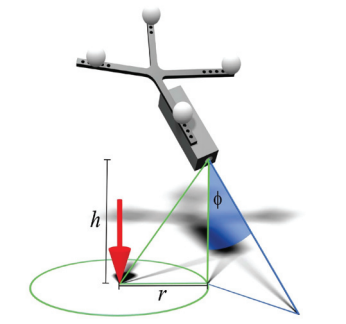
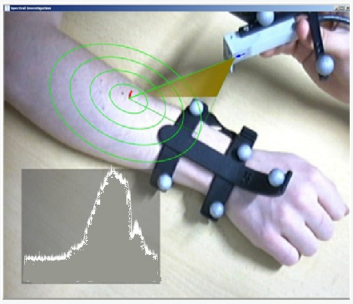
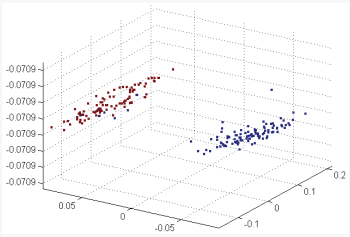
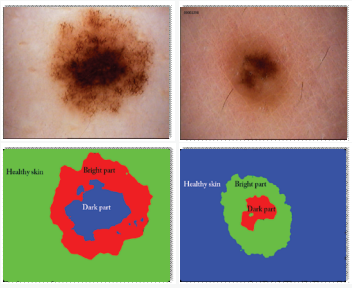
Computer Aided Medical Procedures
(CAMP)
Prof. Dr. Nassir Navab









Dissertation

Towards Computer–Aided Diagnosis of Pigmented Skin Lesionson

Asad Ali Safi

Fakultät für Informatik
Technische Universität München

TECHNISCHE UNIVERSITÄT MÜNCHEN

Chair for Computer-Aided Medical Procedures & Augmented Reality

Towards Computer–Aided Diagnosis of Pigmented Skin Lesions

Asad Ali Safi

Vollständiger Abdruck der von der Fakultät für Informatik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. Klaus A. Kuhn

Prüfer der Dissertation:

1. Univ.-Prof. Dr. Nassir Navab
2. apl. Prof. Dr. Alexander Horsch

Die Dissertation wurde am 15.02.2012 bei der Technischen Universität München eingereicht und durch die Fakultät für Informatik am 08.05.2012 angenommen.

Abstract

Skin cancer is one of the most frequently encountered types of cancer in the Western world. According to the Skin Cancer Foundation Statistics, one in every five Americans develops skin cancer during his/her lifetime. Today, the incurability of advanced cutaneous melanoma raises the importance of its early detection. Since the differentiation of early melanoma from other pigmented skin lesions is not a trivial task, even for experienced dermatologists, computer aided diagnosis could become an important tool for reducing the mortality rate of this highly malignant cancer type.

In this thesis, a computer aided diagnosis system based on machine learning is proposed in order to support the clinical use of optical spectroscopy and dermatoscopy imaging techniques for skin lesions quantification and classification. The thesis is divided into two parts; the first part focuses on a feasibility study of optical spectroscopy. To this end, data acquisition protocols for optical spectroscopy are defined and detailed analysis of feature vectors is performed. Different techniques for supervised and unsupervised learning are explored on clinical data, collected from patients with malignant and benign skin lesions. A mole mapping technique is proposed for hand-held optical spectroscopy devices with tracking where spectral information is acquired synchronously with position and orientation. Furthermore, an augmented reality guidance system is presented which allows to find a previously examined point on the skin with an accuracy of 0.8 $[mm]$ and 5.0 $[deg]$ (vs. 1.6 $[mm]$ and 6.6 $[deg]$ without guidance).

The second part is based on modeling the visual assessment of the dermatologist. To this end, detailed feature sets are derived based on the well-known diagnostic rules in dermatology, such as the ABCD rule. Several supervised and unsupervised classification methods; i.e *k*-Nearest Neighbors, Logistic Regression, Artificial Neural Networks, Decision Trees, and SVM, have been tested in combination with the developed feature extraction technique. Therefore, a dermatoscopy database consisting of 42.911 patient datasets is utilized which are acquired in routine check-ups and show skin lesions of different grades.

The contributions of this work are twofold. The feasibility study of the optical spectroscopy demonstrates the requirements for its clinical use and suggests that it can improve the diagnostic accuracy when utilized in combination with other imaging techniques such as multi-spectral imaging. The second contribution is the development of new feature vectors based on the modeling of expert's visual perception. This results in high classification accuracy of several skin lesions. Thus, this thesis presents a step towards computer aided solutions in order to improve dermatological diagnosis in the near future.

Keywords: Skin Cancer, Optical Spectroscopy, Dermatoscopy, Machine Learning, Supervised Learning, Unsupervised Learning.

Zusammenfassung

Hautkrebs ist eine der am weitesten verbreiteten Krebsarten in der westlichen Welt. Nach Statistiken der Skin Cancer Foundation entwickelt jeder fünfte Amerikaner Hautkrebs im Laufe seines Lebens. Die Unheilbarkeit des fortgeschrittenen kutanen Melanoms verstärkt heutzutage die Bedeutung der frühzeitigen Erkennung. Da es selbst für erfahrene Dermatologen schwierig ist, das Melanom in einer frühen Phase von anderen pigmentierten Hautveränderungen zu unterscheiden, könnte die computergestützte Diagnose zu einem wichtigen Werkzeug werden, um die Sterblichkeitsrate dieser höchst bösartigen Krebsart zu verringern.

In dieser Arbeit wird ein computergestütztes Diagnosesystem vorgestellt, das auf maschinellem Lernen beruht und die klinische Anwendung von optischer Spektroskopie und Dermatoskopie für die Quantifikation und Klassifikation von Hautveränderungen unterstützt. Die Arbeit besteht aus zwei Teilen; im ersten Teil findet sich eine Anwendbarkeitsstudie zur optischen Spektroskopie. Hierzu werden Aufnahmeprotokolle definiert und Vektoren zur Merkmalsextraktion im Detail untersucht. Verschiedene Methoden des überwachten und unüberwachten maschinellen Lernens werden auf klinischen Daten angewendet, die von Patienten mit gutartigen und bösartigen Hautveränderungen stammen. Eine Mole-Mapping-Methode für Spektroskopie-Handgeräte wird vorgestellt, bei der mithilfe von optischem Tracking die Spektralinformation synchron mit der Position und Orientierung des Geräts erfasst wird. Außerdem wird ein Unterstützungssystem basierend auf Augmented Reality präsentiert, das es ermöglicht, einen vorher untersuchten Punkt auf der Haut mit einer Genauigkeit von 0,8 [*mm*] und 5,0 [*Grad*] wiederzufinden (gegenüber 1,6 [*mm*] und 6,6 [*Grad*] ohne Unterstützung).

Im zweiten Teil wird die visuelle Beurteilung durch den Hautarzt modelliert. Hierzu werden detaillierte Sätze von charakteristischen Merkmalen auf Basis bekannter dermatologischer Diagnoseregeln hergeleitet, wie der ABCD Regel. Mehrere überwachte und unüberwachte Klassifizierungsmethoden werden in Verbindung mit der entwickelten Merkmalsextraktion erprobt (*k*-Nearest Neighbors, Logistische Regression, Neuronale Netze, Entscheidungsbäume und SVM). Zu diesem Zweck wird eine Dermatoskopie-Datenbank verwendet mit 42,911 Patientendatensätzen, die bei Routineuntersuchungen entstanden sind und Hautveränderungen verschiedenen Grades beinhalten.

Diese Arbeit bietet zweierlei Beiträge. Die Anwendbarkeitsstudie behandelt die Anforderungen für den klinischen Einsatz von Spektroskopie und zeigt auf, dass das Verfahren die diagnostische Genauigkeit verbessern kann, wenn es in Verbindung mit anderen bildgebenden Verfahren zum Einsatz kommt, z.B. Multispektralkameras. Der zweite Beitrag besteht aus der Entwicklung neuartiger Merkmalsvektoren durch die Modellierung der visuellen Wahrnehmung eines Arztes. Es ergibt sich eine hohe Klassifizierungsgenauigkeit für mehrere Arten von Hautveränderungen. Die Arbeit stellt also einen Schritt in Richtung computergestützter medizinischer Lösungen dar zur Verbesserung der dermatologischen Diagnose in naher Zukunft.

Keywords: Skin Cancer, Optical Spectroscopy, Dermatoscopy, Machine Learning, Supervised Learning, Unsupervised Learning.

Acknowledgments

First of all, I want to thank God for giving me the life to complete an important step in my professional and personal development.

I would like to express my deepest gratitude to my supervisor, Prof. Dr. Nassir Navab, for his extraordinary guidance, valuable encouragement, and thoughtful supervision. I feel honored to work with such a great human being with an outstanding vision for novel technological ideas. I thank very much to my group leader, Dr. Tobias Lasser, for sharing with me his knowledge and valued time. I also want to thank my collaboration partners, Dr. med. Mahzad Ziai and Prof. Dr. Johannes Ring at Klinikum Rechts der Isar, Munich.

I am very grateful to all my colleagues at the chair of Computer Aided Medical Procedures for their great support and friendship over the last years of my PhD. My special thanks to my Latin-American colleagues Victor Castañeda, Diana Mateus, and José Gardiazabal and Turkish friend Dzhoshkun Ismail Shakir for their affection. I wish to thank my best friend Anabel Martin Gonzalez for her unconditional love, care, and moral support. I also want to thank Martin Horn and Martina Hilla for their kindness, friendship, and all the technical and administrative support. The group has been a source of friendships as well as good advice and collaboration.

I would like to thank my family for all their love and encouragement. For my parents, Abdul Hanan Safi and Malika Safi, who raised me with love and supported me in all my pursuits. For my brothers, Nadeem and Naeem, who made me feel a loved brother. For my sisters Sema, Azra and Huma who are lights in the family. For the constant love and support of my friends at all time, Sabookh, Nassir, Asif, Ahmad, Shakir and Sarfaraz. Thank you all.

I would like to express my thanks to COMSATS Institute of Information Technology (CIIT) Abbottabad, Pakistan for granting me ex-Pakistan leave for PhD studies.

I gratefully acknowledge the generous support of my funding sources that made my Ph.D. work possible, the Higher Education Commission (HEC) Pakistan and the German Academic Exchange Service (DAAD) Germany.

Finally, I thank the Technische Universität München for the opportunity of becoming part of a great university.

Asad Ali Safi

Contents

List of Figures

List of Tables

Part I.

Introduction and Theory

CHAPTER 1

Introduction

SKIN cancer is among the most frequent types of cancer and one of the most malignant tumors. The incidence of melanoma in the general population is increasing worldwide [107], especially in countries where the ozone layer is thinning. Its incidence has increased faster than that of almost all other cancers, and the annual rates have increased on the order of 3% to 7% in the fair–skinned population in recent decades [107]. Currently, between 2 and 3 million non–melanoma skin cancers and 132.000 melanoma skin cancers occur globally each year. One in every three cancers diagnosed is a skin cancer, and according to the Skin Cancer Foundation statistics, one in every five Americans will develop skin cancer during their lifetime [124]. Because advanced cutaneous melanoma is still incurable, early detection, by means of accurate screening, is an important step toward mortality reduction. The differentiation of early melanoma from other pigmented skin lesions (e.g. benign neoplasms that simulate melanoma) is not trivial, even for experienced dermatologists. In several cases, primary care physicians seem to underestimate melanoma in its early stage [126] which attracted the interest of many researchers, and lead to the development of systems for automated detection of malignancy in skin lesions.

At present most dermatologists rely on their experience of visual assessment to distinguish benign and malign skin lesions [91] like pigmented nevi, seborrhoeic keratosis or basal cell carcinoma and malignant melanoma, as well as requiring pathology of the affected skin. To complicate matters, Cutaneous T-Cell Lym-

phoma (CTCL) is a blood cancer type with symptoms that are exhibited as skin lesions as well. Again a timely diagnosis and staging is very crucial for a successful treatment [83]. The experienced dermatologist relies initially on pattern recognition, second on history, and later on laboratory parameters. Generally, Physicians follow the "ABCD" criteria; Asymmetry, Border irregularity, Color variation, and Diameter, in their assessment [4][82][95][20][145].

However, many particular aspects of the skin cannot be evaluated effectively with the naked eye e.g. morphology, as skin is composed of many superimposed layers, with different characteristics, properties and functions that cannot be differentiated by the naked eye but are clearly delineated by imaging methods. Advances in digital dermoscopy, microscopy, imaging, and photography have formed an impressive arsenal with which dermatologists can better diagnose [51][59][114][1][60][12][132][142].

New technologies to assist the dermatologists in identifying and diagnosing skin lesion, such as hand-held magnification devices and computer-aided image analysis. Colored image processing methods have been introduced for detecting the melanoma [34] which focused on non-constant visual information of skin lesions. Neural network diagnosis of skin lesion has been applied by classifying extracted features from digitized dermoscopy images of lesions [98] [131]. The extracted features are based on geometry, colors, and texture of the lesions, involving complex image processing techniques. Many other attempts have been made to automate the detection and classification of melanoma from the digital color and surface reflectance images [132][14][45][12][144]. Those attempts involve the initial segmentation of the skin lesion from the surrounding skin followed by the calculation of classification features [52][13][131][77][146][161]. Accurate description and measurement of image features cannot be achieved without accurate image segmentation [80]. Therefore, a wide range of algorithms have been proposed in the past for color image segmentation [86], broadly categorized as pixel-based segmentation [152], region-based segmentation and edge detection [153]. However, in the case of optical spectral reflectance images, the research is still limited due to the recent introduction of the imaging technology in dermatology. The contributions of this thesis is two fold: first, a feasibility study of spectroscopy as a tool to aid the diagnosis of skin lesions is performed. Second, modeling of the visual perception of dermatologist experts a new feature extraction method and performance comparison among different classification models have been developed. In next sections we give more details.

1.1. Problem definition

1.1.1. Optical spectroscopy

One of the substantial features for the diagnosis of malignant melanoma is the skin lesion color [49]. In most of the related research, skin lesion color was investigated to disintegrate malignant melanoma lesions from benign lesions in clinical images [122]. Human skin is a variegated surface, with fine scale geometry, which makes its appearance difficult to model. Furthermore, the conditions under which the skin surface is viewed and illuminated greatly affect its appearance.

As we know that light of different wavelengths access the skin in different depths (as shown in Figure 1.1). This fact led the researchers to evaluate pigmented lesions under specific wavelengths of light from visible spectrum to near infrared range. Through multi-spectral imaging we can capture light from frequencies beyond the visible light range which allows us to extract additional information that the human eye fails to capture with its receptors for red, green and blue. Furthermore, the spectral information can be employed for the analysis and the information retrieval about the consistence and the concentration of absorbers and reflectors in the skin. Different pigments of the skin absorb different wavelength of optical spectrum, which helps in determining the reflectance coefficient of the area of the skin.

One of the most significant features of spectral reflectance is the property that the spectral reflectance curve is based on the material composition of the object surface, color, biochemical composition and cellular structure. This property can be utilized for recognizing objects and segment regions. Currently there exist only a small number of systems, e.g. spectrophotometric intracutaneous analysis (SIA) scope [112], MelaFind [117] and SpectroShade [118], which use multispectral dermoscopic images as the inputs for subsequent computer analysis.

The best of our knowledge, the system which has already developed for the analysis of skin lesion from multispectral images, is based on the images of selected wavelength without keeping record of reflectance spectra. However, as different skin lesions can be investigated more in detail by observing their reflectance, we analyze the feasibility of spectroscopy as a tool to distinguish benign and malign skin lesions.

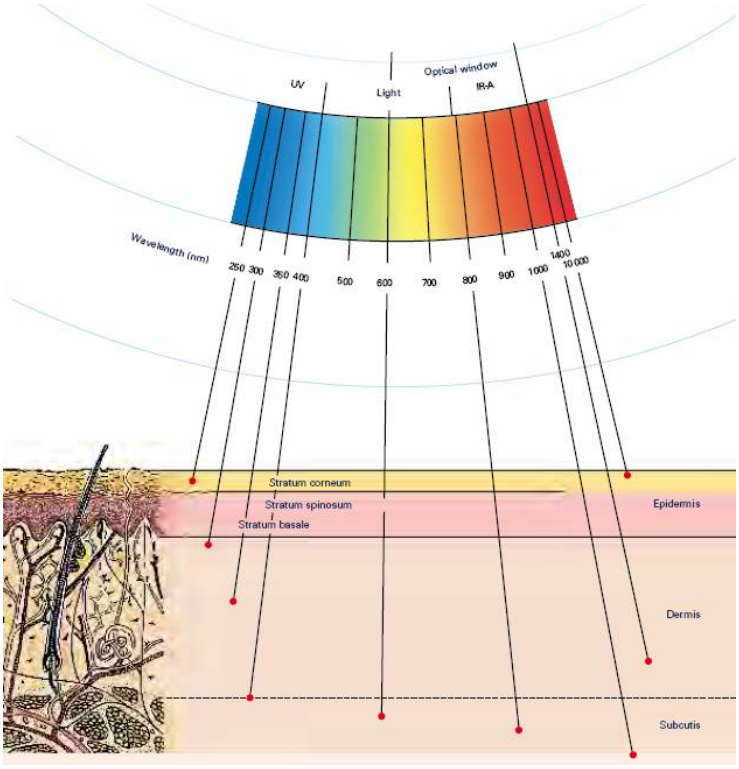


Figure 1.1.: Different wavelengths penetrates the skin to different depths. Visible light and near infrared penetraion in skin is more then other wavelengths (Image source: [43])

1.1.2. Dermoscopy

Dermoscopy consists of visual examination of skin lesion that are optically enlarged and illuminated by halogen light. This is a non–invasive in vivo technique to assists the clinician in detecting the melanoma in its early stage [22]. This technique permits the visualization of new morphologic features and thus cases facilitates early diagnosis. However, evaluation of the many morphologic characteristics is often extremely complex and subjective [131][132][142].

The Second Consensus Meeting on Dermoscopy, held in 2000, resulted in the conclusion of four algorithms as suitable means for evaluating skin lesions using dermoscopy: pattern analysis, ABCD rule, Menzies scoring method and the 7–point check list [82][130]. All four methods share some common concepts and allow for selection of specific features, which can be done with the aid of computers. The ABCD rule specifies a list of visual features associated to malignant lesions, from which a score is computed [116]. This methodology provided clinicians with a useful quantitative basis, but it did not prove to be efficient enough for clinically doubtful lesions (CDL). The main reason for this is the difficulty in visually characterizing the lesion's features. Setting an adequate decision threshold for the score is also a difficult problem. Many authors claim that these thresholds may lead to high rates of false diagnoses [95].

Collaboration of dermatologists, computer scientists and image processing specialists has led to significant automation of analysis of dermoscopic images and improvement in their classification [135][68][62]. The computerized analysis of dermoscopic images can be an extremely useful tool to measure and detect sets of features from which dermatologists derive their diagnosis. It can also be helpful for primary screening campaigns, increasing the possibility of early diagnosis of melanoma and training new practicing dermatologist. Our conclusive aim is to model visual perception of experienced dermatologist for the identification of early–stage melanoma, based on images obtained by digital dermoscopy. This would enable supervised classification of melanocytic lesions. The result of such classification procedure will separate the screened lesions into two groups. The first group corresponds to lesions that were classified with low confidence level which requires subsequent inspection by an experienced dermatologist for the final decision, while the second one corresponds to those lesions for which the confidence level is high and thus there is no need for examination by a dermatologist.

1.2. Research Objectives

The principal objective of this work is to investigate the integration of advanced imaging methods (Optical spectroscopy and digital dermoscopy) into computer aided diagnosis (CAD) of skin lesions.

- The first goal consists of the acquisition of spectral information from patient skin synchronously with position and orientation.
- Classification of spectroscopy with high accuracy based on the labelling provided by a dermatology expert.
- Improving the feasibility of spectroscopy based on the clustering and the dimensionality reduction for visualization of acquired datasets.
- Substantially important objective is to model the of visual perception of dermatologist considering dermoscopic image dataset.
- The final Objective is the performance comparison among different stat of the art models for computer-aided diagnosis system in skin lesions classification.

1.3. Contributions

The first contribution of this thesis is in the feasibility study of optical spectroscopy. To this end, spectroscopy data are acquired from the patients visiting department of dermatology for their routine checkup and its protocols are setup. Labeling of data based on the prescription of physician. Furthermore, experiments are perform for the verification of spectroscopy data. Different methods are applied for dimension reduction, classification and clustering.

Another contribution is the development of a model based on the analysis of dermatologist's visual perception. Dermoscopic images are labeled with the coordination of expert physicians which are used as a input and ground truth for classification. In literature variety of statistical and machine learning approaches for classification are available, but few comparisons among different models have been done on the same datasets.

The work presented in this thesis spawned a series of publications presented at major conferences in the field of medical imaging and medical augmented reality:

- A. Safi, T. Lasser, D. Mateus, A. Horsch, M. Ziai and N. Navab. A Comprehensives study of Advanced Computer-aided Diagnosis System for Skin Lesions Characterization . In review process at *Journal of IEEE Transactions on Biomedical Engineering* 2011

- A. Safi, M. Baust, O. Pauly, V. Castaneda, T. Lasser, D. Mateus, N. Navab, R. Hein and M. Ziai. Computer-Aided Diagnosis of Pigmented Skin Dermoscopic Images. In MICCAI Workshop on *Medical Content-based Retrieval for Clinical Decision Support* Volume: 7075, Toronto, Canada, 2011.

