in registering multi-modal images with partial data.

Chapter 3 introduces a spectral-based global shape fingerprint that addresses the limitations tied to popular cotangent-based shape descriptors, namely, dependency on triangular mesh representation and high intra-class quality of 3D models. Besides the state of being invariant to rotations and isometries, this study presents a straightforward and easy-to-compute normalization method that takes out the parameter of scale from the shape descriptor. Unlike methods with a built-in scale normalization step, gives the opportunity to preserve the parameter of scale in problems in which scale is a determining feature, such as tumor classification.

Chapter 4 demonstrates the novel application of manifold learning to characterize pulmonary nodules aiming to specify whether a nodule is benign or malignant. Building on the idea presented in Chapter 3 combined with recent advances in deep learning on point sets, we propose a descriptive model that outperforms a number of state-of-the-art methods with accurate predictions. The model shows high sensitivity in identifying cancerous nodules, as well as high specificity that reduces false-positive diagnosis. Using the LIDC-IDRI dataset, we obtained 88.54% accuracy, 87.23% true-positive rate, 89.80% true-negative rate, and 92.30% area under the receiver operating characteristic curve (AUC).

The common theme throughout these chapters has been using manifold learning to study the surface structure of a smooth manifold, whether it is generated from image patches or provided by 3D surface meshes. Manifold learning along with other computational methods including manifold alignment, scale normalization, and deep networks characterizes the underlying structure of the problem of interest and offers an accurate and robust solution.

5.2 Future Work

This research can be extended from both computation and application perspectives. Below, I point interested readers to a few areas of research.

Computational perspective. In Chapter 4, we proposed a descriptive model that benefits from the structural differences between benign and malignant nodules to extract discriminative features. However, in our model, a nodules' structure is represented in the form of a point cloud. With recent advances in generalizing deep architectures to non-Euclidean structured data, intrinsic convolutional neural networks have the potential to learn and describe non-rigid deformations of the surface. Such deep architectures give an opportunity to enrich the non-Euclidean structured data used in the learning process by including other related features, e.g., nodule calcification, vertex color, and point normal.

Segmentation of 3D structures. One application of manifold learning is in graph clustering, which is to partition a graph in such a way that it optimizes a cut criterion. Therefore, these learning techniques can be used for image segmentation. Whether the smooth manifold is constructed using image patches or other forms of image representatives, such as SIFT features, the computational complexity of manifold learning methods hampers the segmentation of 3D structures from image volumes. A possible approach to this problem could be incorporating an adaptive region-growing technique into segmentation of 2D image slices using manifold learning.

Dermatosis Conditions. A skin condition, a condition in which the integumentary system of the body is affected, can be categorized into many different types. In many types, the skin condition can be detected with naked eyes. However, even with the help of an expert, diagnosis of a skin condition may require a skin biopsy or special imaging technique called dermoscopy. People with these conditions may face a number of challenges such as limited access to specialized physicians supplied with advanced imaging devices, the cost, and convenience of screening on a regular basis. Recent advances in machine intelligence with computational techniques that have been developed in computer vision and learning

domains can maximize the chance of identifying, characterizing, and effective treatment of these conditions including lesion outlining and classification, accurate measurement of a healing wound, and Melanoma image analysis, to name a few.

Bibliography

[1] Driving data growth in healthcare. Report, EMC Digital Universe with Research and Analysis by IDC, 2014. [Online] https://www.emc.com/analyst-report/digital-universe-healthcare-vertical-report-ar.pdf [Accessed: 11.13.2017].

[2] Kunio Doi. Diagnostic imaging over the last 50 years: research and development in medical imaging science and technology. *Physics in medicine and biology*, 51(13):R5, 2006.

[3] Chuanren Liu, Fei Wang, Jianying Hu, and Hui Xiong. Temporal phenotyping from longitudinal electronic health records: A graph based framework. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, pages 705–714. ACM, 2015.

[4] Wei-Qi Wei and Joshua C Denny. Extracting research-quality phenotypes from electronic health records to support precision medicine. *Genome medicine*, 7(1):41, 2015.

[5] Rimma Pivovarov, Adler J Perotte, Edouard Grave, John Angiolillo, Chris H Wiggins, and Noémie Elhadad. Learning probabilistic phenotypes from heterogeneous EHR data. *Journal of biomedical informatics*, 58:156–165, 2015.

[6] Jonathan D Mosley, M Benjamin Shoemaker, Quinn S Wells, Dawood Darbar, Christian M Shaffer, Todd L Edwards, Lisa Bastarache, Catherine A McCarty, Will Thompson, Christopher G Chute, et al. Investigating the genetic architecture of the PR in-

terval using clinical phenotypes. *Circulation: Cardiovascular Genetics*, 10(2):e001482, 2017.

[7] Sameer Pradhan, Noémie Elhadad, Wendy W Chapman, Suresh Manandhar, and Guergana Savova. Semeval-2014 task 7: Analysis of clinical text. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pages 54–62, 2014.

[8] Ahmad P Tafti, Eric LaRose, Jonathan Badger, Zhan Ye, Neel Shimpi, Fereshteh S Bashiri, Elham Sagheb, Huong McLean, David Page, and Peggy Peissig. Big data deep neural network to analyze adverse vaccine reactions. AMIA 2018 Informatics Summit, 2018.

[9] Hyunwoo J Kim, Nagesh Adluru, Heemanshu Suri, Baba C Vemuri, Sterling C Johnson, and Vikas Singh. Riemannian nonlinear mixed effects models: Analyzing longitudinal deformations in neuroimaging. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[10] Dinggang Shen, Guorong Wu, and Heung-Il Suk. Deep learning in medical image analysis. *Annual review of biomedical engineering*, 19:221–248, 2017.

[11] Mohammad Havaei, Axel Davy, David Warde-Farley, Antoine Biard, Aaron Courville, Yoshua Bengio, Chris Pal, Pierre-Marc Jodoin, and Hugo Larochelle. Brain tumor segmentation with deep neural networks. *Medical image analysis*, 35:18–31, 2017.

[12] Ming Xu and Zeyun Yu. 3D image segmentation based on feature-sensitive and adaptive tetrahedral meshes. In *Image Processing (ICIP), 2016 IEEE International Conference on*, pages 854–858. IEEE, 2016.

[13] Looking back on the millennium in medicine. *New England Journal of Medicine*, 342 (1):42–49, 2000. doi: 10.1056/NEJM200001063420108.

[14] IBM unveils watson-powered imaging solutions for healthcare providers. [Online] http://www-03.ibm.com/press/us/en/pressrelease/51146.wss, 2016. [Accessed: 11.14.2017].

[15] IBM Watson Health. Behind the scenes: The future of cognitive computing for medical imaging. [Online] https://www.ibm.com/watson/health/imaging/, 2017. [Accessed: 11.14.2017].

[16] Kunio Doi. Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Computerized medical imaging and graphics*, 31(4):198–211, 2007.

[17] Ryohei Takahashi and Yuya Kajikawa. Computer-aided diagnosis: A survey with bibliometric analysis. *International Journal of Medical Informatics*, 101:58–67, 2017.