- A. Safi, V. Castaneda, T. Lasser, D. Mateus and N. Navab. Manifold Learning for Dimensionality Reduction and Clustering of Skin Spectroscopy Data. In Proceedings of *SPIE Medical Imaging* Volume: 7963, Pages 1192, Florida, USA, 2011.

- A. Safi, T. Lasser and N. Navab. Skin Lesions Classification with Optical Spectroscopy. In MICCAI workshop on *Medical Imaging and Augmented RealityIn (MIAR2010)*. Beijing, China, pages 411-418, 2010.

- A. Duliu, T. Lasser, A. Safi and N. Navab. Navigated Tracking of Skin Lesion Progression with Optical Spectroscopy. In Proceedings of *SPIE Medical Imaging* Volume: 7624, Pages 76243, San Diego, USA, 2010.

1.4. Overview of the Dissertation Organization

The remaining of this thesis is organized as follows: In **Chapter 2**, we provide the medical background about skin lesions, its types and clinical diagnosis methods. **Chapter 3** deals with mole mapping technique and acquisition of spectral information. **Chapter 4-5** describe supervised and unsupervised learning of spectroscopy together with a discussion on our experiments. In **Chapter 6**, we present our feature extraction method, segmentation and classification for dermatoscopic images. Comparison of different methods is presented in **Chapter 7**. Finally, in **Chapter 8**, we discuss the future work and potential improvements of the presented methods and conclusion.

CHAPTER 2

Human Skin Medical Background, Imaging Techniques and Systems

DERMATOLOGY is often termed as a visual specialty wherein a majority of diagnoses can be made by visual inspection of the skin. Diagnosis of skin disease in dermatology is largely noninvasive. The physician diagnosis is based on the anatomic distribution, color, configuration, and visible surface changes of a lesion. In some cases, a skin pathology is performed which again offers the opportunity for a microscopic visual examination of the lesion in question, but there exist limitations in the assessment of depth and size of skin lesions as well as internal features of superficial lesions. Such limitations results in need for an objective noninvasive means of assessing the skin. Digital dermatoscopic images firstly have to be parameterized for automatic classification. The deep study of skin nature has to be done before to parameterize it.

2.1. Introduction to Skin Lesions

The skin consists of an epidermis, dermis and subcutaneous fat. Many skin diseases characteristically affect a particular layer of the skin. For extraction skin or lesion optical features it is very useful to use multi layer skin model. The most common is four-layer skin model: Stratum Corneum, Epidermis, Papillary dermis and Reticular dermis. Stratum Corneum is top thin layer, which is a pro-

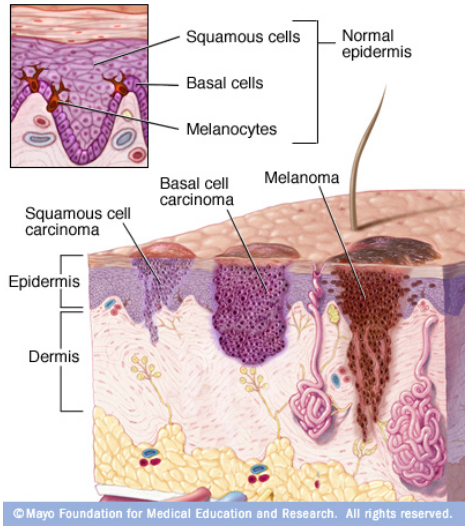


Figure 2.1.: Skin layers and cancer generation (image source: [31])

tective layer consisting of keratin-impregnated cells and it varies considerably in thickness. Apart from scattering the light, it is optically neutral.

2.1.1. Epidermis

The epidermis (Figure 2.1) is largely composed of connective tissue. It also contains the melanin producing cells, the melanocytes, and their product, melanin. In this layer there is strong absorption of blue and ultraviolet light as shown in Figure 1.1. Melanocytes absorb most of this light. It behaves like blue and ultraviolet filter, which characteristics depend on concentration of melanocytes. Within the epidermal layer there is very little scattering, with the small amount that occurs being forward directed. The result is that all light not absorbed by melanin can be considered to pass into the dermis [3].

2.1.2. Dermis

Dermis consists of two sub layers: papillary dermis and reticular dermis. Dermis itself consists of collagen fibers and, in contrast to the epidermis; it contains sensors, receptors, blood vessels and nerve ends. In papillary dermis the collagen fibers are thinner and they behave as highly backscattering layer. Any incident light is backscattered towards surface. Scattering is greater in red spectrum and

going greater to infrared. Because infrared is not absorbed by melanin and blood, this part of spectrum is best for assessing thickness of papillary dermis [3].

2.1.3. Subcutaneous fat

The subcutaneous fat is composed of adipose tissue separated by connective tissue trabeculae containing blood vessels, nerves and lymphatics. It serves both as insulation and a caloric reservoir. Its thickness also varies depending upon anatomic location, sex and body habits [3].

2.2. Types of Skin Lesions

Skin cancer is the abnormal growth of skin cells. Skin cancer begins in the cells that make up the outer layer (epidermis) of skin as shown in Figure 2.1. Manly skin cancer are three types **(i)** basal cell carcinoma, **(ii)** squamous cell carcinoma and **(iii)** melanoma [32].

Basal cell carcinoma begins in the basal cells, which make skin cells that continuously push older cells toward the surface. As new cells move upward, they become flattened squamous cells, where a skin cancer called squamous cell carcinoma can occur. Squamous cell carcinoma rarely causes further problems when identified and treated early. Untreated, squamous cell carcinoma can grow large or spread to other parts of your body, causing serious complications. A patient with squamous cell tumor has an increased chance of developing another, especially in the same skin area or nearby.

Melanoma is the most serious form of skin cancer. If it is recognized and treated early, it is almost always curable, but if it is not, the cancer can advance and spread to other parts of the body, where it becomes hard to treat and can be fatal. While it is not the most common of the skin cancers, it causes the most deaths. Melanoma originates in melanocytes as (see Figure 2.1), which is a pigment-producing cell in the skin, hair and eye that determines their color. The pigment that melanocytes make is called melanin. The major determinant of color is not the number but rather the activity of the melanocytes. Melanin production takes place in unique organelles (tiny structures within the cell) known as melanosomes. Darkly pigmented skin, hair and eyes have melanosomes that contain more melananin [110] [81] [160].

2.3. Skin Imaging Techniques

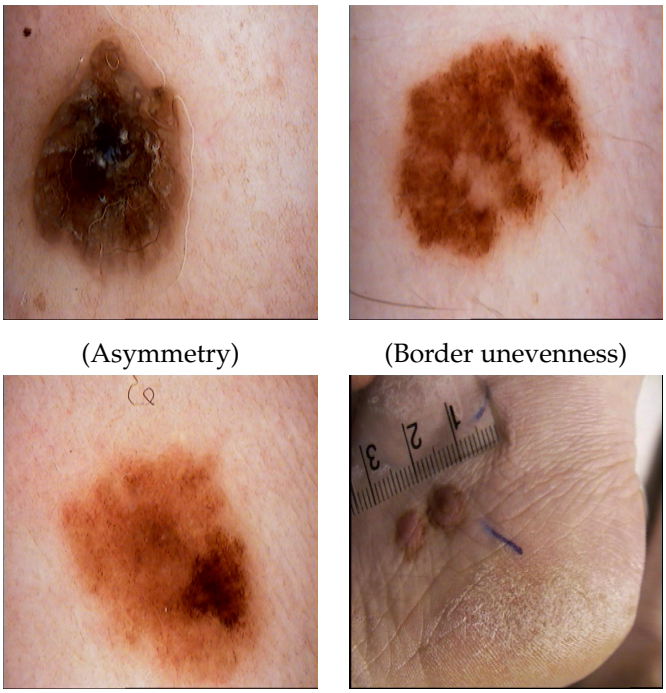
Skin imaging includes various modifications of electromagnetic wave imaging such as optical, infrared, nuclear magnetic resonance, multispectral imaging, acoustical wave imaging and mechanical wave imaging. Tomographic images i.e. 2D cross sectional images are acquired with medical imaging system. Depending on the spatial orientation of the cross section, these images depict information about the tissue over depth or any other direction. Three dimensional tissue volumes are usually imaged by acquiring a stack of consecutive 2D images [157].

Different illumination method called epiluminence microscopy (ELM, or dermoscopy) can be used in order to get the image from deeper skin layers . The light is directed straight in to these layers and reflected goes back through lesion giving more information about consistence of light absorbers in these layers [79]. Another appealing solution of getting more information from skin is using multi spectral photography, which uses narrow frequency bands of light illumination. Those images give information about consistence and concentration of absorbers and reflectors in the skin. The idea is that different pigments of skin absorb different light waves, determining the color of our skin. When those photos are made with range of light waves, we can calculate the reflectance frequency characteristics of skin. And comparing to normal skin characteristic there can be made diagnostic decisions about skin pigment consistency [81]. Some of the methods based upon the above classification of skin imaging are described in next coming sections.

2.3.1. Dermatoscopy

Dermatoscopy also known as Dermoscopy is a diagnostic technique that is used mostly in dermatology for the identification and diagnosis of skin lesions [96]. This diagnostic tool permits the recognition of structures not visible by the naked eye in other words its skin surfacing microscopy, which is noninvasive diagnostic technique for the observation of pigmented skin lesions, allowing a better visualization of surface.

Dermatologist contemplates visual signs of the lesion. The Second Consensus Meeting on Dermoscopy was held in 2000 and its main conclusions were that four algorithms: pattern analysis, ABCD rule, Menzies scoring method and the 7–point[82] check list are good ways of evaluating skin lesions using dermoscopy.



(Asymmetry) (Border unevenness)

(Color deviation) (Diameter)

Figure 2.2.: Dermoscopic images according to ABCD rule

All four methods share some common concepts and allow for selection of specific features, which can be done with the aid of computers. These methodology provided clinicians with a useful quantitative basis, but it did not prove efficient enough for clinically doubtful lesions (CDL). The main reason for this is the difficulty in visually characterizing the lesion's features. Setting an adequate decision threshold for the score is also a difficult problem; by now it has been fixed based on several years of clinical experience. Many authors claim that these thresholds may lead to high rates of false diagnoses [81] [160] [95].

2.3.1.1. The ABCD Rule

The ABCD rule was proposed in 1985 by Friedman et al. [57] as a guideline both for clinicians and laypeople to visually recognition potential melanomas in the early stages of development. The ABCD rule specifies a list of visual features (see Figure 2.2) associated to malignant lesions (Asymmetry, Border unevenness, Color deviation, Diameter and Elevation), from which a score is computed [116].

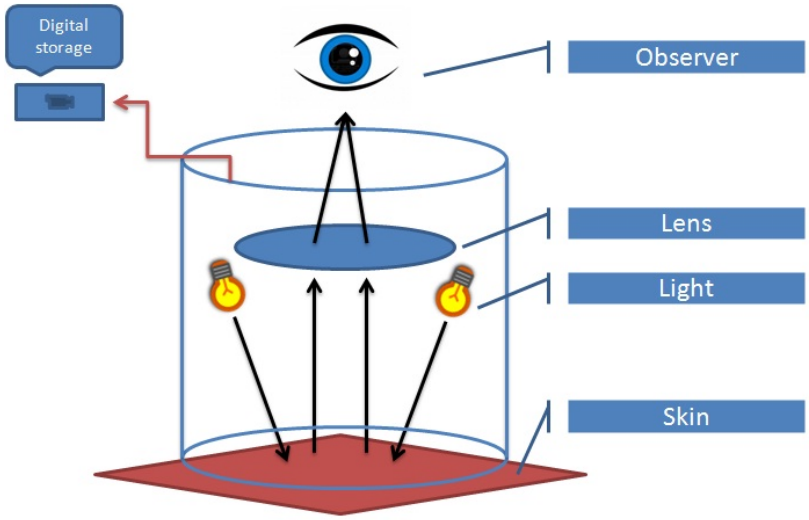


Figure 2.3.: Dermatoscopy Principle

- **A (Asymmetry):** Usually malignant skin lesions are asymmetric instead of the normal moles, which are symmetric.
- **B (Border):** Usually the melanocytic lesions have blurry and/or jagged edges.
- **C (Color):** The melanocytic lesion has different colors inside the mole.
- **D (Diameter):** The lesions does not exceed a diameter of a pencil eraser (6 mm), otherwise it is suspicious.

2.3.1.2. Dermatoscopy Principle

The functionally of dermatoscopy is similar to a magnifying lens but with the added features of an inbuilt illuminating system, a higher magnification which can be adjusted, the ability to assess structures as deep as in the reticular dermis, and the ability to record images. These phenomena are influenced by physical properties of the skin (Figure 2.3). Most of the light incident on dry, scaly skin is reflected, but smooth, oily skin allows most of the light to pass through it, reaching the deeper dermis.

Dermoscope can mainly be classified as:

- Dermoscope without image capturing facility

Table 2.1.: Commercial Dermoscopic Devices.

S.No	Device Name	Function	Manufacturing Company
1	MoleMax	ABCD score & comparison with reference bank	Derma Instruments LP (Austria)
2	Fotofinder Dermoscope	Comparison with reference bank	Edge system corpand teachscreen GmbH
3	DB_Mips	ANN and similarity classifier	Scientific Information (Italy)
4	DermGenius Ultra	ABCD score & comparison with reference bank	LINOS Photonics Inc
5	MicroDerm	ANN classifier	VisoMed (Germany)

- Dermoscope with image capturing facility
- Dermoscope with image capture facility and analytical capability.

2.3.1.3. Contact and Non-contact Technique

Two different ways of dermoscopy can be perform by contact or non-contact technique. In the contact technique, the glass plate of the instrument comes in contact with the surface of the linkage fluid applied lesion. In contrast, in the non-contact technique, there is no contact of the lens with the skin; the cross-polarized lens absorbs all the scattered light and hence allows only light in a single plane to pass through it (Figure 2.1). While the non-contact technique ensures that there are no nosocomial infections, this advantage is overshadowed by the disadvantages of decreased illumination and poor resolution [81] [160]. Table 2.1 shows some of the available commercial dermoscopic devices [121] .

A recently introduced method of ELM imaging is side-transillumination (transillumination). In this approach, light is directed from a ring around the periphery of a lesion toward its center at an angle of $45°$, forming a virtual light source at a focal point about $1cm$ below the surface of the skin, thus making the surface and subsurface of the skin translucent. The main advantage of transillumination is its sensitivity to imaging increased blood flow and vascularization and also to viewing the subsurface pigmentation in a nevus. This technique is used by a prototype

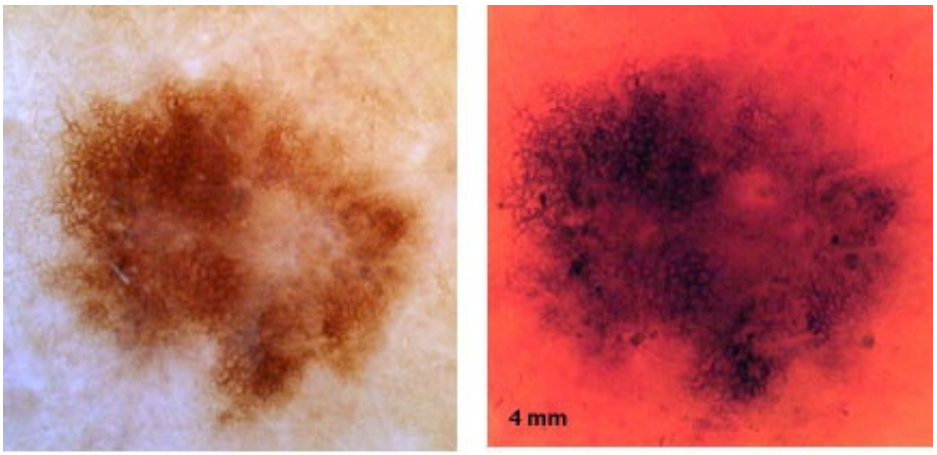


Figure 2.4.: Epiluminescence imaging of a pigmented lesion with the Dermatoscope (left) and transillumination imaging with the Nevoscope (right) (image source: [150])

device, called Nevoscope, which can produce images that have variable amount of transillumination and cross-polarized surface light [163] [127] [81] [160]. Comparison of pigmented skin lesion is shown in Figure 2.4.

2.3.2. Multispectral imaging

Surface-spectral reflectance of an object is an inherent physical property of its surface. One important role that the surface-spectral reflectance plays is to supply the physical basis for the perception of an object's color.

Another important aspect of surface-spectral reflectance is the property that the spectral reflectance curve is based on the material composition of the object surface. These can be helpful to recognize objects and segment regions in the illumination invariant way. The usual camera system with three channels of RGB has difficulty in estimating surface-spectral reflectances of objects because surface reflectances in natural scenes are spectrally high dimensional. The knowledge that light of different wavelengths penetrates the skin to different depths led investigators to evaluate pigmented lesions under specific wavelengths of light from the infrared to near UV range (Figure 1.1). Multispectral images is subdivided into absorption, transmission, and reflectivity spectroscopy. Currently there are only few systems, spectrophotometric intracutaneous analysis scope (SIAscope) and

Table 2.2.: Commercial Multispectral imaging Devices.

S.No	Device Name	Function	Manufacturing Company
1	SIAScope	Spectral Imaging	Astron Clinica (UK)
2	MelaFind	Contact spectral imaging with diagnosis Algo	Electro Optical sciences, Inc (USA)

MelaFind, which use multispectral dermoscopic images as the inputs for subsequent computer analysis [106] [81] [160]. For further details see Table 2.2

2.3.2.1. Optical Spectroscopy

Optical spectroscopy (also known as Reflectance spectroscopy) is the study of light that has been reflected or scattered from a solid, liquid, or gas. As photons enter a mineral, some are reflected from grain surfaces, some pass through the grain, and some are absorbed. Those photons that are reflected from grain surfaces or refracted through a particle are said to be scattered. Scattered photons may be detected and measured by device called spectrometer. The reflection and scattering properties of tissue in general depend on biochemical composition, cellular structure and the wavelength of light. It has been shown that malignant tissues have different optical properties from those of normal tissue [56].

2.3.2.2. Other Image Acquisition Techniques

The use of commercially available photographic cameras is also quite common in skin lesion inspection systems, particularly for telemedicine purposes [93], [139]. However, the poor resolution in very small skin lesions, i.e., lesions with diameter of less than 0.5 cm, and the variable illumination conditions are not easily handled, and therefore, high-resolution devices with low-distortion lenses have to be used. In addition, the requirement for constant image colors (necessary for image reproducibility) remains unsatisfied, as it requires real time, automated color calibration of the camera, i.e., adjustments and corrections to operate within the dynamic range of the camera and always measure the same color regardless of the lighting conditions. The problem can be addressed by using video cameras [1] that are parameterizable online and can be controlled through software [100]

Table 2.3.: Image acquisition methods along with the respective detection from literature [99]

S.No	Image Acquisition Technique	Detection
1	Tissue Microscopy	Melanoma Recognition [132][142]
2	Still CCD Camers	Wound Healing [76]
3	Video RGB Camrera	Melanoma Recognition [149][52], Tumor, crust, scale, shiny ulcer of skin lesion [153] [152], Skin erythema [122], Burn scars [151]
4	Ultraviolet illumination	Melanoma Recognition [14][28]
5	Video microscopy	Melanoma Recognition [1][60][12]
6	Multi-frequency electical impedance	Melanoma Recognition [8]
7	Side or Epi-transillumination using Nevoscope	Melanoma Recognition [163][128][159]
8	Raman Spectra	Melanoma Recognition [137]
9	Epiluminescence microscopy	Melanoma Recognition [51][59][114][13][77][20] [145]

[63]. In addition, improper amount of immersion oil or misalignment of the video fields in the captured video frame, due to camera movement, can cause either loss or quality degradation of the skin image. Acquisition time error detection techniques have been developed [63] in an effort to overcome such issues. Computed tomography (CT) images have also been used [141] in order to detect melanomas and track both progress of the disease and response to treatment. Positron emission tomography (PET) employing fluorodeoxyglucose (FDG) [129] has also been proven to be a highly sensitive and suitable diagnostic method in the staging of various neoplasms, including melanoma, complementing morphologic imaging. FDG uptake has been correlated with proliferation rate, and thus the degree of malignancy of a given tumor. MRI can also be used for tumor delineation [39]. Such methods are utilized mostly for examining the metastatic potential of a skin melanoma and for further assessment. Finally, alternative techniques such multifrequency electrical impedance [8] nor Raman spectra [137] have been proposed as potential screening methods. The electrical impedance of a biological material reflects momentary physical properties of the tissue. Raman spectra are obtained by pointing a laser beam at a skin lesion sample. The laser beam excites molecules in the sample, and a scattering effect is observed. These frequency shifts are functions of the type of molecules in the sample; thus, the Raman spectra hold useful information on the molecular structure of the sample. Table 2.3 summarizes the most common image acquisition techniques found in literature along with the respective detection goals.

Part II.

Feasibility Study of Spectroscopy

CHAPTER 3

Optical Spectroscopy based Navigation and Tracking of Skin Lesion

RECENTLY, optical spectroscopy has been proposed as a potential non-invasive screening method for skin lesions. This chapter is focused on use of spectroscopy as a diagnosis process of malignant melanoma form benign skin lesions. In the following sections we introduce a computer-assisted scheme together with the required hardware for valid spectral quantification of skin disease progression based on the tracking of optical spectroscopy probe and an augmented reality guidance system.

3.1. Introduction

Optical spectroscopy has been proposed for quantification of minimal changes in skin offering an interesting tool for monitoring skin lesions [83]. In order to keep track of the skin lesion in the follow-up of the patient, the measurements on the lesions have to be taken from the same position with the same orientation in each examination. Combining hand-held optical spectroscopy devices with advanced realtime tracking (ART) and acquiring synchronously spectral information with position and orientation, we introduce a novel computer-assisted scheme for spectral quantification of disease pro-gression. We further present an an augmented reality guidance system that allows for finding a point previously ana-

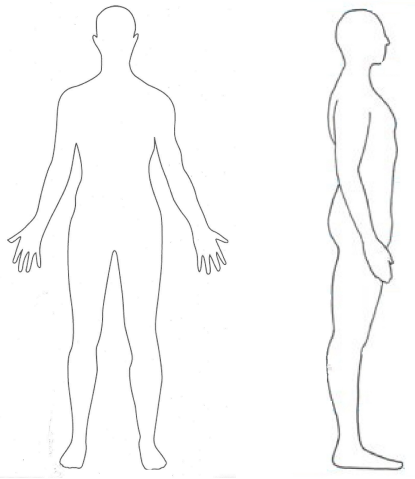
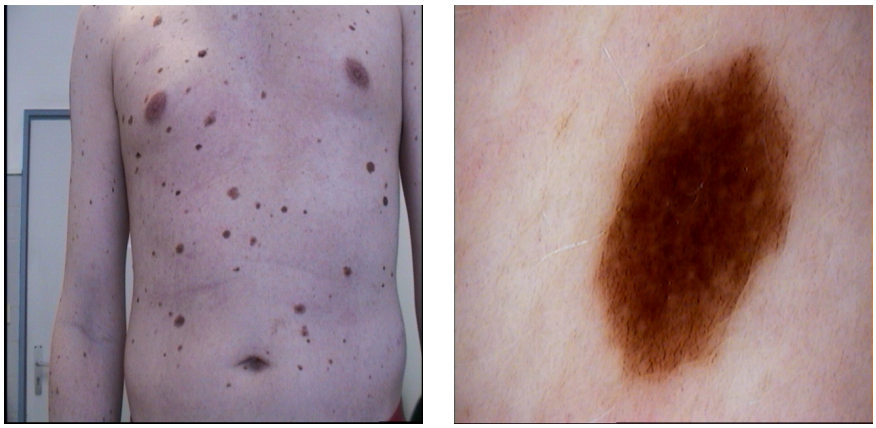


Figure 3.1.: Phantom drawing of the body to map the mole. The physician marks the location for where the image is taken

lyzed with an accuracy of $0.8[mm]$ and $5.0[deg]$ (vs. $1.6[mm]$ and $6.6[deg]$ without guidance). The intuitive guidance, as well as the preliminary results shows that the presented approach has great potential towards innovative computer-aided methods for quantification of disease progression.

3.2. Mole Mapping

The word 'mole mapping' has been used in numerous different ways. However, it usually refers to a surveillance program for those at high risk of malignant melanoma. It may include a clinical skin examination and dermoscopy to identify and evaluate lesions of concern. Mole mapping might simply involve marking spots on a phantom drawing of the body (see in Figure 3.1) to indicate the position of skin lesions of concern, particularly moles and freckles or refer to the conventional print photographs or digital images of the whole body's skin surface (see in (Figure 3.2)). These can be reviewed at a later date to see if there are any new skin lesions, or whether pre-existing skin lesions have grown or changed color or shape.



(a) (b)

Figure 3.2.: (a) Patient's whole body image (b) microscopic image of the single mole

3.3. Mole Mapping (State of the Art)

Duke research grope has focused on mole mapping[119]. heir approach is based on the use of 33 cameras to photograph from different angles to cover as much of the skin's surface as possible. There are also other commercial system; e.g. of DB-Mips, Mole-Max II and Molemax 3, which also rely on multiple video and still digital cameras for capturing whole body images. In the case of multispectral image, to the best of our knowledge, there exists currently no system for mole mapping. In next section we describe our optical spectroscopy for navigation and tracking of skin lesion in detail.

3.4. System Setup

A hand-held reflectance spectroscopy probe (StellarNet Inc., Oldsmar, FL, USA) (see Figure 3.4), consisting of $6 \times 200\mu m$ illumination fibers arrayed around one $600\mu m$ acquisition fiber as shown in Figure 3.3, was attached to an infrared optical tracking target in order to be able to determine its position and orientation in real-time. The selected tracking system consists of four ARTtrack2 infrared cameras (A.R.T. GmbH, Weilheim, Germany) positioned to be able to track a volume of $2 \times 2 \times 2 \left[m^3\right]$. According to the manufacturer the positional accuracy for such a configuration is $0.4[mm]$ with a maximum error of $1.4[mm]$ (for angle $0.002[rad]$

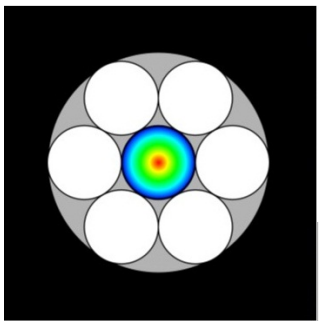


Figure 3.3.: Schematic of the fiber arrangement in the spectroscopy probe: $6 \times 200\mu m$ illumination fibers arrayed around one $600\mu m$ acquisition fiber.

and $0.007[rad]$ respectively).

A $178 - 1132[nm], 2048[px], 12bit$ CCD spectrometer (StellarNet Inc., Oldsmar, FL, USA) was connected to the acquisition fiber, and a $12[W]$ tungsten lamp was connected to the illumination fibers as a light source. The spectrometer was controlled by a data processing unit to acquire spectra synchronously with the tracking information of the probe. The data-processing unit was also used to run the augmented reality application that combined spectra, positions and orientations. An overview of the entire setup is displayed in Figure 3.4

3.4.1. Target Calibration

In order to calculate the position and orientation of the probe the fixed transformation from its tracking target to its tip was determined. For that a calibration construct was custom-built, in which the shaft of the probe is mechanically aligned with two infrared markers see Figure 3.5. By acquiring the position of these markers using the optical tracking system, both the axis and the tip of the probe can be deduced and thus the desired transformation can be calculated.

3.4.2. Proposed scheme for quantification of disease progression

Based on the literature [83],[115] there is much potential in using reflectance spectroscopy for quantitatively analyzing skin. In a disease progression setup, a valid comparison of spectra is only possible if acquisitions are taken at the same position and with the same orientation. For this we propose the use of tracking in the periodic tests. The propose system works to an accuracy of with an accuracy of

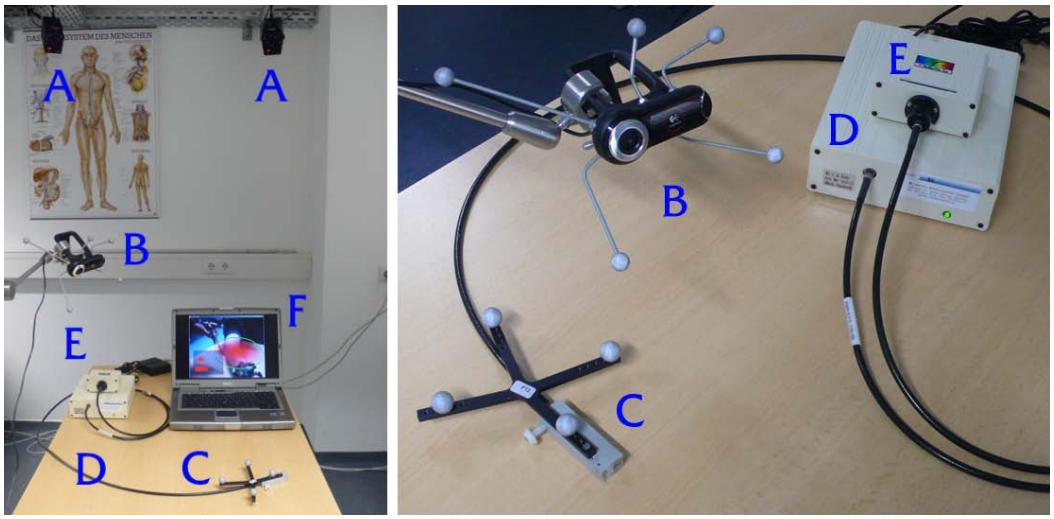


Figure 3.4.: System setup: (a) tracking cameras, (b) augmented camera, (c) tracked probe, (d) spectrometer, (e) light source, and (f) data-processing unit.

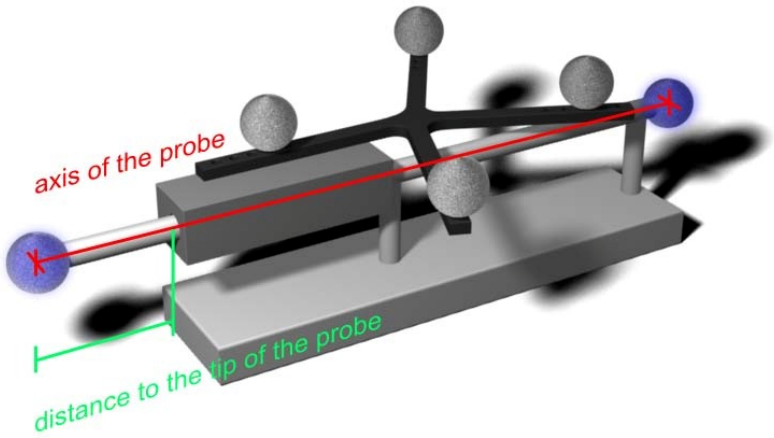


Figure 3.5.: Calibration of tip of spectroscopic probe.

$0.8[mm]$ and $5.0[deg]$ (vs $1.6[mm]$ and $6.6[deg]$ with-out guidance). The workflow of the procedure would include affixing a tracking target to the patient and then recording not only the spectra of the analyzed points, but also the position and orientation of the probe in a coordinate system fixed at the patient tracking target.

On return in the next examination, the patient tracking target can be affixed again and the positions and orientations as well as the spectra of the previous session can be loaded. The positions and orientations are used for guiding the physician to place the probe at the correct position and the spectra are used to quantify the changes of that skin location over time.

3.5. Guidance System

In order to create an intuitive guidance system for correct positioning of the probe we used augmented reality visualization. The image of a calibrated and tracked camera is augmented by different indications of the position and orientation of the previous scan as compared to the current pose of the probe.

To facilitate the three-dimensional visualization of the current error in the position, cylindrical coordinates were employed where the cylinder main axis is the vector that represents the orientation of the previous scan see Figure 3.6 (a). The error in the radius r is displayed as a circle with the previous scan position as center, whereas the error in the direction of the main axis h is shown as a line connecting the circle to the probe see Figure 3.6 (b). Additionally, h is also visualized as a cylinder using the error circle of the radius r as a base for growing up and downwards. For the angular error, the angle ϕ is used for the aperture of a cone located at the tip of the probe (Figure 3.6 (b)).

Furthermore, a 2D visualization of the radius r, height h and angle ϕ is shown as a growing/shrinking bar representation in one corner of the screen. For both the 2D and 3D visualizations color encoding is used to provide intuitive quantification of the error size, ranging from green for small errors to red for big errors.

In order to initiate the required spectral data scan both an 'on-click' acquisition and an automatic acquisition mode were implemented. In automatic acquisition mode spectral scans are acquired continuously until there is a set number of scans within a predefined tolerance interval with regard to position and orientation. After the acquisition has been completed, the best scan is selected and added to the database. The automatic acquisition mode was implemented in order to ensure

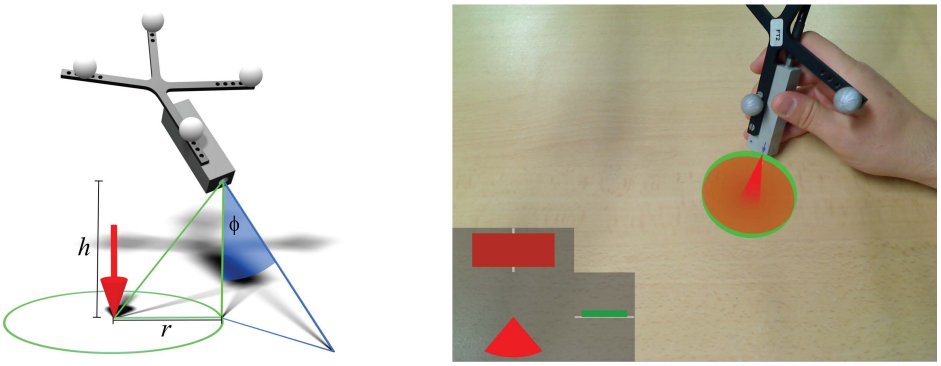


Figure 3.6.: Guidance system: (a) Visualized parameters. (b) AR guidance. (a) Mock-up of visualized parameters assuming the red arrow is the target, i.e. the position and orientation of the previous scan. (b) Augmented reality guidance, error is color-encoded in order to achieve a more intuitive impression: red means big error, green small error.

that a follow-up scan was taken close enough to the scan of interest as well as to make the process less user-dependent.

3.5.1. Analysis of Progression

Once the spectra are acquired at the right position and with the proper orientation, the application implements a spectrum comparison. For this a database of spectra is loaded and after each new acquisition, the closest spectra according to user-definable distance measures (e.g. non-Euclidean norms) are shown. This allows the physician not only to compare the spectra with the patient, but also to find similarities with previously examined patients.

3.6. Experiment Design

To validate the accuracy of the guidance procedure a phantom was custom-built with 45 target sites designed to yield easily differentiable spectral signatures via colored inks, see Figure 3.7. Four reflective markers were attached to facilitate tracking of the phantom.

Three series of experiments were conducted by three different persons. In each series, all 45 targets were acquired once to serve as reference scans. After that each of the 45 targets was acquired again three times, first using no guidance software

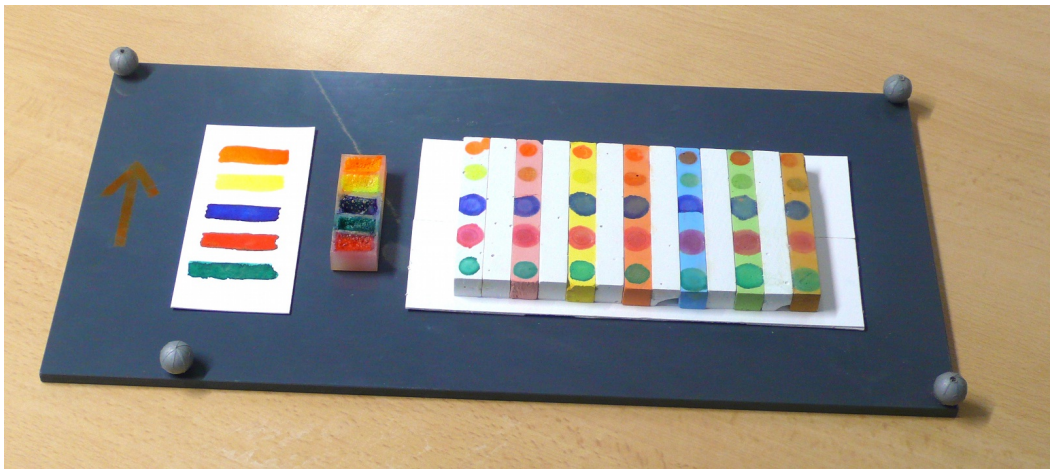


Figure 3.7.: Phantom used for guidance experiments.

at all going only from memorized positions (as is current practice in examinations) to serve as ground truth. In the second and third run, each of the targets was scanned again using the guidance system in manual acquisition mode respectively in auto acquisition mode.

3.7. Experimental results

For each of the three runs (no guidance, using guidance without and with auto acquisition) in the three series the positional accuracy was evaluated as provided by the optical tracking system. For a reference position with coordinates P_{ref} and the corresponding reacquired position with coordinates P_q the Euclidean norm $\epsilon_{pos} = \|P_{ref} - P_q\|_2$ was calculated. For the angular error the formula

$$\epsilon_\phi = cos^{-1}\left(0.5\left(tr\left(R_{ref}R_q^t\right) - 1\right)\right) \tag{3.1}$$

was used[136], where $R_{ref}, R_q$are the rotation matrices for the reference and reacquired positions as reported by the tracking system.

Furthermore the time until a target was reacquired was measured (in seconds). For validation purposes, the difference in spectra of the reference σ_{ref} and reacquired positions σ_q was calculated using

$$\epsilon_\sigma = \sum_i \left|\sigma_{ref}^i - \sigma_q^i\right| \tag{3.2}$$

Table 3.1.: Results of the guidance accuracy measurements on the phantom (from left to right: position, angle, acquisition time and spectral difference.

	$\epsilon_{pos}(mm)$	$\epsilon_{\phi}(deg)$	acq.time(sec)	ϵ_{σ}
1^{st} series				
No Guidance	2.4 ± 1.2	$14 - 6 \pm 11.4$	2.3 ± 0.5	5.5 ± 9.2
Guidance	0.6 ± 0.3	9.1 ± 8.8	9.6 ± 3.7	2.7 ± 1.2
Auto Guidance	0.7 ± 0.2	6.9 ± 8.5	15.5 ± 10.4	4.8 ± 2.3
2^{nd} series				
No Guidance	1.5 ± 0.9	3.2 ± 1.4	2.1 ± 0.3	2.2 ± 1.7
Guidance	0.6 ± 0.4	3.5 ± 1.8	7.8 ± 2.5	2.5 ± 2.4
Auto Guidance	0.8 ± 0.2	3.9 ± 1.3	16.7 ± 15.9	2.9 ± 1.4
3^{nd} series				
No Guidance	0.8 ± 0.4	2.0 ± 0.8	2.8 ± 0.4	2.1 ± 1.4
Guidance	1.2 ± 0.7	2.7 ± 1.1	8.6 ± 7.1	2.3 ± 1.5
Auto Guidance	0.8 ± 0.2	3.7 ± 1.9	16.9 ± 27.5	2.1 ± 0.9

The results of the experiments are listed in Table 3.1. The values are the mean and the standard deviation of the respective quantity, computed over the 45 targets of each run.

The results show that positional accuracy of reacquiring a reference target improves markedly using guidance. The auto acquisition mode (auto guidance) further improves on that result, especially in tightening up any outliers and thus achieving the goal of enforcing a user-independent accuracy standard. Orientational accuracy seems mostly user-dependent, and even the auto guidance fails to improve on that in this particular case the reason for that was probably a very lenient parameter setting of $7.5[deg]$ as allowable angular error. Acquisition time increases markedly when using guidance (especially with the auto acquisition mode), which is as expected. For validation the spectral readings were compared as well and confirm that the same sites were scanned again. Here the results of all methods are comparable and satisfactory.

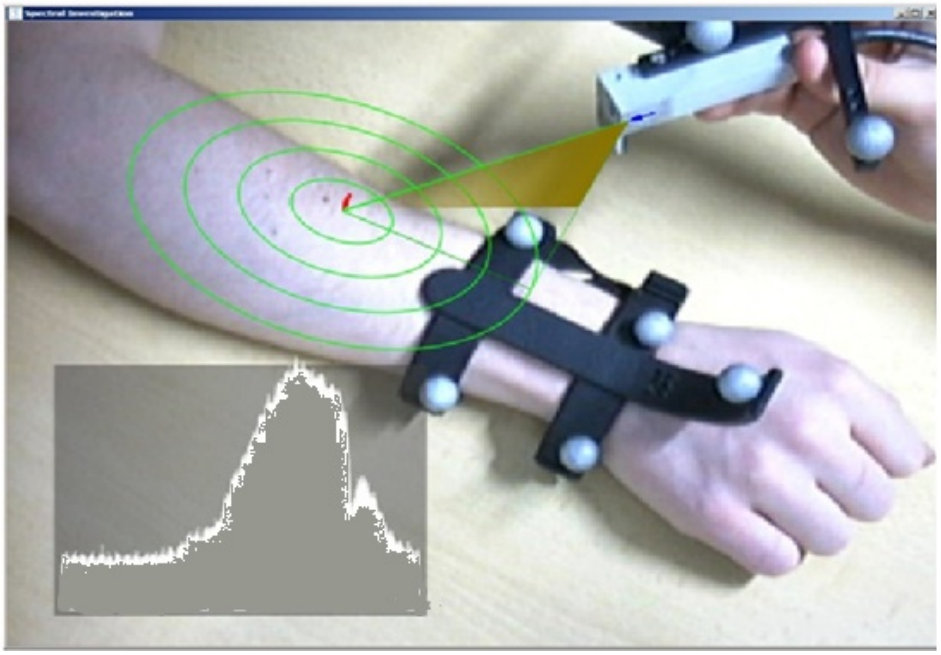


Figure 3.8.: View of navigation guidance of the prob with spectral reading from patients hand

3.8. Conclusion

This work introduced novel ideas towards valid quantification of disease progression. Among its contributions first, the combination of optical spectroscopy with tracking and the use of it for guided acquisition of spectral measurements at previously analyzed positions and with the previously acquired orientation; second, a proposed preliminary clinical methodology based on the said method; and third, initial results on the performance of a prototypic implementation.