[18] Diana Mateus, Christian Wachinger, Selen Atasoy, Loren Schwarz, and Nassir Navab. Learning manifolds: design analysis for medical applications. In *Machine Learning in Computer-Aided Diagnosis: Medical Imaging Intelligence and Analysis*, pages 374–402. IGI Global, 2012.

[19] Paul Aljabar, Robin Wolz, and Daniel Rueckert. Manifold learning for medical image registration, segmentation, and classification. In *Machine learning in computer-aided diagnosis: Medical imaging intelligence and analysis*, pages 351–372. IGI Global, 2012.

[20] Laurens Van Der Maaten, Eric Postma, and Jaap Van den Herik. Dimensionality reduction: a comparative review. *Journal of Machine Learning Research*, 10:66–71, 2009.

[21] Karl Pearson. LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):559–572, 1901.

[22] Warren S Torgerson. Multidimensional scaling: I. theory and method. *Psychometrika*, 17(4):401–419, 1952.

[23] Trevor F Cox and Michael AA Cox. *Multidimensional scaling*. CRC press, 2000.

[24] Ronald A Fisher. The use of multiple measurements in taxonomic problems. *Annals of human genetics*, 7(2):179–188, 1936.

[25] Jeanny Herault and Christian Jutten. Space or time adaptive signal processing by neural network models. In *AIP conference proceedings*, volume 151, pages 206–211. AIP, 1986.

[26] Jerome H Friedman and John W Tukey. A projection pursuit algorithm for exploratory data analysis. *IEEE Transactions on computers*, 100(9):881–890, 1974.

[27] Edsger W Dijkstra. A note on two problems in connexion with graphs. *Numerische mathematik*, 1(1):269–271, 1959.

[28] David DeMers and Garrison W Cottrell. Non-linear dimensionality reduction. In *Advances in neural information processing systems*, pages 580–587, 1993.

[29] Bernhard Schölkopf, Alexander Smola, and Klaus-Robert Müller. Kernel principal component analysis. In *International Conference on Artificial Neural Networks*, pages 583–588. Springer, 1997.

[30] Joshua B Tenenbaum, Vin De Silva, and John C Langford. A global geometric framework for nonlinear dimensionality reduction. *science*, 290(5500):2319–2323, 2000.

[31] Vin D Silva and Joshua B Tenenbaum. Global versus local methods in nonlinear dimensionality reduction. In *Advances in neural information processing systems*, pages 721–728, 2003.

[32] Sam T Roweis and Lawrence K Saul. Nonlinear dimensionality reduction by locally linear embedding. *science*, 290(5500):2323–2326, 2000.

[33] Lawrence K Saul and Sam T Roweis. Think globally, fit locally: unsupervised learning of low dimensional manifolds. *Journal of machine learning research*, 4:119–155, Dec 2003.

[34] Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, 15(6):1373–1396, 2003. doi: 10.1162/089976603321780317.

[35] David L Donoho and Carrie Grimes. Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data. *Proceedings of the National Academy of Sciences*, 100(10):5591–5596, 2003.

[36] Matthew Brand. Charting a manifold. In *Advances in neural information processing systems*, pages 985–992, 2003.

[37] Zhenyue Zhang and Hongyuan Zha. Principal manifolds and nonlinear dimensionality reduction via tangent space alignment. *SIAM journal on scientific computing*, 26(1):313–338, 2004.

[38] Kilian Q Weinberger and Lawrence K Saul. Unsupervised learning of image manifolds by semidefinite programming. *International journal of computer vision*, 70(1):77–90, 2006.

[39] Ronald R Coifman and Stéphane Lafon. Diffusion maps. *Applied and computational harmonic analysis*, 21(1):5–30, 2006.

[40] Tong Lin and Hongbin Zha. Riemannian manifold learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(5):796–809, 2008.

[41] Li Yang. Alignment of overlapping locally scaled patches for multidimensional scaling and dimensionality reduction. *IEEE transactions on pattern analysis and machine intelligence*, 30(3):438–450, 2008.

[42] Yunqian Ma and Yun Fu. *Manifold learning theory and applications*. CRC press, 2011.

[43] Lawrence Cayton. Algorithms for manifold learning. *Univ. of California at San Diego Tech. Rep*, 12:1–17, 2005.

[44] Ali Ghodsi. Dimensionality reduction a short tutorial. *Department of Statistics and Actuarial Science, Univ. of Waterloo, Ontario, Canada*, 37:38, 2006.

[45] John A Lee and Michel Verleysen. *Nonlinear dimensionality reduction*. Springer Science & Business Media, 2007.

[46] Xiaoming Huo, Xuelei Ni, and Andrew K Smith. A survey of manifold-based learning methods. In *Recent Advances In Data Mining Of Enterprise Data: Algorithms And Applications*, pages 691–745. World Scientific, 2007.

[47] Robert Pless and Richard Souvenir. A survey of manifold learning for images. *IPSJ Transactions on Computer Vision and Applications*, 1:83–94, 2009.

[48] Alireza Sarveniazi. An actual survey of dimensionality reduction. *American Journal of Computational Mathematics*, 4(02):55, 2014.

[49] John P Cunningham and Zoubin Ghahramani. Linear dimensionality reduction: survey, insights, and generalizations. *Journal of Machine Learning Research*, 16(1):2859–2900, 2015.

[50] Shusen Liu, Dan Maljovec, Bei Wang, Peer-Timo Bremer, and Valerio Pascucci. Visualizing high-dimensional data: Advances in the past decade. *IEEE transactions on visualization and computer graphics*, 23(3):1249–1268, 2017.

[51] Yan Zhang, Zhao Zhang, Jie Qin, Li Zhang, Bing Li, and Fanzhang Li. Semi-supervised local multi-manifold isomap by linear embedding for feature extraction. *Pattern Recognition*, 76:662–678, 2018.

[52] Biao Hou, Xiangrong Zhang, Qiang Ye, and Yaoguo Zheng. A novel method for hyperspectral image classification based on laplacian eigenmap pixels distribution-flow. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(3):1602–1618, 2013.

[53] Maxime Devanne, Hazem Wannous, Stefano Berretti, Pietro Pala, Mohamed Daoudi, and Alberto Del Bimbo. 3-D human action recognition by shape analysis of motion trajectories on riemannian manifold. *IEEE transactions on cybernetics*, 45(7):1340–1352, 2015.

[54] Christian Wachinger and Nassir Navab. Manifold learning for multi-modal image registration. In *BMVC*, pages 1–12, 2010.

[55] Yizhou Yu, Chaowei Fang, and Zicheng Liao. Piecewise flat embedding for image segmentation. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1368–1376, 2015.

[56] Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps and spectral techniques for embedding and clustering. In *Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic*, pages 585–591. MIT Press, 2002.

[57] MR Brito, EL Chavez, AJ Quiroz, and JE Yukich. Connectivity of the mutual k-nearest-neighbor graph in clustering and outlier detection. *Statistics & Probability Letters*, 35(1):33–42, 1997.

[58] Fan RK Chung. *Spectral graph theory*, volume 92. American Mathematical Soc., 1997.

[59] Avinash Sharma. *Representation, Segmentation and Matching of 3D Visual Shapes using Graph Laplacian and Heat-Kernel*. PhD thesis, Institut National Polytechnique de Grenoble-INPG, 2012.