There are, however, issues that have to be considered for its further development of this technology and its clinical use. Firstly, the proposed method includes the use of a tracking target fixed at the patient relative to which the guidance should be performed as shown in Figure 3.8. It is observed that it's very difficult to place the tracking target on the same location where once it is placed in first acquisition. As a solution in a realistic setup this can be replaced by high accuracy non-invasive patient registration methods like the ones being developed for radiation therapy and navigated surgery [92] of particular interest are methods that do not require ionizing radiation or high logistic costs. The current approach at

our lab is the evaluation of surface registration based strategies for these means. Using such a strategy the positioning of the tracking target would be arbitrary as long as it can be transformed rigidly to its position in the previous session(s) via the surface registration.

A further constraint might apply with regard to the last point: The skin surface is not allowed to deform with respect to the tracking target. A solution for this is the development of adequate deformable surface registration methods. In our research group (CAMP) one of our colleague is working on it as Ph.D. project.

In summary, despite the preliminary nature, the introduced computer-assistance method is very promising in its results and its potential applications. This work opens new ways for the 'computer-assisted intervention' community, where even with simple approaches; a big impact can be done in the quality of diagnostics, prognosis, therapy and follow-ups.

In conclusions with the introduction of computer-aided navigation for screening of skin lesions with an optical spectral imaging modality (spectroscopy) allows tracking of lesion progression over time. Extending this system with a macroscopic imaging device would enable a solution for computer-aided diagnosis, documentation and quantitative analysis of skin lesion progression.

CHAPTER 4

Skin Lesions Classification with Optical Spectroscopy

In this chapter we present a framework for acquiring spectroscopic data of skin lesions from the patients. We propose a protocol for data acquisition and define the rules for labelling the data with the assistance of dermatology experts. The experiments are performed for classification of the data using support vector machines (SVM). We report the classification results obtained from the skin lesions (benign and malignant) of 148 patients. In the following section, we describe the materials and methods used in this experiments with their results. Before explaining our approach in detail, as next, we discuss the state-of-the art of related methods.

4.1. Introduction of Spectroscopy

Spectroscopy is a new imaging technology which is increasingly used to derive significant information about tissue. Due to its multi-spectral nature, this imaging method allows to detect and classify multiple physiological changes like those associated with increased vasculature, cellular structure, oxygen consumption or edema in tumors [108], [85]. The hardware setup for data acquisition is explained in more detail in section 3.4.

Optical spectra in different wavelengths and amplitude is shown in Figure 4.1

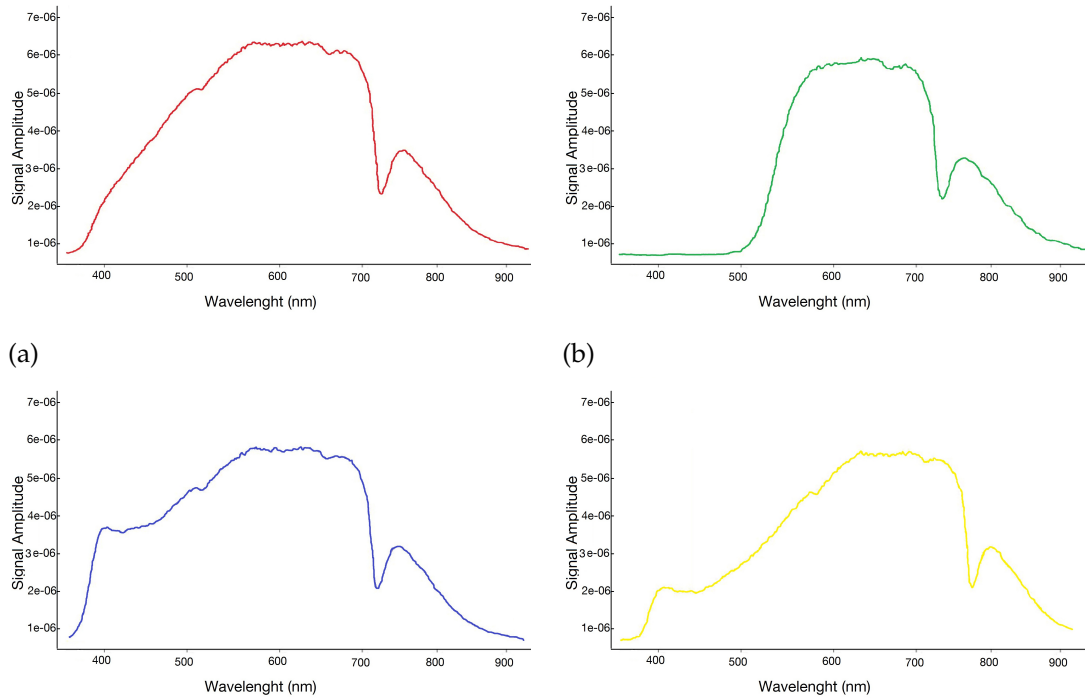


Figure 4.1.: Optical spectra from chalk with color inks in wavelength and amplitude. (a) Red color, (b) Green color, (c) Blue color, (d) Yellow color

makes the differences between four colors (red, green, blue and yellow). The experiment is perform on the phantom (as illustrated in Figure 3.7). Chalk colored with four different inks are used in the experiment. Figure 4.1 clearly demonstrate that variation in color produces difference in optical spectroscopy.

We design an experiment, to observe the difference between objects based on internal structure. In this experiment we gather six different fruits (Apple, Blueberry, Kiwi, Strawberry, Plum and Orange). Data was collected from each fruit after 12 hours for 7 days consecutively. Due to the change in the internal structure of the fruits the cure was changed, but the main shape of the cure was always constant. Stander deviation of each fruit in wavelength and amplitude is shown in Figure 4.2.

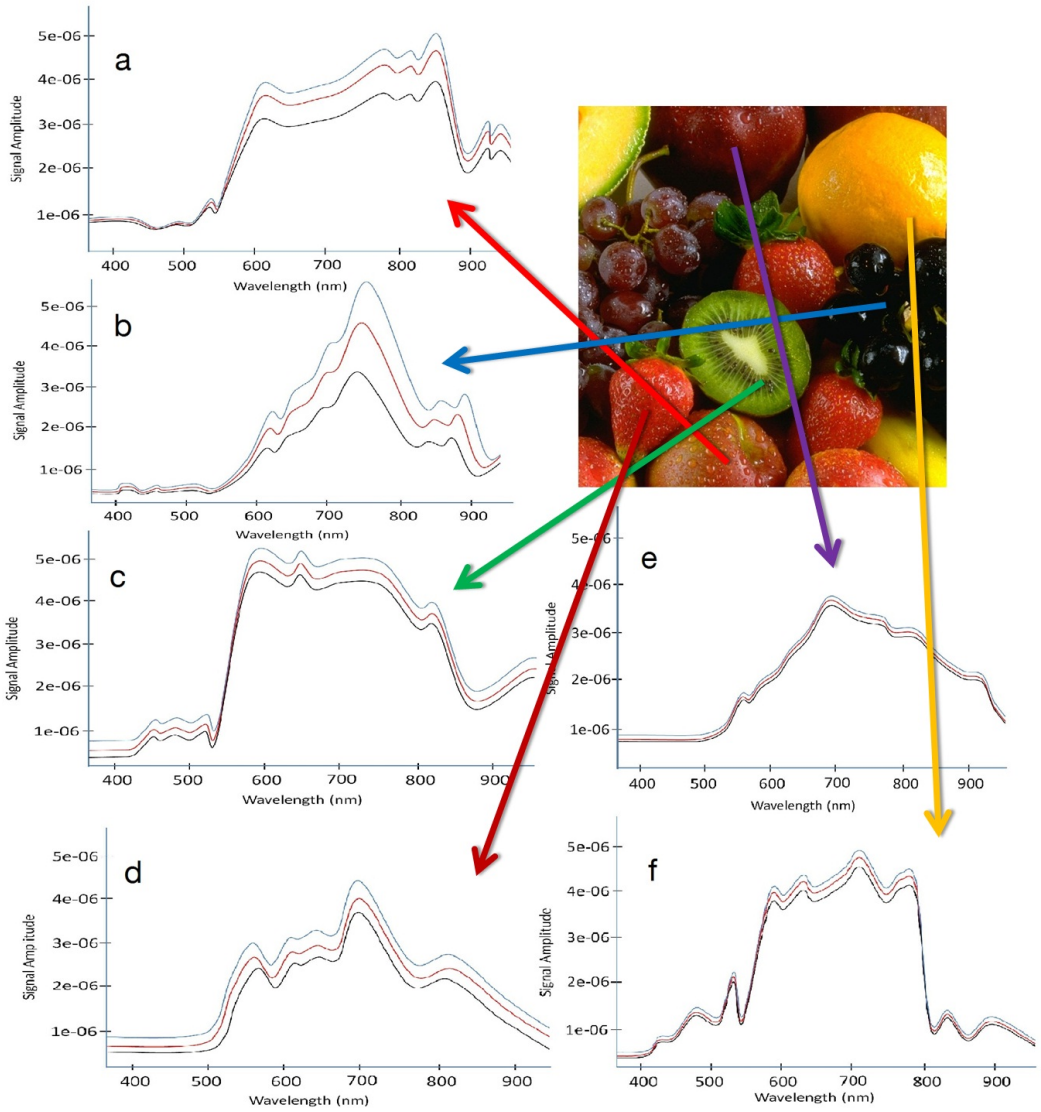


Figure 4.2.: Spectral stander deviation of each fruit in wavelength and amplitude. Lower cure (Black colored) is the minimum, the upper cure (Blue colored) represents the maximum and the middle cure (Red colored) represents the mean. (a) Apple, (b) Blueberry,(c) Kiwi, (d) Strawberry, (e) Plum, (f) Orange

4.2. State of the Art

Skin color measurement through reflectance spectroscopy has received significant attention in the literature [49][87][3][55]. It has been used to provide a numerical index for color, which in turn allows for the study of constriction of a blood vessel and abnormal redness of the skin due to local congestion, such as in inflammation [44]. Dawson et al. [40] worked on the reflectance spectroscopy for the measurement of skin tissue to exemplify the spectral properties. Farrell et al. [53] and Kienle et al. [84] addressed the problem of reflectance measurements to determining in vivo tissue optical properties. Another approach for measuring the optical reflectance over a broad range of wavelengths spectroscopy has been utilized for assessing the skin type and gestation age of newborn infants by Lynn et al. [97].

The first work to evaluate the possibilities of using reflectance spectrophotometry for discriminating between benign and malignant skin lesions was done by Marchesini et al. [104]. The authors experiments show that the wavelengths between 400 and 800 nm; were highly significant to show the differences between the reflectance spectra of benign and malignant melanomas. Consequently, the authors report a discrimination between 31 primary melanoma and 31 benign lesions with a sensitivity of 90.3% and a specificity of 77.4%, a stepwise discriminate analysis of reflectance spectral features [105].

Moreover the concluding remarks of Bono et al. [15] are that color is the most important parameter in discriminating melanomas from benign in spectrophotometric imaging of skin lesions using 420-1020 nm. Recently with Raman spectroscopy the molecular structure of skin lesions are explored [137], but due to the side effect of the laser beam on the sensitive skin surface, this technique is not preferred in the dermatology practice.

4.3. Data acquisition protocol

In our protocol, the the mole selection for the data acquisition is purely based on the doctor's (or physician's) choice based on a visual examination. The labeling of mole is performed using two classes: suspicious skin lesion (possibility of malignant melanoma) and normal skin moles based on physician's diagnosis.

The data is stored as a plot of wavelength and amplitude (as shown in Figure 4.1) by spectrometer without taking into account the mole structure. The time of data acquisition and the number of measurements depend on the number of

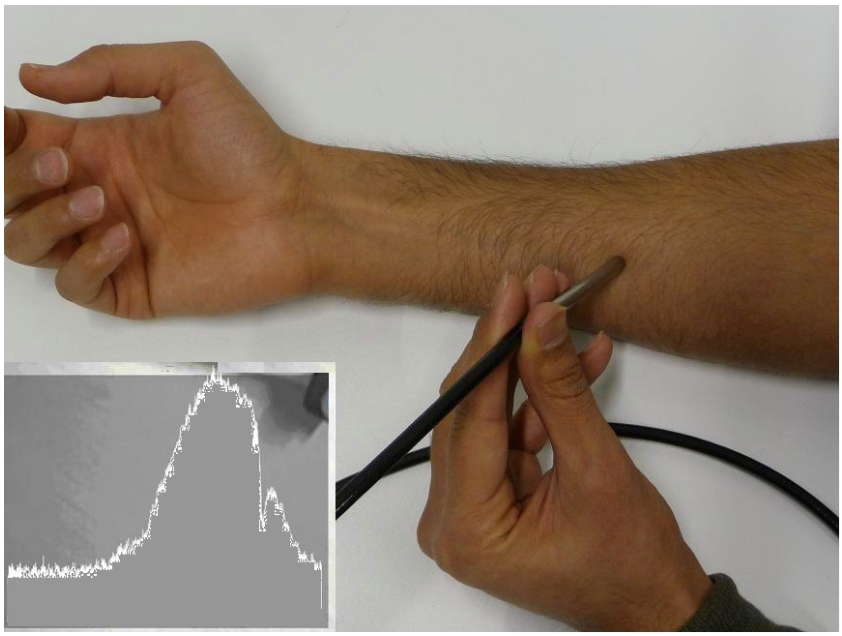


Figure 4.3.: Covering all the surface of prob tip by contacting skin surface.

moles defined on patients, where the time for whole body skin checkup was approximately 20 minutes.

The spatial resolution of sampling region is 1 mm diameter which permits the study of smaller lesions and sampling of several regions within bigger lesions. For mole size bigger than 3 mm and smaller then 6 mm we take 5 measurements (4 from the edges 1 from the center). If the mole sizes exceeds 6 mm then we take 7 measurements (6 from the edges 1 from the center). To make sure that the database is the consistent and not biased, we only use the measurements which were taken once per each mole.

The data acquisition time for one mole is 100 ms. It is important to contact the surface of the mole by the probe tip and keep the probe in a way that no light goes in from outside to ensure that the spectra are only obtained from the lesion itself, as shown in Figure 4.3). Hair, nails and tattoos are avoided during data acquisition.

4.4. Data acquisition

The data collection for this study was performed in collaboration with the dermatology department at Klinikum Rechts der Isar München; Germany. All lesions in this analysis were selected by dermatology experts. In total, 3072 spectroscopic data vectors were collected from 148 patients, where 2926 measurements were of normal skin moles and 146 measurements from malignant skin lesions. The schematic of data acquisition system is shown in Figure 4.4. Out of 146 malignant skin lesions 9 cases were histological proven melanoma. The remaining 137 are kept under observation. The details of the 9 cases of melanoma were: average Breslow thickness was 1.1 mm, the minimum being 0.1 mm and the maximum 2.8 mm, the average diameter of the lesions was 3 mm, the minimum being 2 mm and the maximum 5 mm. The average age of patients was 40, where the youngest and oldest patients were 2 and 82 years old, respectively. 70% of the examined patients were female. The collected data consists of the following clinical cases:

- Normal skin: spectra were obtained from the inside of the upper arm, groin and inside thigh, a region defined as skin that is not normally exposed to sunlight (i.e. not tanned).
- Normal skin moles: in average 19 spectra per patient were obtained from benign skin moles. Normal skin moles can be visually very similar to malignant moles, as illustrated in Figure 4.5.
- Malignant skin mole: one spectra were obtained from middle positions on the lesion. Multiple spectra were taken depending on size of the mole as discussed in data acquisition protocol section 4.3.

Immediately prior to each patient data collection session the spectrophotometer probe end was placed in the disinfectant substance to prevent migration of any diseases.

To make sure of reproducibility and accuracy of data acquisition one concern was that the pressure of the probe on the skin might cause blanching by forcing blood out of local vessels. To test a novel approach to reducing this effect and to assess the magnitude of this problem a study was performed by Osawa et al. [125]. In their study the probe was held in contact with a flat area of skin and the pressure slowly increased beyond that which would be applied normally for taking skin reflectance measurements. Increasing the pressure caused a decrease

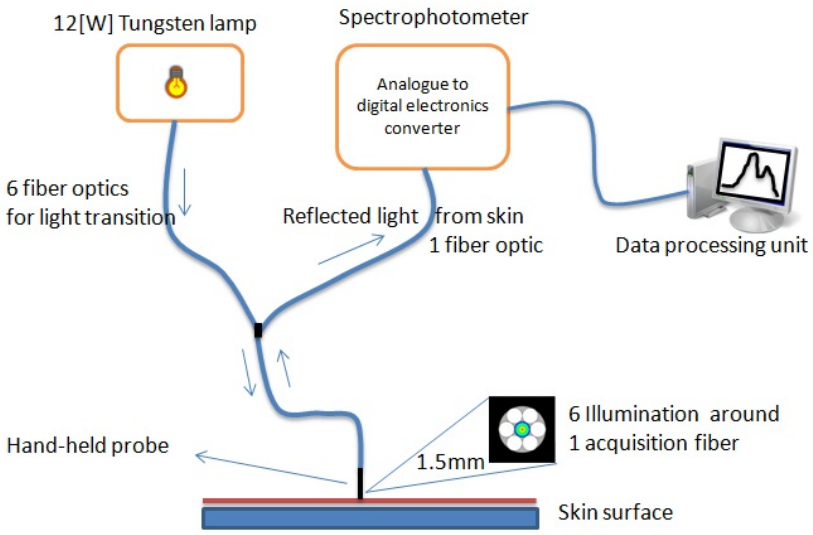


Figure 4.4.: Schematic of data acquisition system.

in overall reflectance. Osawa et al. suggested three methods for eliminating the effect: **(a)** a sensor to determine the pressure being applied, **(b)** an adhesive pad to just hold the probe against the skin, and **(c)** an electrical contact sensor to feed back information on when the probe makes contact with the skin. In our study the pressure on the skin was reduced by increasing the surface area of contact with a probe holder that was designed to slide in the probe which was also used to keep the tracking points (see in figure 3.4).

4.5. Data Processing

The spectral data is acquired as a $2048D$ vector of the floating points values $x_i \in R^{2028}, i = 1, ..., n$ where n denotes the number of measurements. Each x_i represents the discretized reflective spectrum from $178[nm]$ to $1132[nm]$ (due to limitation of hardware) of the ith measurement and is stored normalized as

$$\widehat{x}_i = \frac{x_i}{\|x_i\|_2} where i = 1, ..., n \tag{4.1}$$

To reduce the dimensions of the input data, principal components analysis (PCA) is applied. The resulting spectrum of eigenvalues $\left(e_j^i\right)_{j=1,...,2048}$ is sorted descending by magnitude. Since the highest eigenvalues represent the most relevant com-

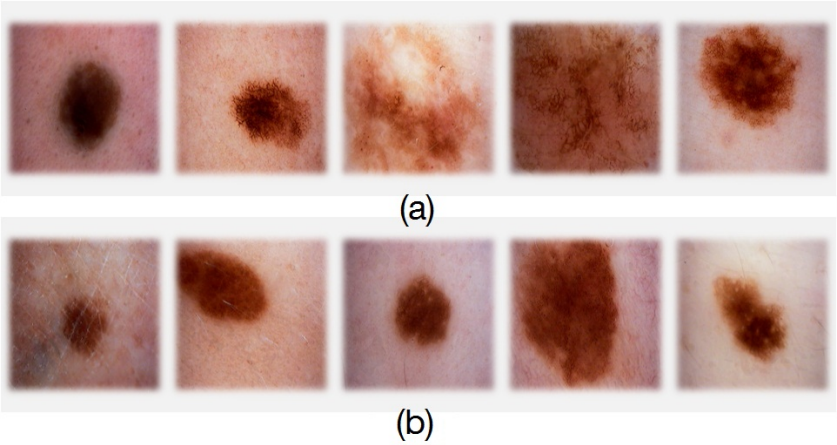


Figure 4.5.: Skin lesions: (a) Malignant skin lesions, (b) Normal skin lesions.

ponents, a cut-off value C_{PCA} is chosen, such that the final input data y_i for the classification algorithm from measurement $x_i (i = 1, ..., n)$ is

$$y_i = \left(e_j^i\right)_{j=1,...,C_{PCA}} \tag{4.2}$$

The cut-off value C_{PCA} is chosen empirically from the data. Figure 4.6 is showing a representative example of $\left(e_j^i\right)_{j=1,...,2048}$ from which C_{PCA} was selected as one of $\{2, 3, 4, 5\}$.

4.6. Classification

Classification is performed by a support vector machine (SVM) [37]. SVM was selected as the method of choice as it allows to linearly classify data in a high-dimensional feature space that is non-linearly related to the input space via the use of specific kernel functions, such as polynomial functions or radial basis functions (RBF). This way we can build complex enough models for skin lesion classification while still being able to compute directly in the input space.

The SVM classifier needs to be trained first before using it, thus we partition our already reduced input data $(y_i), i = 1, ..., n$ into two partitions, $T \subset \{1, ..., n\}$ the training set and $V \subset \{1, ..., n\}$ the testing (or validation) set with $T \cup V = \{1, ..., n\}$ and $T \cap V = \{\}$. The training data set T is labeled manually into two classes with the ground truth, $l(y_i) = \pm 1$. Once the classifier is trained, a simple evaluation of the decision function $d(y_i) = \pm 1$ will yield the classification of any data y_i.