[60] Ron Larson. *Elementary linear algebra*. Nelson Education, 2016.

[61] Peter Buser. *Geometry and spectra of compact Riemann surfaces*. Springer Science & Business Media, 2010.

[62] Alston S Householder and Friedrich L Bauer. On certain methods for expanding the characteristic polynomial. *Numerische Mathematik*, 1(1):29–37, 1959.

[63] James Hardy Wilkinson. *The algebraic eigenvalue problem*. Clarendon Press Oxford, New York, NY, USA, 1988.

[64] Klaus-Jürgen Bathe and Edward L Wilson. Large eigenvalue problems in dynamic analysis. *Journal of the Engineering Mechanics Division*, 98(6):1471–1485, 1972.

[65] Klaus-Jürgen Bathe. *Finite element procedures in engineering analysis*. Prentice-Hall, 1982.

[66] PC Chowdhury. The truncated Lanczos algorithm for partial solution of the symmetric eigenproblem. *Computers & Structures*, 6(6):439–446, 1976.

[67] Richard B Lehoucq, Danny C Sorensen, and Chao Yang. *ARPACK users' guide: solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods*, volume 6. Siam, 1998.

[68] VRS Mani and S Rivazhagan. Survey of medical image registration. *Journal of Biomedical Engineering and Technology*, 1(2):8–25, 2013.

[69] Jignesh N Sarvaiya, Suprava Patnaik, and Kajal Kothari. Image registration using log polar transform and phase correlation to recover higher scale. *Journal of Pattern recognition research*, 7(1):90–105, 2012.

[70] A Ardeshir Goshtasby. *Image registration: Principles, tools and methods*. Springer Science & Business Media, 2012.

[71] Barbara Zitova and Jan Flusser. Image registration methods: a survey. *Image and vision computing*, 21(11):977–1000, 2003.

[72] Aristeidis Sotiras, Christos Davatzikos, and Nikos Paragios. Deformable medical image registration: A survey. *IEEE transactions on medical imaging*, 32(7):1153–1190, 2013.

[73] Lisa Gottesfeld Brown. A survey of image registration techniques. *ACM computing surveys (CSUR)*, 24(4):325–376, 1992.

[74] Petra A Van den Elsen, Evert-Jan D Pol, and Max A Viergever. Medical image matching-a review with classification. *IEEE Engineering in Medicine and Biology Magazine*, 12(1):26–39, 1993.

[75] JB Antoine Maintz and Max A Viergever. A survey of medical image registration. *Medical image analysis*, 2(1):1–36, 1998.

[76] Hava Lester and Simon R Arridge. A survey of hierarchical non-linear medical image registration. *Pattern recognition*, 32(1):129–149, 1999.

[77] Karl Rohr. Elastic registration of multimodal medical images: A survey. *KI*, 14(3): 11–17, 2000.

[78] Fatma El-Zahraa Ahmed El-Gamal, Mohammed Elmogy, and Ahmed Atwan. Current trends in medical image registration and fusion. *Egyptian Informatics Journal*, 17(1): 99–124, 2016.

[79] Senthil Periaswamy and Hany Farid. Medical image registration with partial data. *Medical image analysis*, 10(3):452–464, 2006.

[80] Fereshteh S Bashiri, Ahmadreza Baghaie, Reihaneh Rostami, Zeyun Yu, and Roshan M D'Souza. Multi-modal medical image registration with full or partial data: A manifold learning approach. *J. of Imaging*, 5(1), 2018.

[81] Jingfeng Han, Min Qiao, Joachim Hornegger, Torsten Kuwert, Werner Bautz, and Wolfgang Römer. Partial image data registration using stochastic optimization. *Frontiers in Simulation*, pages 644–645, 2006.

[82] Senthil Periaswamy and Hany Farid. Elastic registration with partial data. In *Biomedical Image Registration*, pages 102–111. Springer, 2003.

[83] Paul Viola and William M Wells III. Alignment by maximization of mutual information. *International journal of computer vision*, 24(2):137–154, 1997.

[84] Josien PW Pluim, JB Antoine Maintz, and Max A Viergever. Mutual-information-based registration of medical images: a survey. *IEEE transactions on medical imaging*, 22(8):986–1004, 2003.

[85] Nathan D Cahill, Julia A Schnabel, J Alison Noble, and David J Hawkes. Revisiting overlap invariance in medical image alignment. In *Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on*, pages 1–8. IEEE, 2008.

[86] Colin Studholme, Corina Drapaca, Bistra Iordanova, and Valerie Cardenas. Deformation-based mapping of volume change from serial brain MRI in the presence of local tissue contrast change. *IEEE transactions on Medical Imaging*, 25(5):626–639, 2006.

[87] Alexis Roche, Grégoire Malandain, Xavier Pennec, and Nicholas Ayache. The correlation ratio as a new similarity measure for multimodal image registration. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 1115–1124. Springer, 1998.

[88] Christoph Guetter, Chenyang Xu, Frank Sauer, and Joachim Hornegger. Learning based non-rigid multi-modal image registration using Kullback-Leibler divergence. In

International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 255–262. Springer, 2005.

[89] Rui Liao, Christoph Guetter, Chenyang Xu, Yiyong Sun, Ali Khamene, and Frank Sauer. Learning-based 2D/3D rigid registration using Jensen-Shannon divergence for image-guided surgery. In *International Workshop on Medical Imaging and Virtual Reality*, pages 228–235. Springer, 2006.

[90] Bilge Karaçali. Information theoretic deformable registration using local image information. *International journal of computer vision*, 72(3):219–237, 2007.

[91] Alexis Roche, Xavier Pennec, Grégoire Malandain, and Nicholas Ayache. Rigid registration of 3-D ultrasound with MR images: a new approach combining intensity and gradient information. *IEEE transactions on medical imaging*, 20(10):1038–1049, 2001.

[92] Wolfgang Wein, Shelby Brunke, Ali Khamene, Matthew R Callstrom, and Nassir Navab. Automatic CT-ultrasound registration for diagnostic imaging and image-guided intervention. *Medical image analysis*, 12(5):577–585, 2008.

[93] JB Antoine Maintz, Petra A van den Elsen, and Max A Viergever. 3D multimodality medical image registration using morphological tools. *Image and vision computing*, 19 (1):53–62, 2001.

[94] Eldad Haber and Jan Modersitzki. Intensity gradient based registration and fusion of multi-modal images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 726–733. Springer, 2006.

[95] Torsten Butz and Jean-Philippe Thiran. Affine registration with feature space mutual information. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 549–556. Springer, 2001.

[96] Yangming Ou, Aristeidis Sotiras, Nikos Paragios, and Christos Davatzikos. DRAMMS: Deformable registration via attribute matching and mutual-saliency weighting. *Medical image analysis*, 15(4):622–639, 2011.

[97] Mattias P Heinrich, Mark Jenkinson, Manav Bhushan, Tahreema Matin, Fergus V Gleeson, Michael Brady, and Julia A Schnabel. MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration. *Medical Image Analysis*, 16(7):1423–1435, 2012.

[98] Christian Wachinger and Nassir Navab. Entropy and laplacian images: Structural representations for multi-modal registration. *Medical Image Analysis*, 16(1):1–17, 2012.

[99] Dong Hye Ye, Jihun Hamm, Dongjin Kwon, Christos Davatzikos, and Kilian M Pohl. Regional manifold learning for deformable registration of brain MR images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 131–138. Springer, 2012.

[100] Jihun Hamm, Dong Hye Ye, Ragini Verma, and Christos Davatzikos. GRAM: A framework for geodesic registration on anatomical manifolds. *Medical image analysis*, 14(5):633–642, 2010.

[101] Liang Hu, Manning Wang, and Zhijian Song. Manifold-based feature point matching for multi-modal image registration. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 9(1):e10–e18, 2013.

[102] Gemma Piella. Diffusion maps for multimodal registration. *Sensors*, 14(6):10562–10577, 2014.

[103] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[104] Raheleh Kafieh, Hossein Rabbani, Michael D Abramoff, and Milan Sonka. Intra-retinal layer segmentation of 3D optical coherence tomography using coarse grained diffusion map. *Medical image analysis*, 17(8):907–928, 2013.

[105] Amit Singer, Radek Erban, Ioannis G Kevrekidis, and Ronald R Coifman. Detecting intrinsic slow variables in stochastic dynamical systems by anisotropic diffusion maps. *Proceedings of the National Academy of Sciences*, 106(38):16090–16095, 2009.

[106] Chang Wang and Sridhar Mahadevan. Manifold alignment using procrustes analysis. In *Proceedings of the 25th international conference on Machine learning*, ICML '08, pages 1120–1127. ACM, 2008.

[107] Chang Wang and Sridhar Mahadevan. Heterogeneous domain adaptation using manifold alignment. In *Proceedings of the International Joint Conference on Artificial Intelligence*, IJCAI'11, pages 1541–1546, 2011.

[108] Yuru Pei, Tae-Kyun Kim, and Hongbin Zha. Unsupervised random forest manifold alignment for lipreading. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 129–136, 2013.

[109] Devis Tuia, Michele Volpi, Maxime Trolliet, and Gustau Camps-Valls. Semisupervised manifold alignment of multimodal remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 52(12):7708–7720, 2014.

[110] Ben Bellekens, Vincent Spruyt, Rafael Berkvens, Rudi Penne, and Maarten Weyn. A benchmark survey of rigid 3D point cloud registration algorithm. *Int. J. Adv. Intell. Syst*, 8:118–127, 2015.

[111] B Srinivasa Reddy and Biswanath N Chatterji. An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE transactions on image processing*, 5(8):1266–1271, 1996.

[112] Stefan Klein, Marius Staring, Keelin Murphy, Max A Viergever, and Josien PW Pluim. Elastix: a toolbox for intensity-based medical image registration. *IEEE transactions on medical imaging*, 29(1):196–205, 2010.

[113] Denis P Shamonin, Esther E Bron, Boudewijn PF Lelieveldt, Marion Smits, Stefan Klein, and Marius Staring. Fast parallel image registration on CPU and GPU for diagnostic classification of alzheimer's disease. *Frontiers in neuroinformatics*, 7:50, 2014.

[114] Chris A Cocosco, Vasken Kollokian, Remi K-S Kwan, G Bruce Pike, and Alan C Evans. Brainweb: Online interface to a 3D MRI simulated brain database. In *NeuroImage*, volume 5, Copenhagen, 1997. Citeseer.

[115] Remi K-S Kwan, Alan C Evans, and G Bruce Pike. MRI simulation-based evaluation of image-processing and classification methods. *IEEE transactions on medical imaging*, 18(11):1085–1097, 1999.

[116] Remi K-S Kwan, Alan C Evans, and G Bruce Pike. An extensible MRI simulator for post-processing evaluation. In *Visualization in biomedical computing*, pages 135–140. Springer, 1996.

[117] D Louis Collins, Alex P Zijdenbos, Vasken Kollokian, John G Sled, Noor J Kabani, Colin J Holmes, and Alan C Evans. Design and construction of a realistic digital brain phantom. *IEEE transactions on medical imaging*, 17(3):463–468, 1998.

[118] Jay B West, J Michael Fitzpatrick, Matthew Yang Wang, Benoit M Dawant, Calvin R Maurer, Robert M Kessler, Robert J Maciunas, Christian Barillot, Didier Lemoine, Andre MF Collignon, et al. Comparison and evaluation of retrospective intermodality image registration techniques. In *Medical Imaging 1996: Image Processing*, volume 2710, pages 332–348. International Society for Optics and Photonics, 1996.

[119] Jay West, J Michael Fitzpatrick, Matthew Y Wang, Benoit M Dawant, Calvin R Maurer Jr, Robert M Kessler, Robert J Maciunas, Christian Barillot, Didier Lemoine, André Collignon, et al. Comparison and evaluation of retrospective intermodality brain image registration techniques. *Journal of computer assisted tomography*, 21(4): 554–568, 1997.

[120] J Michael Fitzpatrick, Jay B West, and Calvin R Maurer. Predicting error in rigid-body point-based registration. *IEEE transactions on medical imaging*, 17(5):694–702, 1998.

[121] Hsuan Chang and J Michael Fitzpatrick. A technique for accurate magnetic resonance imaging in the presence of field inhomogeneities. *IEEE Transactions on medical imaging*, 11(3):319–329, 1992.

[122] David Mattes, David R Haynor, Hubert Vesselle, Thomas K Lewellyn, and William Eubank. Nonrigid multimodality image registration. In *Medical imaging 2001: image processing*, volume 4322, pages 1609–1621. International Society for Optics and Photonics, 2001.

[123] Martin Styner and Guido Gerig. Evaluation of 2D/3D bias correction with 1+1 ES-optimization. *Rapport de recherche*, 179, 1997.

[124] Emad Omrani, Ahmad P Tafti, Mojtaba F Fathi, Afsaneh Dorri Moghadam, Pradeep Rohatgi, Roshan M D'Souza, and Zeyun Yu. Tribological study in microscale using 3D SEM surface reconstruction. *Tribology International*, 103:309–315, 2016.

[125] Lydia Ng, Sayan Pathak, Chihchau Kuan, Chris Lau, Hong-wei Dong, Andrew Sodt, Chinh Dang, Brian Avants, Paul Yushkevich, James Gee, et al. Neuroinformatics for genome-wide 3-D gene expression mapping in the mouse brain. *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, 4(3):382–393, 2007.

[126] Zhanheng Gao, Reihaneh Rostami, Xiaoli Pang, Zhicheng Fu, and Zeyun Yu. Mesh generation and flexible shape comparisons for bio-molecules. *Molecular Based Mathematical Biology*, 4(1):1–13, 2016.

[127] Germán D Sosa, Sebastián Rodríguez, Javier Guaje, Jorge Victorino, Manuel Mejía, Luz Stella Fuentes, Angelica Ramírez, and Hugo Franco. 3D surface reconstruction of entomological specimens from uniform multi-view image datasets. In *2016 XXI Symposium on Signal Processing, Images and Artificial Vision (STSIVA)*, pages 1–8, Aug 2016. doi: 10.1109/STSIVA.2016.7743319.