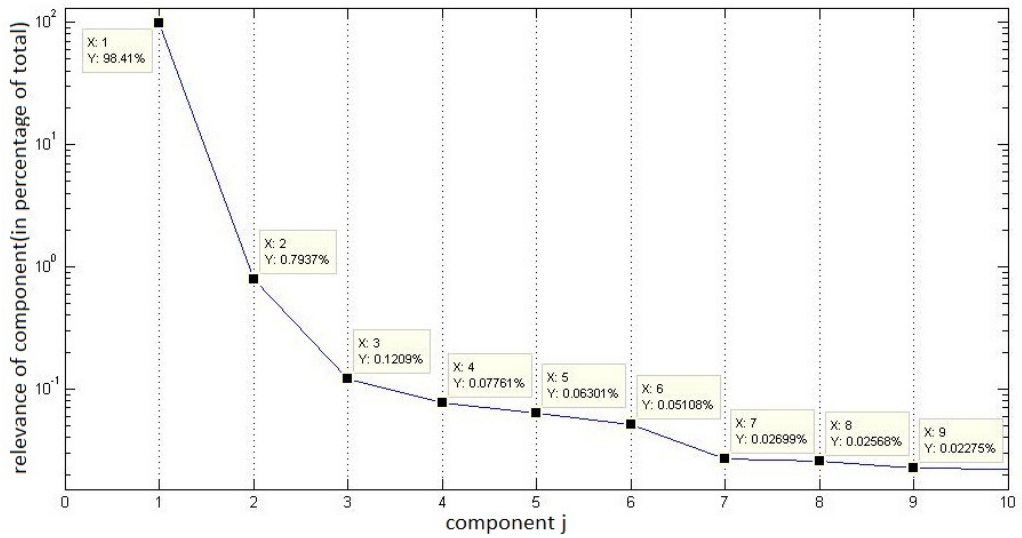


Figure 4.6.: Representative example of the first part of the sorted PCA eigenvalue spectrum $\left(e_j^i\right)$, the $y - axis$ shows the values of the component as a percentage of the total in log scale.

In detail, SVM is trying to separate the data $\phi(y_i)$ mapped by the selected kernel function ϕ by a hyperplane $w^T\phi(y_i) + b = 0$ with w the normal vector and b the translation. The decision function then is $d(y_i) = sgm\left(w^T\phi(y_i) + b\right)$. Maximizing the margin and introducing slack variables $\xi = (\xi_i)$ for non-separable data, we receive the primal optimization problem:

$$\min_{w,b,\xi} = \frac{1}{2}w^T w + C\sum_{i\in T}\xi_i \tag{4.3}$$

with constraints $l(y_i)(w^t\phi(y_i) + b) \geq 1 - \xi_i,\ \xi \geq 0$ for $i \in T$. C is a user–determined penalty parameter. Switching to the dual optimization problem allows for easier computation,

$$\min_{\alpha} = \frac{1}{2}\alpha^T Q\alpha - e^T\alpha \tag{4.4}$$

with constraints $0 \leq \alpha_i \leq C$ for $i \in T,\ \sum_{i\in T} y_i\alpha_i = 0$. The $\alpha = (\alpha_i)$ are the so–called support vectors, $e = [1, ...1]^T$ and Q is the positive semidefinite matrix formed by $Q_{jk} = l(y_j)l(y_k)K(y_j, y_k)$, and $K(y_j, y_k) = \phi(y_j)^T\,\phi(y_k)$ is the kernel function built from ϕ. Once this optimization problem is solved, we determine the hyperplane parameters w and b, w directly as $w = \sum_{i\in T}\alpha_i l(y_i)\phi(y_i)$ and b

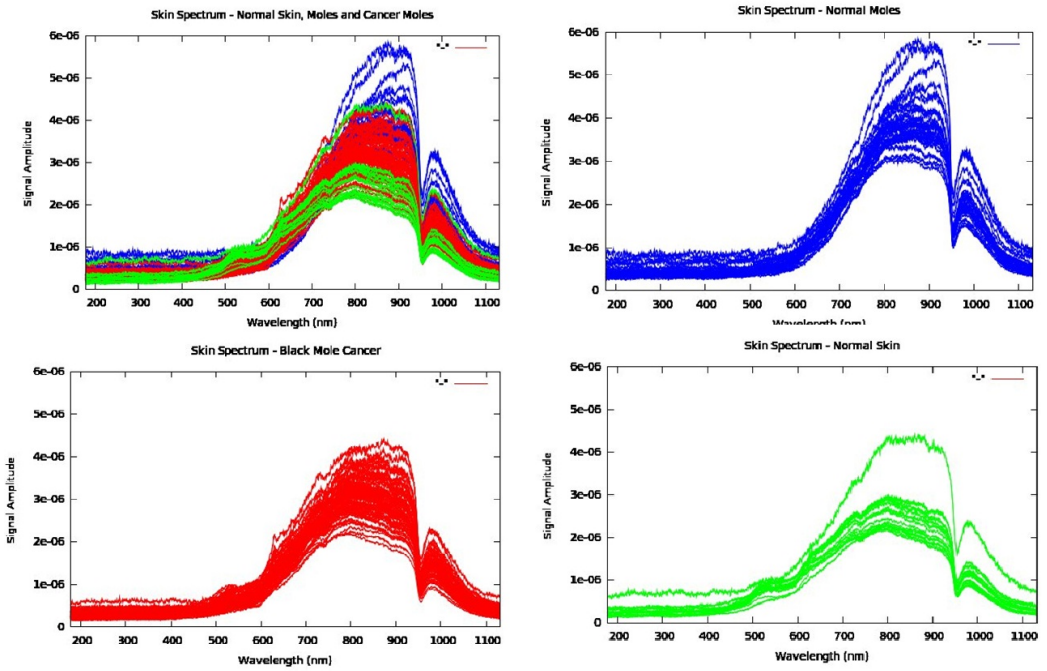


Figure 4.7.: Plot of all normalized spectra $\hat{x}_i$ from the training data set T, color-coded as blue for normal skin moles, red malignant mole and green for normal skin. One cure represent one skin lesion data.

via one of the Karush-Kuhn-Tucker conditions as $b = -l(y_i)y_i^T w$, for those i with $0 < \alpha_i < C$. Thus the decision function of the trained SVM classifier ends up as

$$d(y_i) = \text{sgn}\big(w^T \phi(y_i) + b\big) = \text{sgn}\left(\sum_{j \in T} \alpha_i l(y_i) K(y_j, y_i) + b\right). \tag{4.5}$$

4.7. Experiments

Data collection of 3072 spectroscopic is define as $(x_i), i = 1, ..., 3072$ labeled into the two classes normal skin $l(x_i) = 1$ and lesion $l(x_i) = -1$. The 3072 data points were randomly separated into a training data set T and a testing (validation) data set V with $|T| = 2072$ and $|V| = 1000$, however retaining the balance of both sets containing 50% each of the two classes. A color-coded representation of the normalized skin spectra $\hat{x}_i, i \in T$ of the training data set T is shown in Figure 4.7.

Before classification, PCA was applied to the $\widehat{x}_i$ for dimension reduction to yield

Table 4.1.: Results of the cross-validation using the training dataset T.

SVM Traning				
Parameters	Linear Kernel	Poly Kernel	RBF Kernel	Sigmoid Kernel
$C_{PCA} = 2$	95 ± 9.2	96 ± 8.3	95 ± 7.5	95 ± 10.1
$C_{PCA} = 3$	95 ± 8.3	96 ± 6.7	97 ± 9.5	96 ± 10.5
$C_{PCA} = 4$	95 ± 11.5	97 ± 7.2	97 ± 8.7	96 ± 8.6
$C_{PCA} = 5$	96 ± 9.2	97 ± 10.5	97 ± 8.3	97 ± 7.7

our classification inputy_i. The eigenvalue cut-off C_{PCA} was empirically chosen as one of $C_{PCA} \in 2, 3, 4, 5$.

The SVM classifier (we used LibSVM, [26]) was then trained using the training data set T. As there are multiple parameters to be selected, like for example the penalty parameter C, we performed a cross-validation of 10 folds via parallel grid search. The average accuracy on the prediction of the validation fold is the cross validation accuracy.

4.8. Results

The cross-validation of the training data set T determined, among others, the parameters $C = -5$ and $\gamma = -7$. For the further parameters C_{PCA} and the choice of the kernel (linear, polynomial, radial basis function (RBF) or sigmoid) we performed cross validation of the training data set T, the results are shown in Table 4.1. The best results were received consistently by using the RBF kernel, while for C_{PCA} the value of 5 turned out to be the best choice with an accuracy of 97 ± 8.3.

With the training of the classifier completed, we studied the accuracy of the testing (validation) data set V. We compared the manual ground truth labeling $l(y_i)$ for data point y_i with the computed decision function $d(y_i)$ to compute the accuracy as follows

$$Accuracy = \frac{\text{\# of correctly predicted data}}{\text{\# of total data}} \times 100\% = \frac{i\,|l(y_i)d(y_i)| > 0}{|V|} \times 100\% \quad (4.6)$$

The results are shown in Table 4.2. We achieve the same accuracy of 94.9% for the kernels RBF the C_{PCA} values of 4 and 5. This corresponds to Figure 4.6, where it is clear that between C_{PCA} 4 and 5 there is only very little difference. In total we received the best results using the RBF kernel and $C_{PCA} = 5$.

Table 4.2.: Classification accuracy results using the testing dataset V.

	Testing			
Parameters	Linear Kernel	Poly Kernel	RBF Kernel	Sigmoid Kernel
$C_{PCA} = 2$	86.8%	90.3%	89.9%	88.8%
$C_{PCA} = 3$	89.3%	92.5%	91.8%	90.3%
$C_{PCA} = 4$	91.9%	92.9%	94.9%	94.1%
$C_{PCA} = 5$	92.1%	93.6%	94.9%	94.6%

4.9. Discussion

In this chapter, we have presented a simple, portable and affordable setup for reflectance spectroscopy and SVM-based classification of skin lesions. Our study presents an enhancement in system hardware and software design, techniques for data processing and measured performance in comparison to previously reported studies. Our experiments on patient dataset served as a base to choose and tune various parameters for classification. The results of 94.9% accuracy in distinguishing normal skin mole from malignant skin lesion are comparable to those of a dermatologist using visual inspection [113]. We use spectroscopic data collected form normal skin mole as well as malignant skin mole. IThe ground truth for this study was created by the visual assessment of a dermatology expert without taking the pathological information into account. The experiment is performed with participation of 4 dermatologists with different levels of expertise. We observed that our algorithm performed comparable to an experienced dermatologist and led to higher classification accuracy compared to less experienced physicians. This results also suggests that our algorithm can be utilized for training purposes.

Marchesini et al. [104] suggest a normalization of the malignant skin lesions with respect to the normal skin individually for each patient. However, our discussion with dermatologist experts revealed that this is in fact contradictory with their clinical experience, where no relation has been observed between the color of the skin and of the lesion. Furthermore, in our detailed literature study we did not find any supporting evidence suggesting the necessity of this normalization. Our own studies showed that the most variability in the spectrum is present between the normal and malignant skin lesions this normalization might reduce the ability of our method to distinguish between lesion types. A study designed for a comprehensive analysis of the spectral variance would be required to establish a

final conclusion on this issue.

There were no reflectance spectra features that distinguish between malignant melanoma Breslow's depth, which is difficult to differentiate visually. Also no features discriminated between malignant melanoma and seborrhoeic keratosis, whereas the visual discrepancies between them are normally very apparent to the dermatology expert. The fact that the reflectance spectra are very similar remains at present inexplicable. However, this may provide an explanation for the frequent confusion of this lesion with melanoma by non-experts such as general practitioners.

Our study of skin lesion reflectance spectral classification with no additional information about the lesion creates a basis for the upcoming research in the field spectroscopy. To some extent, this computer aided system provides a second opinion for the dermatologist. High classification accuracy achieved by the use of our method suggest its further clinical evaluation on larger number of case studies. More work is required to fully understand the mechanism behind the interaction between light and skin that results in the observed reflectance spectra.

CHAPTER 5

Manifold Learning for Dimensionality Reduction of Skin Lesions Using Optical Spectroscopy Data

SPECTROSCOPY data is typically very high dimensional (in the order of thousands), which causes difficulties in interpretation and classification [19]. In this chapter, we present different manifold learning techniques to reduce the dimensionality of the input data and get clustering results. Spectroscopic data of 48 patients with suspicious and malignant melanoma lesions is analyzed using ISOMAP, Laplacian Eigenmaps and Diffusion Maps with varying parameters. The results are compared to PCA.

5.1. Introduction

Most recent applications of machine learning in data mining, computer vision, and in other fields require deriving a classifier or function estimate from an large data set. Modern data sets often consist of a large number of examples, each of which is made up of many features. Though access to an abundance of examples is purely beneficial to an algorithm attempting to generalize from the data, managing a large number of features (some of which may be irrelevant or even misleading) is typically a burden to the algorithm. Overwhelmingly complex feature sets will slow the algorithm down and make finding global optima difficult. To lessen this burden on standard machine learning algorithms (e.g. classifiers,

function estimators), a number of techniques have been developed to vastly reduce the quantity of features in a dataset, i.e. to reduce the dimensionality of the data.

Dimensionality reduction has other, related uses in addition to simplifying data so that it can be efficiently processed. The most obvious is visualization; if data lies, for instance, in a 100-dimensional space, one cannot get an intuitive feel for what the data looks like. However, if a meaningful two or three dimensional representations of the data can be found, then it is possible to analyze it more easily. Though this may seem like a trivial point, many statistical and machine learning algorithms have very poor optimality guarantees, so the ability to actually see the data and the output of an algorithm is of great practical interest. In our case, spectroscopic data is typically acquired as a high dimensional vector (in our case a 2048 element vector); this high-dimensionality, however, creates difficulties for visualization and classification of the data. Manifold learning has a significant role in dimensionality reduction and clustering due to its nature of unsupervised learning [24].

There are many approaches to dimensionality reduction based on a variety of assumptions and used in a variety of contexts. We will focus on an approach initiated recently based on the observation that high-dimensional data is often much simpler than the dimensionality would indicate. In this work, we present results of applying different manifold learning techniques such as Isomap [148], Laplacian Eigenmaps [6] and Diffusion Map [33] to spectroscopy data from 48 patients with normal and actually malignant lesions to reduce the dimensionality, and compare them to traditional linear techniques Principal Component Analysis. Clustering results after dimensionality reduction are shown for each technique, where some of the method/parameter combinations yield excellent results on the patient data compared to the diagnosis of the treating physicians.

5.2. Notation

Throughout the chapter, we will be solving the problem of dimensionality reduction of the high-dimensional points to a low-dimensions in Euclidean space.

- Ψ is the mapping function from high-dimension to low-dimension.
- The high-dimensional input points is referred to as $x_1, x_2, ..., x_n$ and its matrix representation is $X = [x_1, \cdots, x_N]$.

- The low-dimensional representations that the dimensionality reduction algorithms find is referred to as $y_1, y_2, ..., y_n$. Y is the matrix representation of these points.
- N is the number of input points (high-dimensional space).
- D is the dimensionality of the input (i.e. $x_i \in \mathbb{R}^D$).
- d is the dimensionality of the manifold (low-dimensional space) and it corresponds to the dimensionality of the output (i.e. $y_i \in \mathbb{R}^d$).
- k is the number of nearest neighbors used by the manifold algorithm.
- $K(i)$ denotes the set of the k-nearest neighbors of x_i.
- E are the edges used by non-linear manifold learning algorithms are introduced whenever some criteria are met between the points.
- Λ is the eigenvalue matrix, which is a diagonal matrix with the ordered eigenvalues. And V is the eigenvector matrix, which has the eigenvectors with the same order as eigenvalues. The ith eigenvalue corresponds to the ith eigenvector.

5.3. Manifold Learning

In the field of machine learning, a very popular research area is manifold learning, which is related to the algorithmic techniques of dimensionality reduction. Manifold learning can be divided into linear and nonlinear methods. Linear methods, which have long been part of the statistician's toolbox for analyzing multivariate data, include Principal Component Analysis (PCA) and multidimensional scaling (MDS). Recently, researchers focus on techniques for nonlinear manifold learning, which includes Isomap, Locally Linear Embedding, Laplacian Eigenmaps, Hessian Eigenmaps, and Diffusion Maps [140]. The algorithmic process of most of these techniques consists of three steps, a nearest-neighbor search, a computation of distances between points, and an eigen-problem for embedding the D-dimensional points in a lower-dimensional space. In this section, we provide basic details of manifold learning: Isomap, Laplacian Eigenmaps and Diffusion Maps. These algorithms will be compared and contrasted with the linear method PCA

for a spectroscopic dataset. The goal is to find a mapping function Ψ from the original D-dimensional data set X to a d-dimensional dataset Y in which distances and information are preserved as much as possible and $d < D$:

$$\Psi : \mathbb{R}^D \rightarrow \mathbb{R}^d \tag{5.1}$$

In our case, we have D = 2048 and thus

$$\Psi : \hat{x}_i \in \mathbb{R}^{2048} \rightarrow y_i \in \mathbb{R}^d \tag{5.2}$$

where X is a matrix, x_i is vector and $\mathbb{R}^d$ is a space. The following section covers some of the basic definitions of linear method (PCA) and non-linear manifold learning methods.

5.3.1. Principal Component Analysis

A linear method such as PCA ignores protrusion or concavity of the data [74]. In order to demonstrate the shortcomings of purely linear methods, we will show results using PCA and compare with nonlinear manifold learning. PCA finds a subspace i.e. which finds an optimal subspace that best preserves the variance of the data [140].

The goal of PCA is to find an optimal subspace i.e. the variance of the data is maximized. In general, manifold learning methods do not care about the variance of the data. Non-linear methods in particular, typically famous on preserving neighborhood properties within the data [140]. The input and output of PCA are defined as in equation 5.1 , given N input points. The algorithm performs the following steps:

1. Calculate the empirical mean vector for each dimension $j \in 1 \cdots D$
   $\mu[j] = \frac{1}{N} \sum_{i=1}^{N} X[i, j]$
2. Subtract μ ($D \times 1$) from each column of the $D \times N$ input matrix X. The subtracted matrix $B = X - \mu h$, where h is a $1 \times N$ vector of 1's.
3. Compute the $D \times D$ covariance matrix $C = \frac{1}{N-1} B \cdot B^\top$
4. Solve the eigenvector problem to find the matrix V of eigenvectors, so that $V^{-1} \cdot C \cdot V = P$ with P being the matrix in which the decreasing eigenvalues (corresponding to their eigenvectors) are on the diagonal and $V^\top = V^{-1}$. All eigenvectors are orthogonal and they form an orthonormal basis.

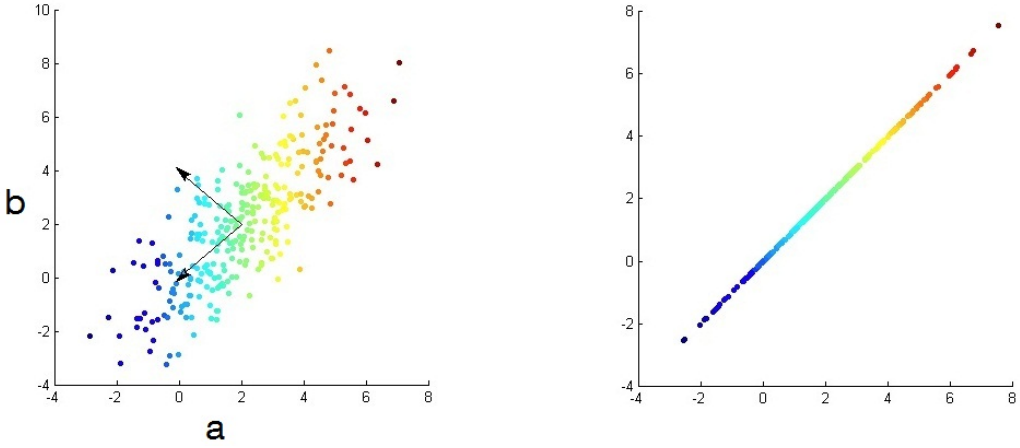


Figure 5.1.: Working example of PCA. The left image shows a Gaussian distribution together with the two principal components. The coloring is dependent on values of a and b. The right side shows the projection on the eigenvector corresponding to the largest eigenvalue [140].

5. Project the data onto the new d-dimensional subspace, using the first d columns of V, where d is chosen according to some measure (data energy or highest variance): $Y = [v_1, \ldots, v_d]^\top \cdot X$

Figure 5.1 shows a Gaussian distribution together with the first (and only) two principal components, calculated by the method described above. The vectors are therefore the eigenvectors of the matrix C.

The coloring is linearly dependent on the values of a and b. The right side shows the projection on the eigenvector corresponding to the largest eigenvalue. As one can see, the variance of the data is preserved.

Figure 5.2 shows that PCA cannot handle non-linear datasets.The left image shows a spiral distribution (2-d Swiss roll) together with the two principal components. The coloring is dependent on the values of t, where the function is given as $f(t) = (tcos(t), tsin(t))$. The right side of Figure 5.2 shows the overlapping projection on the eigenvector corresponding to the largest eigenvalue. One can observe that blue, red and yellow points are all overlapping in the center of the projected line [140].

This means that most geometric information of the data is lost through this projection. In most cases distances are only meaningful in local neighborhoods, fol-

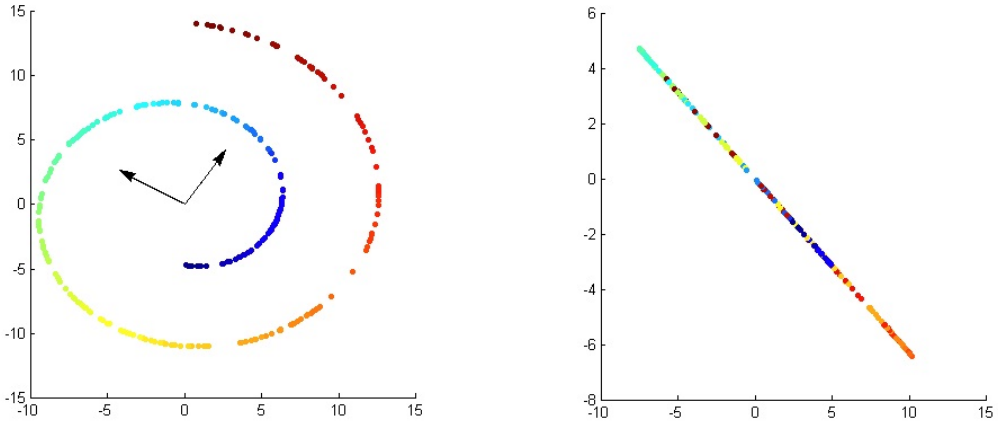


Figure 5.2.: PCA cannot handle non-linear datasets. The left image shows a spiral distribution (2-d Swiss roll) together with the two principal components. The coloring is dependent on the values of t, where the function is given as $f(t) = (tcos(t), tsin(t))$. The right side shows the overlapping projection on the eigenvector corresponding to the largest eigenvalue [140].

lowing Non-linear manifold learning methods address this problem.

5.3.2. Non-linear Manifold Learning Methods

Typical non-linear manifold learning methods are graph-based and perform the following three basic steps.

1. Build undirected similarity graph $G = (V, E)$. where the vertices V are give by the data points x_i

2. Estimate local properties, i.e. the weight matrix W to define the weighted similarity graph $G = (V, E, W)$, where $w_{ij} \geq 0$ represents the weight for the edge between vertex i and j. Weights are obtained by means of a kernel. A weight of 0 means that the vertices are not connected.

3. Derive an optimal global embedding Ψ which preserves these local properties.

There are three often used techniques for building the similarity graph G. First, there is the ϵ-neighborhood graph which connects all vertices with distance $\|x_i - x_j\|^2$ smaller than ϵ. The ϵ graph is naturally symmetric [155] [140].

Contrary to this local connection is the fully connected graph which uses a similarity function that incorporates local neighborhood relations such as the Gaussian function: $w_{ij} = exp(-\|\|x_i - x_j\|^2/(2\sigma^2))$. This leads directly to the third step, since it implicitly defines the weights [140].

k-nearest neighbor (kNN) graphs combine both worlds by connecting each vertex only to its *k*-nearest neighbors.

5.3.2.1. Isomap

Isometric feature mapping was one of the first algorithms introduced for manifold learning [148]. Isomap is a non-linear generalization of multidimensional scaling (MDS) where similarities are defined through geodesic distances, i.e. the path along the manifold. MDS tries to find a low-dimensional projection that preserves pairwise distances by finding the eigenvectors of the distance matrix [36] [140].The Isomap algorithm consists of two main steps:

1. Estimate the geodesic distances (distances along a manifold) between points in the input using shortest-path distances on the data sets k-nearest neighbor graph.

2. Use MDS to find points in low-dimensional Euclidean space whose inter-point distances match the distances found in step 1.

As shown in Figure 5.3 the swiss roll is unfold nicely by keeping the geodesic distance. One general disadvantage of Isomap is that it is governed by the geodesic distances between distant points. In other words, the embedding Ψ preserves the distances of even faraway points. This often leads to distortions in local neighborhoods. Other disadvantage of Isomap is its speed which is quite low due to the complexity of MD, in particular the shortest-path computation.

5.3.2.2. Laplacian Eigenmaps

Laplacian Eigenmaps [7] try to preserve distance relations and that they can be solved by one sparse eigenvalue problem [6] [140].

We compute an embedding Ψ in three steps:

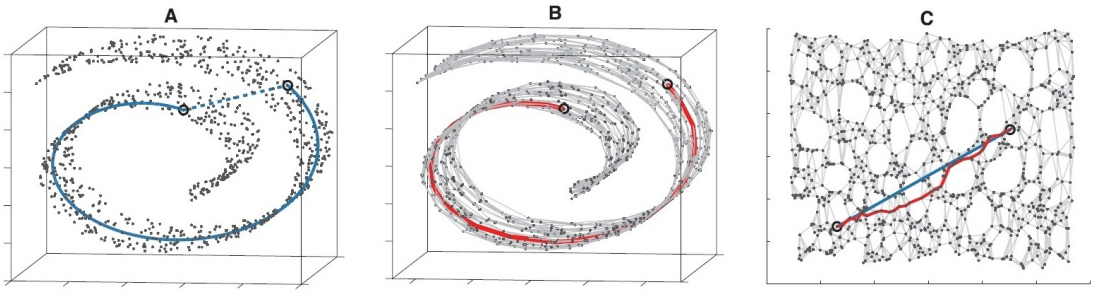


Figure 5.3.: The Swiss roll data set. (A) shows that the Euclidean distance between two points do not reflect their similarity along the manifold. (B) shows the geodesic path calculated in step 1. of the Isomap algorithm (C) displays the 2-dimensional embedding defined by Isomap [140].

1. Build undirected similarity graph $G = (V, E)$.

2. Choose a weight matrix W either by simply setting $W_{ij} = 1$ for all connected vertices or using a heat kernel with parameter $t : w_{ij} = exp(-\|x_i - x_j\|^2 / t)$ If the graph is not fully connected, proceed with step 3 for each connected component.

3. Find the eigenvalues $0 = \lambda_1 \leq ... \leq \lambda_n$ and eigenvectors $v_1, ..., v_n$ of the generalized eigenvalue problem: $L_v = \lambda D v$ Where L is laplacian matrix and D is degree matrix (for every entry ij the number of edges connecting to that node). Define the embedding: $\Psi : x_i \rightarrow (v_2(i), ..., v_d(i))$

Laplacian Eigenmaps are a special case of diffusion maps. This special case handles only manifolds from which the data is sampled uniformly, something that rarely happens in real machine learning tasks. The eigenvalues and eigenvectors of the Laplacian reveal the information about the graph such as whether it is complete or connected [140].

5.3.2.3. Diffusion Maps

Diffusion Maps is another technique for finding meaningful geometric descriptions for data sets even when the observed samples are non-uniformly distributed [33]. It is similar to Laplacian Eigenmaps, but the mapping are defined via diffusion distances. Diffusion Maps achieves dimensionality reduction by re-organizing

data according to parameters of its underlying geometry. The connectivity of the data set, measured using a local similarity measure, is used to create a time-dependent diffusion process. As the diffusion progresses, it integrates local geometry to reveal geometric structures of the data set at different scales. Defining a time-dependent diffusion metric, we can then measure the similarity between two points at a specific scale (or time), based on the revealed geometry. A diffusion map embeds data in (transforms data to) a lower-dimensional space, such that the Euclidean distance between points approximates the diffusion distance in the original feature space. The dimension of the diffusion space is determined by the geometric structure underlying the data, and the accuracy by which the diffusion distance is approximated [41] [140]. To conclude this section, a sketch of the diffusion maps algorithm is shown stepwise:

1. Define a kernel, $c(x; y)$ and create a kernel matrix, C, such that $C_{i,j} = c(X_i, X_j)$.
2. Create the diffusion matrix by normalizing the rows of the kernel matrix.
3. Calculate the eigenvectors of the diffusion matrix.
4. Map to the d-dimensional diffusion space at time t, using the d dominant eigenvectors and values

In the next section, we will apply all above described manifold learning methods for dimensionality reduction and clustering into two classes (in our experiments, malignant and nonmalignant skin lesions).

5.4. System Experiments

For the hardware setup, please refer to section 3.4. The data collection in this study was made possible by the support of the dermatology department of the hospital Klinikum Rechts der Isar München; Germany. We collected 372 spectroscopic data vectors from 48 patients, 326 measurements were of normal skin moles, 46 measurements were malignant skin lesion (as diagnosed by the treating physician). 13 cases out of 46 malignant skin lesions were pathologically verified by the laboratory. All lesions for this experiment were selected by only well-experienced physicians (not by newly joined dermatologists). This was the only additional protocol to the data acquisition protocols as discussed in section 4.3. A color-coded representation of the normalized skin spectra data set is shown in Figure 5.4 and

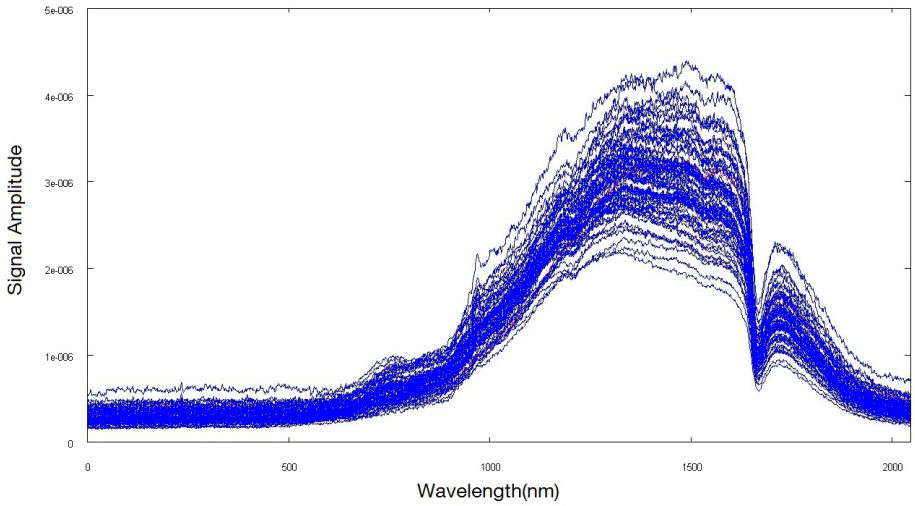


Figure 5.4.: Normalized spectral graph data sets, malignant skin lesions. Each cure is the vector, representing one skin lesion. without labeling of the data the overlaps cures are difficult to separate

Figure 5.5. Figure 5.4 shows malignant skin lesions and Figure 5.5 shows malignant skin lesions combine with normal skin mole. In Figure 5.5 one can observe the overlap between two classes of data set.

The proposed methods were implemented in Matlab 10.1 using libraries for the dimensionality reduction. Clustering was performed by selecting a separating hyperplane in the processed three-dimensional data.

Before applying manifold learning we need to elucidate some parameters that play a significant role in producing meaningful data representation. The parameters for the non-linear dimensionality reduction techniques are:

- k: The k-nearest neighbors specify the number of nearest neighbors used to build the graph for the Isomap, Laplacian eigenmaps and Diffusion maps methods. If k is chosen too large or too small, the local geometry may not be interpreted correctly. Here we used the values of k = 15, 20, 30, 35.
- Alpha: This parameter controls the normalization.
- Sigma: This specifies the width of the Gaussian kernel. The larger Sigma is, the more weight far-away points will exert on the weighted graph. We used Sigma = 20, 30.

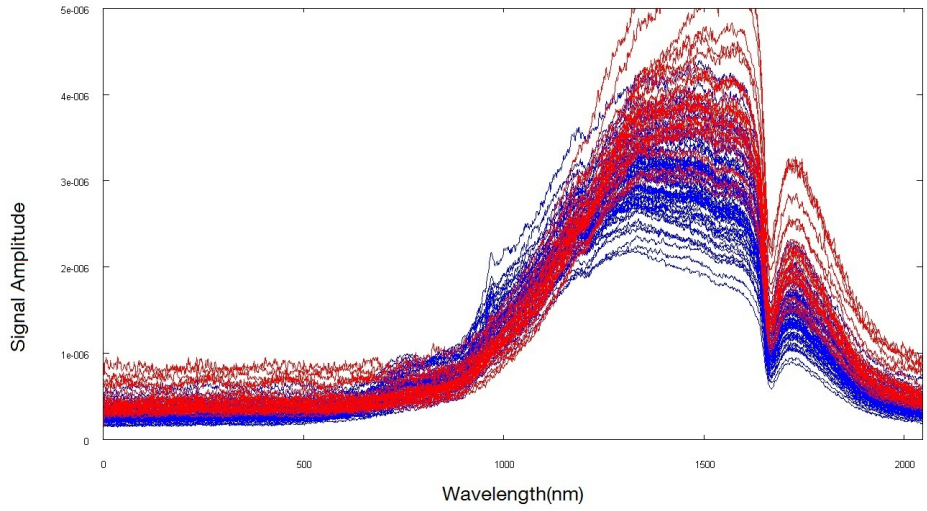


Figure 5.5.: Normalized spectral graph data sets combined form, blue for malignant skin lesions and red for normal skin mole.

5.5. Results

All four studied methods (PCA, Isomap, Laplacian Eigenmaps and Diffusion maps) were applied independently. PCA is applied on 2048 dimensional data vectors, and the first three most significant components are taken. Each point represents one skin lesion (malignant or benign). The data set is labeled which is represented by two color red and blue. Red points are malignant and blue are benign. It is clear from the 3D representation of the data shown in Figure 5.6 that the data is not clearly distinguishable into two clusters. The main reason PCA could not perform well is because PCA maximizes the variance of the data and in our case direction of the variance helps to distinguish between the two classes. The best clustering accuracy PCA achieved is 63%.

The 3D representation of the 2048D data victor after applying Isomap is shown in Figure 5.7. It is clear from the figure that some area of the data is very nicely clustered. We know as discussed in section 5.3.2.1 that Isomap is governed by the geodesic distances between distant points, which causes distortions in local neighborhoods so maybe that is one reason that the data set is not clustered perfectly. Overall Isomap produce better results than PCA.

Figure 5.8, shows that the Diffusion maps is able to preserve the order of clus-

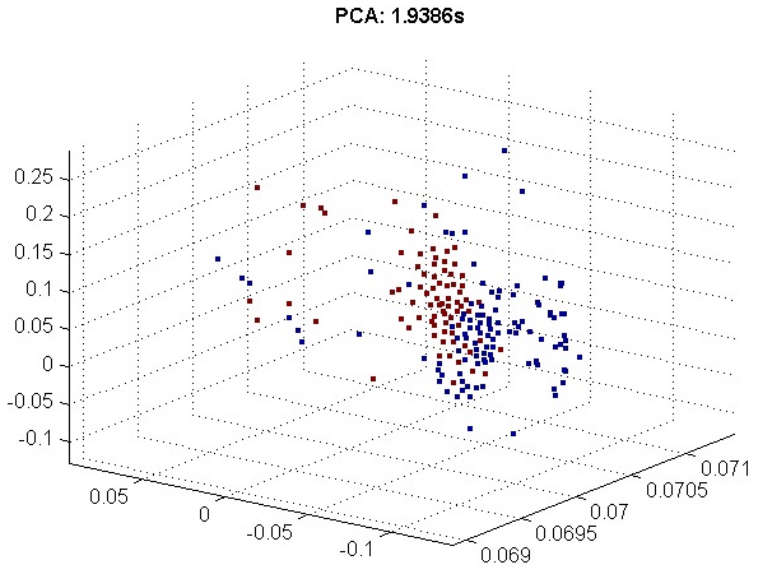


Figure 5.6.: PCA 3D representation of 2048D dataset. The best possible angle to visualize the data points. PCA:1.9386s is the runtime of method

Table 5.1.: Clustering accuracy with different methods and parameters. Where k is k-nearest neighbors , A is for Alpha and S is representing Sigma parameter

Parameters	Isomap	Laplacian eigenmaps	Diffusion maps
$k = 15, A = 2, S = 20$	88%	0%	10%
$k = 20, A = 2, S = 30$	90%	87%	81%
$k = 30, A = 1, S = 20$	86%	92%	90%
$k = 35, A = 1, S = 20$	94%	96%	92%

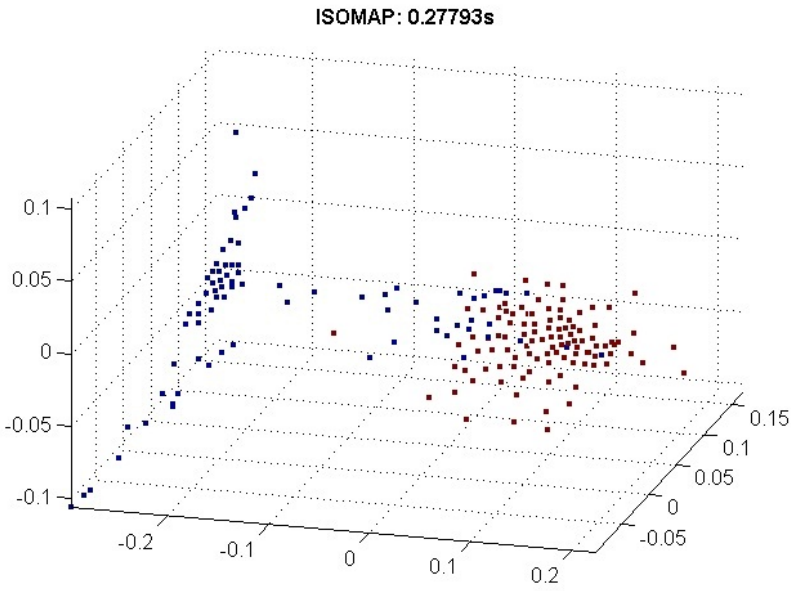


Figure 5.7.: Applying manifold learning by using Isomap and the output 3D representation as a result. The points that corresponds to malignant data example, are well separated from those points corresponds to benign.

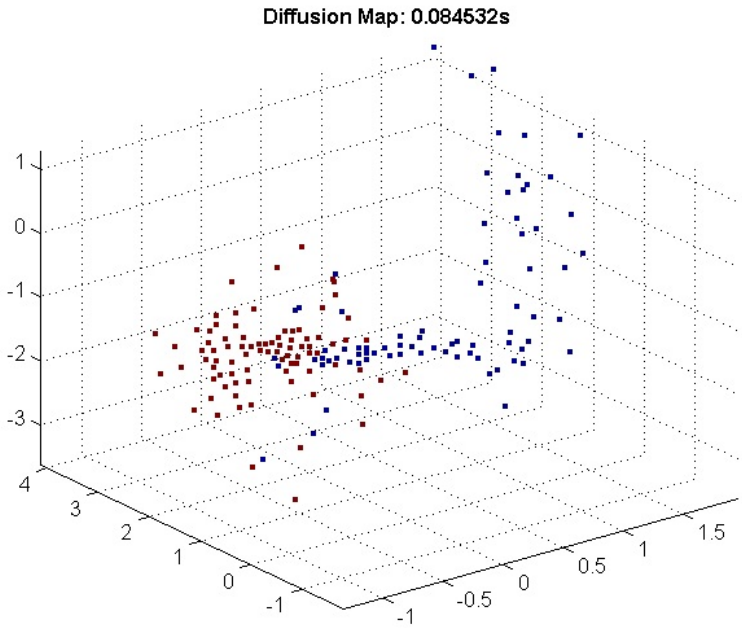


Figure 5.8.: Diffusion maps 3D data representation. The clusters are clearly visible.

ters in three dimensions similar as Isomap. Choosing the right parameter(s) is a difficult stage in manifold learning. Experiments are performed with different parameters shown in Table 5.1. The results were computed as the number of correctly classified points over the total number of points and as a ground truth we have labeling provided by dermatologist. According to the literature [41],[5] Diffusion maps perform better as compared to other manifold learning techniques but in our case Laplacian eigenmaps produces best results by choosing the right parameters shown in Figure 5.9. Laplacian eigenmaps preserves local neighborhood of the points which reflect the geometric structure of the manifold.

In Figure 5.10, all four methods are shown with worst parameters selection. The figure shows that the dataset is not easily distinguishable into two clusters. Variation in parameters for non-linear manifold learning methods are shown in Table 5.1.

Isomap capture local geometry correctly and the dataset is clustered into two parts with an accuracy of 94% out of 100% as shown in Table 5.1. By increasing the neighborhood size to 20 and Sigma to 30, Laplacian eigenmaps and Diffusion maps perform better. Adding even more neighborhood information, Laplacian

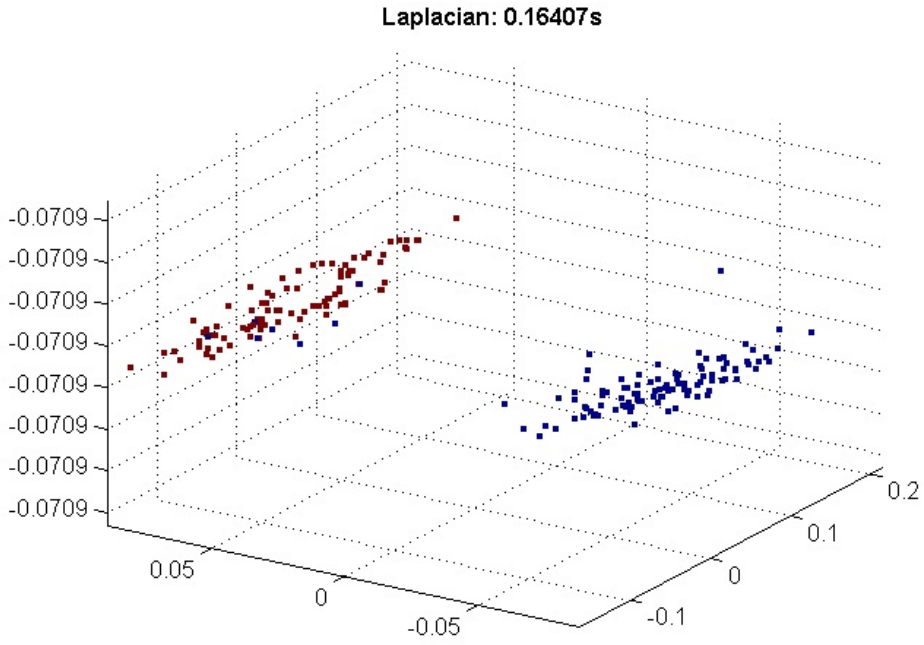


Figure 5.9.: Laplacian Eigenmaps 3D representation of 2048D dataset. Apart from few points which are in wrong cluster, the two clusters are well separated.