[128] Stefan Riehemann, Martin Palme, Peter Kuehmstedt, Constanze Grossmann, Gunther Notni, and Josef Hintersehr. Microdisplay-based intraoral 3D scanner for dentistry. *Journal of Display Technology*, 7(3):151–155, 2011.

[129] Chenglei Wu, Derek Bradley, Pablo Garrido, Michael Zollhöfer, Christian Theobalt, Markus Gross, and Thabo Beeler. Model-based teeth reconstruction. *ACM Transactions on Graphics (TOG)*, 35(6):220, 2016.

[130] Jin Xie, Yi Fang, Fan Zhu, and Edward Wong. Deepshape: Deep learned shape descriptor for 3D shape matching and retrieval. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1275–1283, 2015.

[131] Roberto Toldo, Umberto Castellani, and Andrea Fusiello. Visual vocabulary signature for 3D object retrieval and partial matching. In *Proceedings of the 2nd Eurographics conference on 3D Object Retrieval*, pages 21–28. Eurographics Association, 2009.

[132] Shuhui Bu, Zhenbao Liu, Junwei Han, Jun Wu, and Rongrong Ji. Learning high-level feature by deep belief networks for 3-D model retrieval and recognition. *IEEE Transactions on Multimedia*, 16(8):2154–2167, 2014.

[133] Yonathan Aflalo, Alexander M Bronstein, Michael M Bronstein, and Ron Kimmel. Deformable shape retrieval by learning diffusion kernels. In *International Conference*

on *Scale Space and Variational Methods in Computer Vision*, pages 689–700. Springer, 2011.

[134] Alexander M Bronstein, Michael M Bronstein, Leonidas J Guibas, and Maks Ovsjanikov. Shape google: Geometric words and expressions for invariant shape retrieval. *ACM Transactions on Graphics (TOG)*, 30(1):1, 2011.

[135] Johan WH Tangelder and Remco C Veltkamp. A survey of content based 3D shape retrieval methods. In *Proceedings Shape Modeling Applications*, pages 145–156. IEEE, 2004.

[136] Lisha Zhang, M João da Fonseca, and Alfredo Ferreira. Survey on 3D shape descriptors. Technical report, DecorAR (FCT POSC/EIA/59938/2004); Fundação para a Ciência e a Tecnologia, Lisboa, Portugal, 2007.

[137] Reihaneh Rostami, Fereshteh S. Bashiri, Behrouz Rostami, and Zeyun Yu. A survey on data driven 3D shape descriptors. *Computer Graphics Forum*, 2018. doi: 10.1111/cgf.13536.

[138] Davide Boscaini, Jonathan Masci, Emanuele Rodolà, Michael M Bronstein, and Daniel Cremers. Anisotropic diffusion descriptors. *Computer Graphics Forum*, 35(2):431–441, 2016.

[139] Alexander M Bronstein. Spectral descriptors for deformable shapes. *arXiv preprint arXiv:1110.5015*, 2011.

[140] Martin Reuter, Franz-Erich Wolter, and Niklas Peinecke. Laplace–Beltrami spectra as 'Shape-DNA' of surfaces and solids. *Computer-Aided Design*, 38(4):342–366, 2006.

[141] Jian Sun, Maks Ovsjanikov, and Leonidas Guibas. A concise and provably informative multi-scale signature based on heat diffusion. *Computer Graphics Forum*, 28(5):1383–1392, 2009.

[142] Mathieu Aubry, Ulrich Schlickewei, and Daniel Cremers. The wave kernel signature: A quantum mechanical approach to shape analysis. In *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*, pages 1626–1633. IEEE, 2011.

[143] Raif M Rustamov. Laplace-Beltrami eigenfunctions for deformation invariant shape representation. In *Proceedings of the fifth Eurographics symposium on Geometry processing*, pages 225–233. Eurographics Association, 2007.

[144] Zhanheng Gao, Zeyun Yu, and Xiaoli Pang. A compact shape descriptor for triangular surface meshes. *Computer-Aided Design*, 53:62–69, 2014.

[145] Dan Raviv, Ron Kimmel, and Alfred M Bruckstein. Graph isomorphisms and automorphisms via spectral signatures. *IEEE transactions on pattern analysis and machine intelligence*, 35(8):1985–1993, 2013.

[146] Dela De Youngster, Eric Paquet, Herna Viktor, and Emil Petriu. An isometry-invariant spectral approach for protein-protein docking. In *Bioinformatics and Bioengineering (BIBE), 2013 IEEE 13th International Conference on*, pages 1–6. IEEE, 2013.

[147] Maks Ovsjanikov, Mirela Ben-Chen, Justin Solomon, Adrian Butscher, and Leonidas Guibas. Functional maps: a flexible representation of maps between shapes. *ACM Transactions on Graphics (TOG)*, 31(4):30, 2012.

[148] Gabriel Taubin. A signal processing approach to fair surface design. In *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, pages 351–358. ACM, 1995.

[149] Uwe F Mayer. Numerical solutions for the surface diffusion flow in three space dimensions. *Computational and Applied Mathematics*, 20(3):361–379, 2001.

[150] Guoliang Xu. Discrete Laplace–Beltrami operators and their convergence. *Computer Aided Geometric Design*, 21(8):767–784, 2004.

[151] Martin Reuter, Silvia Biasotti, Daniela Giorgi, Giuseppe Patanè, and Michela Spagnuolo. Discrete Laplace–Beltrami operators for shape analysis and segmentation. *Computers & Graphics*, 33(3):381–390, 2009.

[152] Fereshteh S Bashiri, Reihaneh Rostami, Peggy Peissig, Roshan M D'Souza, and Zeyun Yu. An application of manifold learning in global shape descriptors. *arXiv*, art. arXiv:1901.02508, January 2019.

[153] Michael M Bronstein and Iasonas Kokkinos. Scale-invariant heat kernel signatures for non-rigid shape recognition. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 1704–1711. IEEE, 2010.

[154] Zhenzhong Kuang, Zongmin Li, Qian Lv, Tian Weiwei, and Yujie Liu. Modal function transformation for isometric 3D shape representation. *Computers & Graphics*, 46: 209–220, 2015.

[155] Martin Reuter, Franz-Erich Wolter, and Niklas Peinecke. Laplace-spectra as fingerprints for shape matching. In *Proceedings of the 2005 ACM symposium on Solid and physical modeling*, pages 101–106. ACM, 2005.

[156] Hermann Weyl. Über die asymptotische verteilung der eigenwerte. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 1911:110–117, 1911.

[157] Alexander M Bronstein, Michael M Bronstein, and Ron Kimmel. *Numerical geometry of non-rigid shapes.* Springer Science & Business Media, 2008.

[158] Kaleem Siddiqi, Juan Zhang, Diego Macrini, Ali Shokoufandeh, Sylvain Bouix, and

Sven Dickinson. Retrieving articulated 3-D models using medial surfaces. *Machine vision and applications*, 19(4):261–275, 2008.

[159] Mahsa Mirloo and Hossein Ebrahimnezhad. Non-rigid 3D object retrieval using directional graph representation of wave kernel signature. *Multimedia Tools and Applications*, 77(6):6987–7011, 2018.