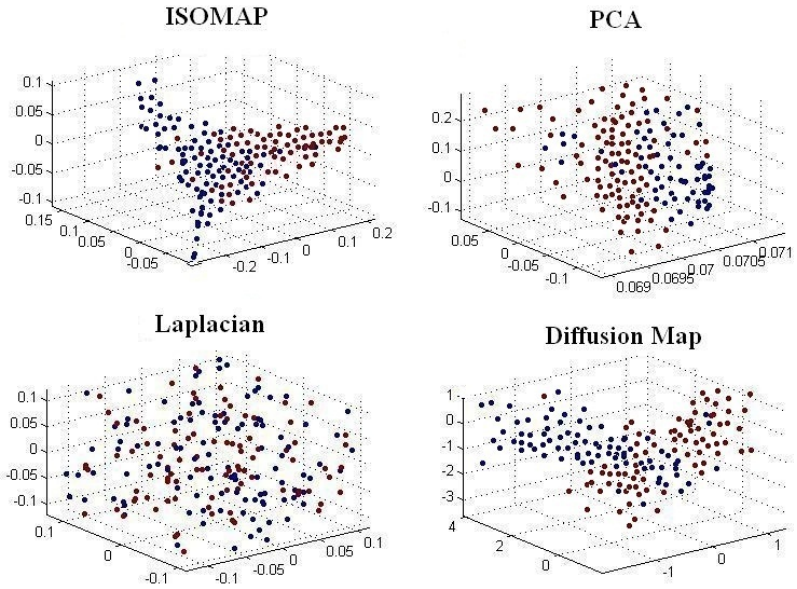


Figure 5.10.: A reduced 3D representation of spectroscopy $2048D$ dataset. The worst selection of parameters for all four methods. Non of the method produced clear clustering of the dataset

eigenmaps clustering accuracy improves to 96%. The parameters shown in the table are the only best combenation for our dataset.

5.6. Concluding Remarks

In this chapter, we applied four manifold learning techniques to the problem of dimensionality reduction and clustering of optical spectroscopic data in dermatology. In contrast to the linear method PCA, all studied manifold learning techniques were able to perform satisfactorily in clustering normal skin mole from malignant skin lesions, provided the parameters were chosen correctly. While especially Laplacian Eigenmaps look very promising for the intended dermatological application.

5.7. Discussion on Feasibility of Optical Spectroscopy for skin lesions classification

In this part of the thesis we analyzed optical spectroscopy for skin lesions classification. Optical spectroscopy by itself produces data, which, due to its high-dimensionality, cannot be directly utilized for classifying skin lesions. In other words, distinguishing between malignant and benign skin lesions is difficult. First the dimension of the data needs to be reduced in a meaningful way. In this respect, we introduced the application of manifold learning techniques to the problem of dimensionality reduction and clustering of spectroscopic data in dermatology. One other problem in dermatology is about quantifying the progress of skin lesions. For this purpose, one needs to be able to numerically compare two or more images of e.g. the same lesion taken during different sessions. This involves accurate registration of all those images. We presented a combination of optical spectroscopy with tracking as a solution to this problem. In our approach, this combination is used as a guidance for acquiring spectral measurements at the same positions and orientations as the first acquisition. We defined several spectroscopic data acquisition protocols for using our system optimally. We also evaluated a patient dataset with an SVM-based classification of skin lesions.

Our system opens a new way for utilizing the real potential of optical spectroscopy for noninvasive diagnosis of skin lesions. In taking optical spectroscopy even one step further using our system, we were able to show that it is a promis-

ing technique for the discrimination of malignant skin lesions from benign ones. Spectroscopy could form the basis of a clinical method to diagnose skin lesions due to the accuracy and reproducibility of its measurements. Acquisition of spectroscopic data causes little or no patient discomfort, does not alter the basic physiology of the skin, poses no hazard to the patient and does not interfere with any other standard clinical diagnostic practices. The scan could be performed by a non-specialist and therefore might be a useful tool for the prescreening of skin lesions. However, before full integration of spectroscopy into the clinical workflow, some further challenges need to be addressed:

- Mole mapping is an essential part in computer aided diagnostic system. It is observed that in our tracking system it is very difficult to place the tracking target on the same location as in the first acquisition. As a solution in a realistic setup, this can be replaced by the high accuracy non-invasive patient registration methods like the ones being developed for radiation therapy and navigated surgery [92]. The future work would be the evaluation of surface registration based strategies. Using a surface registration method, the positioning of the tracking target could be chosen arbitrarily as long as it can be mapped rigidly to a position in the previous examination.

- From our experience, there is need for several spectroscopic probes with different diameter sizes in order **i)** to cover only the area relevant to the lesion during the acquisition, i.e. to avoid getting measurements from the healthy skin region around the lesion and **ii)** to avoid multiple scans of the same lesion.

- In our experiments, we have observed that different samples taken from the same mole led to different spectral readings. A method is required to create a representative measurement from multiple spectroscopic readings for each mole.

- Optical spectroscopy based skin lesion diagnosis systems should be patient specific, since every patient has their own individual pattern of lesions which can be monitored throughout his/her body moles. In our study, we have observed that it is important to perform the classification within patient specific data in order to build a reliable system.

- Combining optical spectroscopy with other imaging technologies, e.g. der-

moscopy imaging, multispectral imaging and hyperspectral imaging, can improve the diagnosis further, since the optical spectroscopy provides complementary information to these techniques.

- Patient age is an important factor which needs to be taken into account during the acquisition of optical spectroscopy data. As the cellular structures can change according to the age of the patient, differences in spectroscopic readings have been observed between young and elderly people, which can be addressed by creating groups of patients accordingly.

- Accurate data acquisition requires constant contact of the probe with the surface of the lesion which is hindered in some cases by ragged skin lesions. Further studies are required to investigate new techniques for data acquisition without touching the skin surface.

- A more in-depth study on data sets with larger variation is required to demonstrate general utility of optical spectroscopy in the clinical setting. Especially, data accompanied by pathological verification of malignant melanoma would be highly desirable to demonstrate the reliability of the presented methods.

Part III.

Modeling Visual Assessment of Dermatologist

CHAPTER 6

Dermoscopic Images classification

DIAGNOSIS of benign and malign skin lesions is currently mostly relying on visual assessment and pathology performed by dermatologists. As the timely and correct diagnosis of these skin lesions is one of the most important factors in the therapeutic outcome, leveraging new technologies to assist the dermatologist seems natural. In this part of the thesis we propose a machine learning approach based on modeling the visual assessment of dermatologist to classify melanocytic skin lesions into malignant and benign from dermoscopic images. The dermoscopic database is composed of 42.911 patients skin lesion image from the department of dermatology, Klinikum Rechts der Isar München.

6.1. Computerized diagnosis of dermoscopic images: State of the art

Computer aided image diagnosis for skin lesions is a comparatively new research field. While the first related work in the medical literature backdates to 1987, the contribution was limited since by that time computer vision and machine learning were both developing fields [23]. One of the first compelling contributions from the image processing community was reported from H. Ganster et. al. [58]. In there work Ganster et al. proposed a classical machine learning approach for dermoscopic image classification. The first stage is automatic, color–based lesion seg-

mentation. Then, more than hundred features are extracted from the image (gradient distribution in the neighborhood of the lesion boundary, shape and color). Feature selection was obtained using sequential forward and sequential backward floating selection. Classification is performed using a 24–NN classifier and delivered a sensitivity of 77% with a specificity of 84%.

Up to our knowledge the best results in semi–automated melanocytic lesion classification where obtained by G. Capdehourat et al. [21]. The image database is composed of 433 benign lesions and 80 malignant melanoma. The learning and classification stage is performed using AdaBoost.M1 with C4.5 decision trees. For the automatically segmented database, classification delivered a false positive rate of 8.75% for a sensitivity of 95%. The same classification procedure applied to manually segmented images by an experienced dermatologist yielded a false positive rate of 4.62% for a sensitivity of 95%.

A summary of the results obtained by the key studies from the decade is presented by Alexander Horsch [78]. This study emphasise on methods of dermoscopic image analysis, commercial dermoscopy systems, evaluation of systems, and there methods. Comparison of some diagnostic system along with their database sizes are discussed by Celebi et al. [25], in there article they also proposed an approach which is based on a simple machine learning methodology. After an Otsu–based image segmentation, a set of global features are computed (area, aspect ratio, asymmetry and compactness). Local color and texture features are computed after dividing the lesion in three regions: inner region, inner border (an inner band delimited by the lesion boundary) and outer border (an outer band delimited by the lesion boundary). Feature selection is performed using Relief [138] and the Correlation-based Feature Selection (CFS) algorithms [67]. Finally, the feature vectors are classified into malignant and benign using SVM with model selection [134]. Performance evaluation gave a specificity of 92.34% and a sensitivity of 93.33%.

Our contribution in this regards is a complete characterization of a skin lesions into a feature vector that contains shape, color and texture information, as well as local and global parameters that try to reflect structures used in medical diagnosis by dermatologists. The learning and classification stage is performed using SVM with polynomial kernels. The classification delivered accuracy of 98.57% with a true positive rate of 0.981 and a false positive rate of 0.019.

6.2. Database

The images of the lesions were obtained by using a digital epiluminescence microscopy system (MoleMax) department of dermatology, Klinikum Rechts der Isar, Technischen Universität München, Germany. Images of lesions were taken with resolution of $640x480$ pixels and stored in 24-bit resolution .jpg formate. the dataset is from patients going for routine checkup since 2000 and there is no record of patients follow-up visits. Our database is composed of 42,911 images of melanocytic lesions out of which we randomly select 7,472 images for labeling. The images are shown (using computer monitor) to four dermatologist, in individual sessions for labeling and we selected the one on which all 4 doctors have consensus. Labeling of the images was into two classes: malignant melanoma and non malignant. out of 7,472 images 532 are diagnosed as malignant and 9 out of 532 were with histopathology record. We know that our database is not according to goal standers mention by Barbara Rosado [130], but it is important to note that we are modeling the visual assessment of dermatologist and our database is synchronize accordingly.

6.3. Segmentation

In order to segment the given image data we adapt the method described by Li et al. in [90]. Let $\Omega \subset \mathbb{R}^2$ denote the image domain. Then we define two soft–labeling functions $u_{1,2} : \Omega \to [0, 1]$ which can be used to define three soft membership functions

$$M_1 = u_1 u_2, \quad M_2 = u_1(1 - u_2), \quad M_3 = 1 - u_1. \tag{6.1}$$

These membership functions provide a soft partitioning of the image domain, because $M_1(x) + M_2(x) + M_3(x) = 1$ holds for all $x \in \Omega$, and allow us to segment the image domain into three areas indicating healthy skin, bright parts of the melanoma, and dark parts of the melanoma. An example is shown in Figure 6.1.

The described partitioning of the image domain is obtained by minimizing the following convex energy

$$E = \frac{1}{2} \int_\Omega |\nabla u_1|^2 + |\nabla u_2|^2 \; dx + \lambda \sum_{k=1}^{3} \int_\Omega d_k M_k \; dx, \tag{6.2}$$

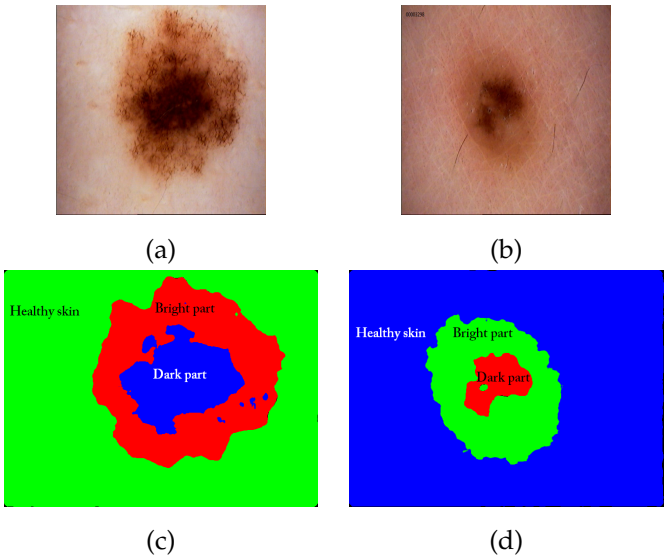


Figure 6.1.: Image (a),(b) Malignant melanoma and image (c),(d) segmented image in three areas

where

$$d_k = |a(x) - \bar{a}_k|^2 + \left|b(x) - \bar{b}_k\right|^2 . \tag{6.3}$$

Here $a, b : \Omega \to \mathbb{R}^3$ are the two color channels of the CIE Lab color space, while $\bar{a}_k, \bar{b}_k$ are the corresponding mean values:

$$\bar{a}_k = \frac{\int_\Omega M_k(x) a(x)\, dx}{\int_\Omega M_k(x)\, dx}, \quad \bar{b}_k = \frac{\int_\Omega M_k(x) b(x)\, dx}{\int_\Omega M_k(x)\, dx}. \tag{6.4}$$

The advantage of using the channels a and b of the CIE Lab color space is that these color channels only contain color and no luminance information making the segmentation more robust with respect to inhomogeneous lighting conditions. For all experiments we chose $\lambda = 2$. Please note that using an approach which minimizes a convex energy allows for a fully automatic segmentation of the data.

6.4. Feature Extraction

The feature extraction is the key point of the classification and has to be adequate in order to obtain a good system detection rate. We selected a group of features which attempts to represent the characteristics observed by the Physician. Each feature is following the idea of the ABCDE rules of skin cancer, which are:

- **A (Asymmetry):** Usually skin cancer moles are asymmetric instead of the normal moles, which are symmetric.
- **B (Border):** Usually the melanocytic lesions have blurry and/or jagged edges.
- **C (Color):** The melanocytic lesion has different colors inside the mole.
- **D (Diameter):** The lesions does not exceed a diameter of a pencil eraser (6 mm), otherwise it is suspicious.
- **E (Elevation):** When the mole is elevated from the normal skin it is suspicious.

Based on this technique, we created a set of features trying to characterize them via computer vision techniques. The list of features selected is as follows: geometric, color, texture and shape properties. The properties obtained by the feature extractor are totally based on the segmentation step and the features have to be independent of the image (size, orientation, etc.) in order to be robust with regard to the image acquisition. This feature property is very important because the physician can take the picture of the lesions in different ways, and lesions can have different sizes, too.

6.4.1. Geometric properties

In this section we try to achieve A and B part of the ABCD rule. From segmentation of the lesions, we obtain a binary image which represents the segmented blobs. Using this binary image, we get the bounding box and we fit an ellipse which has the same second inertia moment of area. Smaller blobs are erased from the binary image. Usually the biggest blob of the image is the segmented lesion and the sparse small moles are only segmentation noise. The bounding box is our metric for the standardization of the lesions. Using the bounding box and the fitted ellipse we reorient the lesions to the biggest ellipse axis and we resize the image to a standard size. The features used to represent the geometric properties are as follows:

- **Relative Area:** Area of segmented mole with respect to the bounding box area. This area represents the size of the mole.

- **Relative Filled area:** Area of the segmented mole with the internal holes filled w.r.t. the bounding box area. It represents how many internal areas of the mole were wrongly segmented.
- **Relative Centroid:** The centroid of the fitted ellipse w.r.t. the bounding box, indicating the distribution of the mole in the bounding box.
- **Eccentricity:** The fitted ellipse eccentricity which represents how circular the mole is.
- **Solidity:** The relation between the convex area and the blob area, representing how irregular the border of the mole is.

We use the fitted ellipse and bounding box to **pre–process** the mole in order to create standard size and orientation to make the classification more robust. The orientation of one mole always will be the same because we apply a reorientation based not only on the orientation of the ellipse, but also on the largest distance of the blob border with regard to the centroid. These properties allow us to reorientate the same mole with different angles to the same orientation as shown in Figure 6.2. The bounding box is resized to a square using the largest side as the value of the square which is cropped and resized to a standard value of 100×100. This standard size allows to compare different moles with different sizes and orientation.

6.4.2. Color properties

The mole color is very important in the classification because it encodes the variety of colors in the mole. When the mole has more colors, the mole has more chance to be malignant [29]. The colors are coded in a color histogram representing all the colors observed in the mole. The histogram is compacted in groups of values named **bins**. The bins allow us to reduce the number of $256^3 \approx 16M$ entries of a sparse histogram to a reasonably small dense histogram. This reduction has the advantage of encapsulating different ranges of colors in only one histogram value and being more robust on lighting changes, but with the disadvantage of losing color precision. The selected number is 8 generating $8^3 = 512$ possible values in the histogram. The color histogram is created using only the pixels of the segmented mole, excluding the skin pixels. The histogram is normalized with the total number of pixels used to create the histogram. In this way, we can compare histograms created from different sized moles.

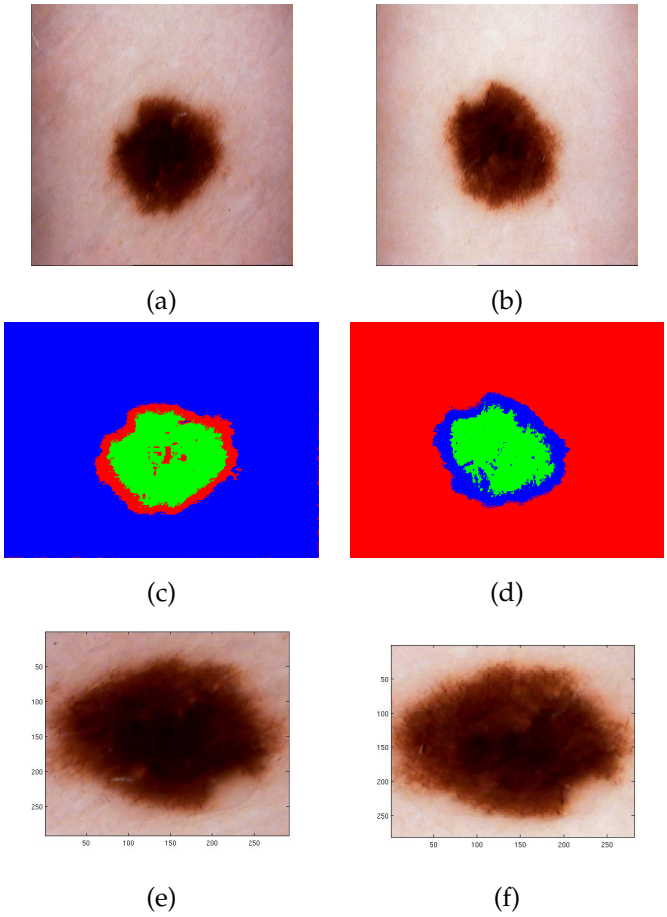


Figure 6.2.: Reorientation management: (a) First screening (b) Second screening (c) Segmentation of image a (d) Segmentation of image b (e) Reorientation of cropped image a (f) Reorientation of cropped image b

6.4.3. Texture properties

This feature describes the differences between the colors of the mole allowing to characterize the discontinuity in the mole colors, which is a tool used by physicians to recognize if a mole is malignant or not. To represent the texture, we use LBP (local binary pattern) of the image which creates a code of the color variability in the neighborhood of each pixel.

6.4.4. Shape properties

This feature represents the shape properties of the mole giving us how elliptic is the mole, circular or irregular, which is a very important feature in the classification of a mole. This section covers B and D of ABCD rule and feature is represented using histogram of oriented gradients (HOG), which counts the occurrences of gradients in portions of the image, coding the variability of the gradient in the image. This feature represents not only the shape of the moles, but also the mole uniformity given the internal shape when the color changes, which is detected by the gradient.

6.5. Classification

The goal of this stage is to classify the feature vectors in two classes: malignant and benign. Our feature vector size is 1682 for each image input. A classification technique that proved very successful in our experiments is support vector machines (SVM, [37]). SVM was selected as the method of choice as it allows to linearly classify data in a high–dimensional feature space that is non–linearly related to the input space via the use of specific polynomial kernels. To reduce the dimensions of the input feature set $x_i \in R^{1682}$, $i = 1, ..., n$, where n denotes the number of measurements (in our case 7472), principal components analysis (PCA) is applied. The resulting feature vector of eigenvalues $(e^i_j)_{j=1,...,1682}$ is sorted descendingly by magnitude. Since the highest eigenvalues represent the most relevant components, a cut–off value C_{PCA} is chosen, such that the final input data y_i for the classification algorithm from measurement x_i $(i = 1, ...n)$ is

$$y_i = (e^i_j)_{j=1,...,C_{PCA}} \tag{6.5}$$

The cut–off value C_{PCA} is chosen empirically, to represent 95% of feature vector of 1682 dimensions, which reduced it to 434 dimensions.