[160] Majid Masoumi, Mahsa Rezaei, and A Ben Hamza. Global spectral graph wavelet signature for surface analysis of carpal bones. *Physics in Medicine & Biology*, 63(3): 35–34, 2018.

[161] Davide Boscaini, Jonathan Masci, Simone Melzi, Michael M Bronstein, Umberto Castellani, and Pierre Vandergheynst. Learning class-specific descriptors for deformable shapes using localized spectral convolutional networks. *Computer Graphics Forum*, 34(5):13–23, 2015.

[162] Chunyuan Li and A Ben Hamza. Spatially aggregating spectral descriptors for nonrigid 3D shape retrieval: a comparative survey. *Multimedia Systems*, 20(3):253–281, 2014.

[163] Zhouhui Lian, Afzal Godil, Benjamin Bustos, Mohamed Daoudi, Jeroen Hermans, Shun Kawamura, Yukinori Kurita, Guillaume Lavoué, Hien Van Nguyen, Ryutarou Ohbuchi, et al. SHREC11 track: Shape retrieval on non-rigid 3D watertight meshes. In *Eurographics Workshop on 3D Object Retrieval (3DOR)*, 2011.

[164] Philip Shilane, Patrick Min, Michael Kazhdan, and Thomas Funkhouser. The princeton shape benchmark. In *Shape modeling applications, 2004. Proceedings*, pages 167–178. IEEE, 2004.

[165] Yong-Jin Liu, Zhanqing Chen, and Kai Tang. Construction of iso-contours, bisectors, and voronoi diagrams on triangulated surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(8):1502–1517, 2011.

[166] Yongjie Zhang, Chandrajit Bajaj, and Guoliang Xu. Surface smoothing and quality improvement of quadrilateral/hexahedral meshes with geometric flow. *International Journal for Numerical Methods in Biomedical Engineering*, 25(1):1–18, 2009.

[167] The American Cancer Society medical and editorial content team. Key statistics for lung cancer. [Online] https://www.cancer.org/cancer/non-small-cell-lung-cancer/about/key-statistics.html, 2017. [Accessed: 09.11.2017].

[168] National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. *New England Journal of Medicine*, 365(5): 395–409, 2011.

[169] University of Rochester Medical Center. Health encyclopedia: Pulmonary nodules. [Online] https://www.urmc.rochester.edu/encyclopedia, 2017. [Accessed: 12.24.2017].

[170] Shu Ling Alycia Lee, Abbas Z Kouzani, and Eric J Hu. Random forest based lung nodule classification aided by clustering. *Computerized medical imaging and graphics*, 34(7):535–542, 2010.

[171] Xingjian Yan, Jianing Pang, Hang Qi, Yixin Zhu, Chunxue Bai, Xin Geng, Mina Liu, Demetri Terzopoulos, and Xiaowei Ding. Classification of lung nodule malignancy risk on computed tomography images using convolutional neural network: A comparison between 2D and 3D strategies. In *Asian Conference on Computer Vision*, pages 91–101. Springer, 2016.

[172] Wenqing Sun, Bin Zheng, and Wei Qian. Computer aided lung cancer diagnosis with deep learning algorithms. In *Medical Imaging 2016: Computer-Aided Diagnosis*, volume 9785, page 97850Z. International Society for Optics and Photonics, 2016.

[173] Bruno Rodrigues Froz, Antonio Oseas de Carvalho Filho, Aristófanes Corrêa Silva, Anselmo Cardoso de Paiva, Rodolfo Acatauassú Nunes, and Marcelo Gattass. Lung

nodule classification using artificial crawlers, directional texture and support vector machine. *Expert Systems with Applications*, 69:176–188, 2017.

[174] David Ost, Alan M Fein, and Steven H Feinsilver. The solitary pulmonary nodule. *New England Journal of Medicine*, 348(25):2535–2542, 2003.

[175] Agnieszka Choromańska and Katarzyna J Macura. Evaluation of solitary pulmonary nodule detected during computed tomography examination. *Polish journal of radiology*, 77(2):22, 2012.

[176] Ayman El-Baz, Garth M Beache, Georgy Gimel'farb, Kenji Suzuki, Kazunori Okada, Ahmed Elnakib, Ahmed Soliman, and Behnoush Abdollahi. Computer-aided diagnosis systems for lung cancer: challenges and methodologies. *International journal of biomedical imaging*, 2013, 2013.

[177] Yutong Xie, Jianpeng Zhang, Sidong Liu, Weidong Cai, and Yong Xia. Lung nodule classification by jointly using visual descriptors and deep features. In *Medical Computer Vision and Bayesian and Graphical Models for Biomedical Imaging*, pages 116–125. Springer, 2017.

[178] Ayman El-Baz, Georgy Gimel'farb, Robert Falk, and Mohamed El-Ghar. Appearance analysis for diagnosing malignant lung nodules. In *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on*, pages 193–196. IEEE, 2010.

[179] Ayman El-Baz, Ahmed Soliman, Patrick McClure, G Gimel'farb, M Abo El-Ghar, and Robert Falk. Early assessment of malignant lung nodules based on the spatial analysis of detected lung nodules. In *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*, pages 1463–1466. IEEE, 2012.

[180] Ayman El-Baz, Matthew Nitzken, Fahmi Khalifa, Ahmed Elnakib, Georgy Gimel'farb, Robert Falk, and Mohammed Abo El-Ghar. 3D shape analysis for early diagnosis

of malignant lung nodules. In *Proceedings of the 22nd International Conference on Information Processing in Medical Imaging*, pages 772–783. Springer-Verlag, 2011.

[181] Fangfang Han, Huafeng Wang, Guopeng Zhang, Hao Han, Bowen Song, Lihong Li, William Moore, Hongbing Lu, Hong Zhao, and Zhengrong Liang. Texture feature analysis for computer-aided diagnosis on pulmonary nodules. *Journal of digital imaging*, 28(1):99–115, 2015.

[182] Ashis Kumar Dhara, Sudipta Mukhopadhyay, Anirvan Dutta, Mandeep Garg, and Niranjan Khandelwal. A combination of shape and texture features for classification of pulmonary nodules in lung CT images. *Journal of digital imaging*, 29(4):466–475, 2016.

[183] Robert M Haralick, Karthikeyan Shanmugam, et al. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, SMC-3(6):610–621, 1973.

[184] Dennis Gabor. Theory of communication. part 1: The analysis of information. *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering*, 93(26):429–441, 1946.

[185] Li Wang and Dong-Chen He. Texture classification using texture spectrum. *Pattern Recognition*, 23(8):905–910, 1990.

[186] Fereshteh S Bashiri, Eric LaRose, Jonathan C Badger, Roshan M D'Souza, Zeyun Yu, and Peggy Peissig. Object detection to assist visually impaired people: A deep neural network adventure. In *International Symposium on Visual Computing*, pages 500–510. Springer, 2018.

[187] Reza Jalil Mozhdehi and Henry Medeiros. Deep convolutional particle filter for visual tracking. In *International Conference on Image Processing (ICIP)*, 2017.

[188] Reza Jalil Mozhdehi, Yevgeniy Reznichenko, Abubakar Siddique, and Henry Medeiros. Convolutional adaptive particle filter with multiple models for visual tracking. In *International Symposium on Visual Computing*, pages 474–486. Springer, 2018.

[189] Sattar Dorafshan, Robert J Thomas, and Marc Maguire. Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete. *Construction and Building Materials*, 186:1031–1045, 2018.

[190] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[191] Devinder Kumar, Alexander Wong, and David A Clausi. Lung nodule classification using deep features in CT images. In *12th Conference on Computer and Robot Vision (CRV)*, pages 133–138. IEEE, 2015.

[192] Kai-Lung Hua, Che-Hao Hsu, Shintami Chusnul Hidayati, Wen-Huang Cheng, and Yu-Jen Chen. Computer-aided classification of lung nodules on computed tomography images via deep learning technique. *OncoTargets and therapy*, 8, 2015.

[193] Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Geert Litjens, Paul Gerke, Colin Jacobs, Sarah J Van Riel, Mathilde Marie Winkler Wille, Matiullah Naqibullah, Clara I Sánchez, and Bram van Ginneken. Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. *IEEE transactions on medical imaging*, 35(5):1160–1169, 2016.

[194] Aiden Nibali, Zhen He, and Dennis Wollersheim. Pulmonary nodule classification with deep residual networks. *International journal of computer assisted radiology and surgery*, 12(10):1799–1808, 2017.

[195] Xiaojie Huang, Junjie Shan, and Vivek Vaidya. Lung nodule detection in CT using 3D

convolutional neural networks. In *Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on*, pages 379–383. IEEE, 2017.

[196] Wei Shen, Mu Zhou, Feng Yang, Caiyun Yang, and Jie Tian. Multi-scale convolutional neural networks for lung nodule classification. In *International Conference on Information Processing in Medical Imaging*, pages 588–599. Springer, 2015.

[197] Kui Liu and Guixia Kang. Multiview convolutional neural networks for lung nodule classification. *International Journal of Imaging Systems and Technology*, 27(1):12–22, 2017.

[198] Qi Dou, Hao Chen, Lequan Yu, Jing Qin, and Pheng-Ann Heng. Multilevel contextual 3-D CNNs for false positive reduction in pulmonary nodule detection. *IEEE Transactions on Biomedical Engineering*, 64(7):1558–1567, 2017.

[199] Ahmad P Tafti, Fereshteh S Bashiri, Eric LaRose, and Peggy Peissig. Diagnostic classification of lung CT images using deep 3D multi-scale convolutional neural network. In *2018 IEEE International Conference on Healthcare Informatics (ICHI)*, pages 412–414. IEEE, 2018.

[200] Guixia Kang, Kui Liu, Beibei Hou, and Ningbo Zhang. 3D multi-view convolutional neural networks for lung nodule classification. *PloS one*, 12(11):e0188290, 2017.

[201] Sarfaraz Hussein, Kunlin Cao, Qi Song, and Ulas Bagci. Risk stratification of lung nodules using 3D CNN-based multi-task learning. In *International conference on information processing in medical imaging*, pages 249–260. Springer, 2017.

[202] Yutong Xie, Jianpeng Zhang, Yong Xia, Michael Fulham, and Yanning Zhang. Fusing texture, shape and deep model-learned information at decision level for automated classification of lung nodules on chest CT. *Information Fusion*, 42:102–110, 2018.

[203] Mario Buty, Ziyue Xu, Mingchen Gao, Ulas Bagci, Aaron Wu, and Daniel J Mollura. Characterization of lung nodule malignancy using hybrid shape and appearance features. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 662–670. Springer, 2016.

[204] Jeremy J Erasmus, John E Connolly, H Page McAdams, and Victor L Roggli. Solitary pulmonary nodules: Part I. morphologic evaluation for differentiation of benign and malignant lesions. *Radiographics*, 20(1):43–58, 2000.

[205] Samuel G Armato, Geoffrey McLennan, Luc Bidaut, Michael McNitt-Gray, Charles Meyer, Anthony Reeves, L Clarke, et al. Data from LIDC-IDRI. the cancer imaging archive, 2015. http://doi.org/10.7937/K9/TCIA.2015.LO9QL9SX.

[206] Samuel G Armato, Geoffrey McLennan, Luc Bidaut, Michael F McNitt-Gray, Charles R Meyer, Anthony P Reeves, Binsheng Zhao, Denise R Aberle, Claudia I Henschke, Eric A Hoffman, et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. *Medical physics*, 38(2):915–931, 2011.

[207] Kenneth Clark, Bruce Vendt, Kirk Smith, John Freymann, Justin Kirby, Paul Koppel, Stephen Moore, Stanley Phillips, David Maffitt, Michael Pringle, et al. The cancer imaging archive (TCIA): maintaining and operating a public information repository. *Journal of digital imaging*, 26(6):1045–1057, 2013.

[208] Anthony P Reeves, Alberto M Biancardi, et al. Lung image database consortium (LIDC) nodule size report. [online] http://www.via.cornell.edu/lidc/, October 2011. [Accessed: 11.30.2018].

[209] Qianqian Fang and David A Boas. Tetrahedral mesh generation from volumetric binary and grayscale images. In *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*, pages 1142–1145. IEEE, 2009.

[210] Fang Qianqian. iso2mesh: a matlab-based 3D tetrahedral mesh generator. [online] http://iso2mesh.sourceforge.net, 2008. [Accessed: 11.30.2018].

[211] CGAL, Computational Geometry Algorithms Library. [Online] http://www.cgal.org, 2007. [Accessed: 11.30.2018].

[212] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. Pointnet: Deep learning on point sets for 3D classification and segmentation. *Proc. Computer Vision and Pattern Recognition (CVPR), IEEE*, 1(2):4, 2017.

[213] Giuseppe Clemente and Massimo DElia. Spectrum of the Laplace-Beltrami operator and the phase structure of causal dynamical triangulations. *Physical Review D*, 97(12): 124022, 2018.

[214] Mikhail Belkin, Jian Sun, and Yusu Wang. Discrete Laplace operator on meshed surfaces. In *Proceedings of the twenty-fourth annual symposium on Computational geometry*, pages 278–287. ACM, 2008.

[215] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An introduction to statistical learning.* Springer, 2013. doi: 10.1007/978-1-4614-7138-7.

Fereshteh S. Bashiri

Curriculum Vitae

Artificial Intelligence, Machine/Deep Learning,
Image Processing, Healthcare Informatics

Professional Summary

A graduate of Electrical Engineering with minor in Computer Science, boosted up my background in applied mathematics and programming skills. I am passionate about data science, machine learning, and artificial intelligence to solve complex industrial problems and bring the best solutions to businesses. I have 5+ years of experience in industry, academia, and research, which has built up my team-work ethics and professional presentation skills.

Education

2013 – 2019 **Ph.D., Electrical Engineering**, *University of Wisconsin-Milwaukee*, WI, USA.
Machine Intelligence for Advanced Medical Data Analysis: Manifold Learning Approach
Advisers: Dr. Zeyun Yu and Dr. Roshan M D'Souza
GPA: 4.00/4.00

2007 – 2009 **Ms.C., Electrical Engineering**, *Sharif University of Technology*, Tehran, Iran.
Thesis: Quantum-Dot Cellular Automata (QCA): Theory and Application
GPA: 17.06/20

2002 – 2007 **Bs.C., Electrical Engineering**, *Shahid Behshti University*, Tehran, Iran.
Thesis: Hardware Implementation of an OCR Algorithm
GPA: 17.12/20

Publications

[1] Jonathan C Badger, Eric LaRose, John Mayer, **Fereshteh S Bashiri**, David Page, and Peggy Peissig. Machine learning for phenotyping opioid overdose events. *Journal of Biomedical Informatics*, 2019. Under review.

[2] Reihaneh Rostami, **Fereshteh S Bashiri**, Behrouz Rostami, and Zeyun Yu. A survey on data-driven 3D shape descriptors. *Computer Graphics Forum*, 2018.

[3] TR Schaid, AH Abdelhafeez, M Ranji, RB Love, SH Audi, S Kaul, **FS Bashiri**, M Masoudi-Motlagh, F Salehpoor, E Jacobs, and JC Densmore. Surface fluorescence studies of tissue mitochondrial redox state in ex-vivo lung perfusion. In *Shock*, volume 41, pages 47–47, 2014.

[4] Ahmad P Taft, Mehdi Assefi, Eric LaRose, Jonathan C Badger, Zhang Ye, Neel Shimpi, **Fereshteh S Bashiri**, Elham Sagheb, Houng McLean, David Page, and Peggy Peissig. Big data deep neural network to analyze adverse vaccine reaction. In *AMIA Informatics Summit*, 2018.

[5] Ahmad P Taft, **Fereshteh S Bashiri**, Eric LaRose, and Peggy Peissig. Diagnostic classification of lung CT images using deep 3D multi-scale convolutional neural network. In *2018 IEEE International Conference on Healthcare Informatics (ICHI)*, pages 412–414. IEEE, 2018.

[6] **Fereshteh S Bashiri**, Jonathan C Badger, Roshan M D'Souza, Zeyun Yu, and Peggy Peissig. Lung nodule classification using combined deep and spectral 3D shape features. In *International Symposium on Mathematical and Computational Oncology (ISMCO)*. Springer, 2019. Under review.

[7] **Fereshteh S Bashiri**, Ahmadreza Baghaie, Reihaneh Rostami, Roshan M D'Souza, and Zeyun Yu. Multi-modal medical image registration with full or partial data: A manifold learning approach. *Journal of Imaging*, 5(1), 2019.

[8] **Fereshteh S Bashiri**, Eric LaRose, Jonathan C. Badger, Roshan M. D'Souza, Zeyun Yu, and Peggy Peissig. Object detection to assist visually impaired people: A deep neural network adventure. In *International Symposium on Visual Computing (ISVC)*, pages 500–510. Springer, 2018.

[9] **Fereshteh S Bashiri**, Eric LaRose, Peggy Peissig, and Ahmad P Tafti. MCIndoor20000: a fully-labeled image dataset to advance indoor objects detection. *Data in Brief*, 17:71–75, 2018.

[10] **Fereshteh S Bashiri**, Reihaneh Rostami, Peggy Peissig, Roshan M D'Souza, and Zeyun Yu. An application of manifold learning in global shape descriptors. *arXiv preprint arXiv:1901.02508*, 2019.

[11] **Fereshteh S Bashiri**, Hesam Tavassoli, Rahim Faez, and Saeed Bagheri Shooraki. A novel 3-input and/or gate in quantum-dot cellular automata with single clock zone and minimum area. In *International Semiconductor Device Research Symposium*, 2009.

[12] **Fereshteh S Bashiri**, Hesam Tavassoli, Rahim Faez, and Saeed Bagheri Shooraki. Quantum-dot cellular automata counter with start/stop and reset inputs. In *International Semiconductor Device Research Symposium*, 2009.

Professional Experience

Sep. 2017 Present	**Research Specialist, MS Informatics**, *Marshfield Clinic Research Institute*, WI, [**Link**]. Medical image analysis, machine and deep learning, and their application in health informatics
Summer 2017	**Graduate Research Assistant Intern**, *Marshfield Clinic Research Institute*, WI, [**Link**]. MCIndoor: A Computer Vision Framework To Assist Navigation of Visually-impaired People
Mar. 2012 Jun. 2013	**Research Engineer**, *Sherkat Kontorsazi Iran (SKI)*, Tehran, Iran. Design and mass production of digital electricity meters Research and development of various parts of electricity meters including metrology system, RTC and metering calibration methods, and test and validation procedure
Feb. 2012 Aug. 2012	**Electronic and Microcontroller Firmware Engineer**, *Daygan Gostaran Sanat (DGS) Co.*, Tehran, Iran – Part time. Design and production of electrical equipment Smart Detection Security System, Automatic Vehicle Locator, Long distance data transmission system through radio frequency, Temperature and Humidity Measuring System

Sep. 2011 Jul. 2013	**Electronic and Firmware Engineer**, *Abzar Teb Iran*, Tehran, Iran – Part time. Design and production of hospital autoclaves for sterilizing surgery room equipment The controller in accordance with international autoclave standards, Temperature and pressure monitoring system and data logger
May 2010 Jan. 2012	**Electronic and Firmware Engineer**, *Hoorpendar C.T.I*, Tehran, Iran. Design and production of industrial and electrical devices in the field of power, gas & oil and water & waste such as: An isolated 8-channel temperature monitoring system, Industrial Battery charger
Sep. 2009 May 2010	**Research Engineer**, *Iran Telecommunication Research Center*, Tehran, Iran. Preparing action plan and project RFPs for customization of nano-technology in optical communica- tions
Jul. 2003 Aug. 2007	**Research Assistant**, *Shahid Beheshti University, Department of Electrical Engineering*, Tehran, Iran. Developing Optical Character Recognition (OCR) algorithm on Persian language: Automatic detection by extraction of primitives (Jan. 2007 – Aug. 2007), Training the system by taking projections from the image (Jul. 2003 – Feb. 2005)

Honors and Awards

2013 – 2017	**Dean's Fellowship**. University of Wisconsin-Milwaukee, WI, USA
Spring 2016	**GE Healthcare Research Poster Competition, 3rd Place**. University of Wisconsin-Milwaukee, WI, USA
Jan. 2016 May 2017	**Graduate Research Assistant – GE Healthcare**. University of Wisconsin-Milwaukee, WI, USA
2010	**Research Grant**. Iran Nano Technology Initiative Council, Tehran, Iran

Technical Skills

Software Skills	Matlab, LaTeX, Weka, Code Vision, Labview, Keil, ModelSim
Programming	C, Python, VHDL, Verilog, and Pascal
Microcontrollers	Atmel family (including AVR and ATXmega), Z80 family
Industrial modules	GPS Module (u-blox), GSM/GPRS module (Siemens MC55), ZigBit module (Microchip MRF24J40), different types of HMIs (such as Beijer, Schneider Electric and Panel Master)

Teaching Experience

Jan. 2015 May 2017	**Matlab Tutor**, *University of Wisconsin-Milwaukee*, USA.
Sep. 2014 Dec. 2015	**Electrical Circuits II (Lab)**, *University of Wisconsin-Milwaukee*, USA.
Jan. 2005 May 2005	**Digital Logic Circuits**, *Shahid Beheshti University*, Iran.