The SVM classifier needs to be trained first before using it, thus we partition our input feature sets (y_i) $i = 1, .., n$, into two partitions, $T \subset \{1, ..., n\}$ the training set and $V \subset \{1, ..., n\}$ the testing (or validation) set with $T \cup V = \{1, ..., n\}$ and $T \cap V = \emptyset$. The training data set T is labeled manually into two classes using the ground truth, $l(y_i) = \pm 1$. Once the classifier is trained, a simple evaluation of the decision function $d(y_i) = \pm 1$ will yield the classification of any data y_i.

In detail, SVM is trying to separate the data $\phi(y_i)$ mapped by the selected kernel function ϕ by a hyperplane $w^T \phi(y_i) + b = 0$ with w the normal vector and b the translation. The decision function then is $d(y_i) = \text{sgn}(w^T \phi(y_i) + b)$. Maximizing the margin and introducing slack variables $\xi = (\xi_i)$ for non-separable data, we receive the primal optimization problem:

$$\min_{w,b,\xi} = \frac{1}{2} w^T w + C \sum_{i \in T} \xi_i \tag{6.6}$$

with constraints $l(y_i)(w^t \phi(y_i) + b) \geq 1 - \xi_i,\ \xi \geq 0$ for $i \in T$. C is a user–determined penalty parameter. Switching to the dual optimization problem allows for easier computation,

$$\min_{\alpha} = \frac{1}{2} \alpha^T Q \alpha - e^T \alpha \tag{6.7}$$

with constraints $0 \leq \alpha_i \leq C$ for $i \in T$, $\sum_{i \in T} y_i \alpha_i = 0$. The $\alpha = (\alpha_i)$ are the so–called support vectors, $e = [1, ...1]^T$ and Q is the positive semidefinite matrix formed by $Q_{jk} = l(y_j) l(y_k) K(y_j, y_k)$, and $K(y_j, y_k) = \phi(y_j)^T \phi(y_k)$ is the kernel function built from ϕ. Once this optimization problem is solved, we determine the hyperplane parameters w and b, w directly as $w = \sum_{i \in T} \alpha_i l(y_i) \phi(y_i)$ and b via one of the Karush-Kuhn-Tucker conditions as $b = -l(y_i) y_i^T w$, for those i with $0 < \alpha_i < C$. Thus the decision function of the trained SVM classifier ends up as

$$d(y_i) = \text{sgn}\big(w^T \phi(y_i) + b\big) = \text{sgn}\left(\sum_{j \in T} \alpha_i l(y_i) K(y_j, y_i) + b \right). \tag{6.8}$$

6.6. Results

Performance evaluation was conducted using a 10–fold cross–validation. The 10–fold cross–validation gives an approximation of the general classifier perfor-

Table 6.1.: Results of the 10 random balanced data sets, and for each dataset 10–fold cross–validation using a SVM classifier (Avg-Std 98.545 ± 0.046).

Variables	Test-1	Test-2	Test-3	Test-4	Test-5
Correctly Classified Instances	98.5772%	98.5743%	98.5765%	98.5614%	98.5167%
Incorrectly Classified Instances	1.4228%	1.4257%	1.4235%	1.4386%	1.4833%
True Positives Rate	0.991%	0.996%	0.993%	0.997%	0.995%
False Positives Rate	0.019%	0.023%	0.034%	0.025%	0.021%
————————	Test-6	Test-7	Test-8	Test-9	Test-10
Correctly Classified Instances	98.4982%	98.5765%	98.5965%	98.4624%	98.5017%
Incorrectly Classified Instances	1.5018%	1.4235%	1.4235%	1.5376%	1.4983%
True Positives Rate	0.981%	0.991%	0.983%	0.991%	0.996%
False Positives Rate	0.059%	0.033%	0.064%	0.020%	0.13%

mance. We created 10 balanced data sets which were generated from the original unbalanced data set of 6840 benign and 532 malignant lesion. The balanced data sets were generated by selecting randomly a similar number of benign and malign images (532) to obtain a more general and balanced training dataset. We assess the feature training and perform 10–fold cross–validation utilizing the 10 balanced datasets. The results of these data sets are shown in Table 6.1

The results show a very good performance in all the random data sets, allowing us to conclude that the selected feature vector of the moles gives meaningful information about the mole in the classification. The correctly classified instances value indicates a performance over 98% in all 10 tested cases and an error of less than 2%. If we observe only the malignant classification, which is the most important, the performance shows a true positives rate greater than 99%, meaning that the classifier recognizes as malignant 99% of the skin cancer moles. Therefore, the number of malignant moles which are not correctly classified is 1%. In addition, the false positives rate is smaller than 3%, showing that the misclassification of the benign images are only 3 in a total of 100 benign images. In the case of recognizing the malignant moles it is important to detect most of the malignant moles even if

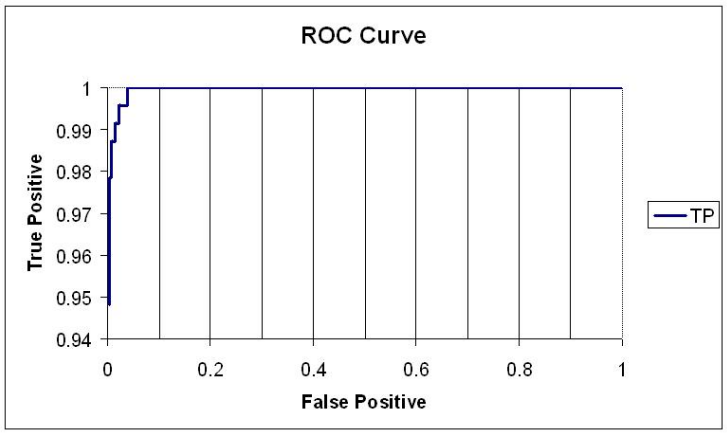


Figure 6.3.: Receiver Operating Characteristic response

the false positive rate is not small, as it is less harmful for the physician to label the mole as suspicious even though it is not.

Receiver operating characteristic (ROC)[71][72] curves were used to determine the performance of the discriminant rules. The area under the ROC curve is a good indication of diagnostic accuracy and should be used when comparing different classification techniques [147].The Figure 6.3 shows the ROC response of our classifier and its consequential performance, having a curve near to the ideal case. The classifier has a high area under the curve being near to 0.99, where the maximum is 1. We believe that our feature vector is a good representation of the dermoscopy characteristics following the ABCDE rule used by the dermatologist in skin cancer diagnosis.

6.7. Summarizing

In this chapter we proposed a methodology for computer–aided classification of dermoscopic images. The learning and classification stage is performed using SVM. According to our medical partners the results are satisfactory and for further research the system can be deployed in dermatology. Concerning our algorithm, to further improve its performance, methods to detect a larger number of geometric or texture–based structures, similar to those used in the 7–points checklist, should be developed. The next important step is sub–classification of malignant categories, which is ongoing research. A rigorous study of this topic,

complemented with the comparison of the weights assigned to visual features in the ABCD and other clinical diagnosis rules, may yield useful recommendations to dermatologist for their medical practice.

CHAPTER 7

Performance Comparison Among Different Models for Computer-aided Skin Lesions Classification

In recent years, computer-aided diagnosis systems have been used in several hospitals and dermatology clinics, targeting mostly the early detection of skin cancer, and more specifically, the recognition of malignant melanoma lesions. In this chapter, we review the state of the art of such systems by presenting the statistical results of the most important implementations that exist in the literature, while comparing the performance of several classifiers on the specific skin lesion diagnostic problem and discussing the corresponding results.

7.1. Comparisons Among Different Computer-aided Diagnosis Systems in Dermatology

The scientific community take an interest in building classification models based on supervised learning due to the increase in electronic medical databases. In literature, numerous machine learning (supervised, unsupervised) and statistical approaches for classification are available, but few comparisons among different models have been done on the same datasets. The potential advantages and disadvantages have been defined theoretically for each of these methods, given cer-

tain assumptions about characteristics of the classification task, data distribution, signal-to-noise ratio, etc., it is often the case that in routine these belief cannot be validate [45]. Under these paradigm, observational comparison of classification performance using standard versification to describe difference and measurement is essential. Function affiliate with misclassification in any direction (e.g., false positives or false negatives) can be built into the models, or treated separately. A final selection of the best model for a given classification task can only be achieve after acknowledge the tradeoffs between classification performance, costs and model illustratesion [45].

Our focus of investigation is the selection of a class of models for a particular dataset. We compare the prejudicial performance of five methods (support vector machines, decision trees, k-nearest neighbors, artificial neural networks, and logistic regression) on the task of classifying pigmented skin lesions (PSLs) as benign or melanoma. The input dataset is same for all models.

The task of classifying PSLs is complicated and involves automated feature calculation extract from digital images, as well as clinical and demographical data collected by dermatologists. The reason for using a PSL dataset as proving ground for the classification algorithms is the fact that the occurrence of melanoma has risen greatly in recent years. Therefore, to a greater extent important to flawlessly diagnose PSLs. The classification task is rigid, as can be seen from the fact that the diagnostic performance of even expert dermatologists is currently far from optimal, the average number of lesions excised per histologically-proven melanoma is 30 [78], which is quite high number. A study by Curley et. al. [38] showed that for three experienced dermatologists the diagnostic accuracy for clinically evaluating pigmented lesions was only 50% when compared with the histological diagnosis. Epiluminescence microscopy was developed as a tool to aid in the diagnostic process, and expert performance increases when using this method [10] [45].

7.2. Materials and Methods

In this section we will examine the definition of features, the most popular methods for skin lesion classification and their results.

7.2.1. Feature Extraction

Computer-based systems are not much different from the conventional visual diagnosis process of dermatologist as regarding feature extraction. The features exploit have to be quantifiable and of high sensitivity, i.e., high correspondence of the feature with skin cancer and high probability of true positive response. Furthermore, the features should have high specificity, i.e., high probability of true negative response. Although in the typical classification model both factors are very important (a tradeoff expressed by maximizing the area under the receiver operating characteristic (ROC) curve), in the case of malignant melanoma detection, the abolition of false negatives (i.e., increase of true positives) is apparently more important. In the conventional procedure, the following diagnosis methods are mainly used [64] (1) ABCD rule of dermoscopy [103][101]; (2) Pattern analysis [89]; (3) Menzies method [46]; (4) Seven-point checklist [4][9]; and (5) Texture analysis [2]. It is a fact that most of the patterns that are used by the pattern analysis, the Menzies method, and the seven-point checklist are very rarely used for automated skin lesion classification, undoubtedly due to their complicatedness. We selected a group of features which attempts to represent the characteristics by following the idea of the ABCD rule [99]. For further details go to section 6.4.

7.2.2. Skin Lesion Classification Methods

In this section we will explain classification models which are used in our experiments.

7.2.2.1. k-Nearest Neighbors

The k-nearest neighbors (kNN) algorithm is a method for classifying objects based on closest training examples in the feature space [47][59]. In comparison to the other methods, the kNN algorithm does not implement a decision boundary, but uses the essential feature of the training set to approximate calculate the density dispersion of the data. They essentially combine this information with class prevalence in the Bayes-rule to obtain the rear(class membership) probability estimates of a data point. The density estimation uses a distance measure (usually Euclidean distance). For a given distance calculated, the only parameter of the algorithm is k, which is the number of neighbors. The parameter k decide the smoothness of the density estimation: bigger values reflect more neighbors, and therefore smooth

over local characteristics. Same values reflect only limited neighborhoods. Generally, the choice of k can only be identify empirically [45]. In our test, we used values of k = 10, 20,...,100. In most of the medical diagnosis systems, k-nearest neighbor algorithms are used as benchmarks for other machine learning techniques [70][50][42].

7.2.2.2. Logistic Regression

The Logistic regression (sometimes called the logistic model or logit model) is an algorithm that assemble a disconnect hyperplane between two data sets, using the logistic function to signify distance from the hyperplane as a presumption of class membership. Logistic regression is extensively used in medical systems for the ease with which the parameters in the model can be illustrate as changes in log odds, for the variable choice methods that are often available in commercial implementations, and for allowing the interpretation of results as probabilities [45]. Although the model is linear-in-parameters and can thus only compute linear decision boundaries, it is a consume predictive model in medical applications [48][143]. In our experiments, the weka open-source classification tool has been used (available from the University of Weka [66]) to derive logistic regression models.

7.2.2.3. Artificial Neural Networks

Artificial neural networks (ANNs) are networks of interconnected nodes composed of various stages that emulate some of the observed properties of biological nervous systems and draw on the analogies of adaptive biological learning. ANNs represent a means to calculate posterior class membership probabilities by minimizing a cross-entropy error function [11][153]. The ANN belongs of several small processing units (the artificial neurons) that are highly link. Information flow in an ANN is modeled the human brain, in that information is pass on between neurons, with the information stored as connection power (called weights) between neurons. The minimization process is implemented as an update rule for the weights in the network. For medical applications, a considerable disadvantage of ANNs is the fact that the parameters in the model are not directly explainable, so that no more understanding of a data set can be derived from a neural network model [45].

7.2.2.4. Decision Trees

Decision trees are an important technique in machine learning techniques. It organize classifiers by dividing the dataset into smaller and more equal groups, depend on a measure of disparity (usually entropy) [128]. Decision Trees does it by finding a variable and a threshold in the domain of this variable that can be used to divide the dataset into two groups. The best pick of variable and threshold is the one that minimizes the disparity measures in the resulting groups. The benefit of decision trees over many of the other methods used here is that small decision trees can be examine by humans as decision rules [45]. Decision trees therefore offer a way to extract decision rules from a database. This makes it particularly well suited for medical applications, and advantages and disadvantages of decision trees in medicine have been extensively investigated [30][109][162].

7.2.2.5. Support Vector Machine

Support vector machines (SVM) is a systems which use hypothesis space of a linear functions in a high dimensional feature space, trained with a learning algorithm from optimization theory that implements a learning bias derived from statistical learning theory [16][18][154]. The most appealing feature of this paradigm is that it is possible to give bounds on the generalization error of the model, and to select the best model from a class using the principle of structural risk minimization [154]. Support vector machine compute dichotomize hyperplanes that maximize the margin between two sets of data points. By using lagrange multipliers, the problem can be formulated in such a way that the only working on the data points are the computation of scalar products [27]. While the fundamental training algorithm can only construct linear separators, kernel functions can be used to compute scalar products in higher dimensional spaces. If the kernel functions are nonlinear, the dichotomize boundary in the original space will be nonlinear. Because there are various kernel functions, there is a wide variety of possible SVM models. In our testing, we used SVM polynomial kernels of degrees 1 to 3 and radial basis function kernels with γ (inverse variance) parameters between 10^{-2} and 10^{-6} [45].

7.3. Results From Existing Systems

Building of automated systems for the diagnosis of skin lesions is acknowledge as a substantial classification task, which consumes many biomedical laboratories and research groups; e.g. [149], [132], [28], [111], [94] and [133]. Most of the surveyed systems focus on the detection of malignant melanoma and its separation from dysplastic or common nevus. However, there exist systems that are used for the identification of different lesions. These lesions include among others tumor, crust, scale and ulcer [153][152], erythema [122], burn scars [151] and wounds [76] [73].

The most ordinary classification models are the statistical and rule-based ones; e.g. [132], [14],[12] and [144]. More advanced techniques such as neural networks are presented in works like [52], [13], [131], [77], [146] and [161], while the *k*-nearest neighborhood classification scheme is applied in [59]. Classification and regression trees (CART) [17] have been used in [158].

The favorable outcome rates for the methods presented in the literature signifies that the work toward automated classification of lesions and melanoma, in specific, may provide good results. Accuracy rates can vary from 70% [159] to 95% [163], whereas sensitivity can score between 82.5% [28] and 100% [1] and specificity between 63.65% [1] and 91.12% [159], respectively. SVM seems to achieve higher performance in terms of sensitivity and specificity [99].

7.4. Experiments

We use the dataset as described in section 6.2 with addition of one extra class (Dysplastic nevus). A total of 8000 PSL images in three classes (common nevi, dysplastic nevi, melanoma) were selected. The distribution of cases in the dataset is 6940 common nevi, 528 dysplastic nevi, and 532 melanoma. We create 10 balanced data sets which are generated from the original unbalanced dataset of 8000 PLS. The balanced datasets were generated by selecting randomly a similar number (500) of common nevi, dysplastic nevi and melanoma images to obtain a more general and balanced training dataset. Each of the five algorithms presented above was run on each of the 10 different datasets as training and test data. A subset of them, e.g., 80% of the images, is used as a learning set, and the other 20% of the samples is used for testing using the trained classifier.

The most usual classification performance assessment in the context of melanoma

detection is the true positive (TP) illustrate the fraction of malignant skin lesions correctly classified as melanoma and the true negative (TN) illustrate the fraction of dysplastic or non-melanoma lesions correctly classified as non-melanoma, respectively [127][128]. A graphical representation of classification performance is the ROC curve, which displays the adjustment between sensitivity (i.e., actual malignant lesions that are correctly identified as such, also known as TP) and specificity (i.e., the proportion of benign lesions that are correctly identified, also known as TN) that results from the overlap between the assessment of lesion scores for melanoma and common nevi [1][51][120]. A good classifier is one with close to 100% sensitivity at a threshold such that high specificity is also acquire. The ROC for such a classifier will plot as a steeply rising curve. When different classifiers are compared, the one whose curve rises fastest should be optimal. If sensitivity and specificity were weighted equally, the greater the area under the ROC curve (AUC), the better the classifier is [65]. The Area Under Curve (AUC) is equal to the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one [54] [45].

Consider a binary classification task with m positive examples and n negative examples. Let G be a fixed classifier that outputs a strictly ordered list for these examples. Let $x_1, ..., x_m$ be the output of G on the positive examples and $y_1, ..., y_n$ its output on the negative examples and denote by 1_Z the indicator function of a set Z. Then, the AUC, A, associated to G is given by [35] [45]:

$$A = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} 1_{x_i > y_j}}{mn} \tag{7.1}$$

AUC is closely related to the ranking quality of the classification. It can be viewed as a measure based on pairwise comparisons between classifications of the two classes. It is an estimate of the probability P_{xy} that the classifier ranks a randomly chosen positive example higher than a negative example [69]. With a perfect ranking, all positive examples are ranked higher than the negative ones and $A = 1$. Any deviation from this ranking decreases the AUC.

For our experiments we use Weka [66], an open source software issued under the GNU General Public License. In some classification methods Weka does not support multi-class problems directly. We therefore reduced the problem to two dichotomous classification tasks: First, to discriminate common nevi from the other two lesion types (dysplastic nevi and melanoma), and second, to discriminate melanoma from common and dysplastic nevi. Standard ROC analysis was

Table 7.1.: Performance comparison of k-nearest neighbors parameters for the task of distinguishing melanoma from common and dysplastic nevi

	$k = 10$	$k = 40$	$k = 70$	$k = 100$
Avg AUC	0.6758	0.9014	0.8431	0.7873
Std dev	0.0357	0.0244	0.0416	0.0432
Min AUC	0.5773	0.8672	0.7833	0.6779
Max AUC	0.7668	0.9356	0.9029	0.8968
Avg sens	0.7195	0.8535	0.8173	0.8173
Avg spec	0.7429	0.9283	0.9182	0.9182

used to summarize the results of classification tasks [71][72].

A compression of the results of all five methods are given in Table 7.2, Table 7.3 and Table 7.4. The entries in the tables are the following, for each method and task: Average AUC over 10 data sets, standard deviation of AUC, maximum and minimum AUC value, as well as average maximum sensitivity and specificity. Furthermore, k-nearest neighbors results with different parameters are shown in Table 7.1. The performance comparisons of the SVM kernels are shown in Table 7.6 and Table 7.5. ROC curves for the best and worst methods (support vector machines and decision trees, respectively) are shown in Figure 7.1 and Figure 7.2. We show only these curves, since the results of the other methods lie between those of decision trees and support vector machines. In particular, since the results of logistic regression and neural networks are almost the same as those of support vector machines, their ROC curves are visually indistinguishable [45]. We now briefly discuss the results of the different methods on the data sets.

7.4.1. k-Nearest Neighbors

Standard euclidian distance was used on vectors as a distance metric for this method and the data had been normalized to zero mean and unit variance, every variable provide equally to the distance measure. It is important to consider that the *k*-nearest neighbors algorithm is very fast on this problem; i.e., the classification results differ with the choice of parameter *k*. The results for $k = 20$, $k = 30$, $k = 50$, $k = 60$, $k = 80$, and $k = 90$ are not displayed in Table 7.1 because these with small changes from those shown in the Table 7.1. The AUC results over all

values of k ranged from 0.6758 (k = 10) to 0.9014 (k = 40). In Table 7.2, Table 7.3 and Table 7.4, it can be seen that the results of the k-nearest-neighbors algorithm are only slightly inferior (3 to 4 percentage points) to those of the better methods.

7.4.2. Decision Trees

For the task of classifying PSL images decision trees are not suitable, it can be seen from the results in Table 7.2, Table 7.3 and Table 7.4. The justification for this is that approximately all the variables in the data set represent continuous data. This makes it hard to find the ideal thresholds needed to construct the decision tree. Given this basic drawback, it is not astonishing to see that decision trees perform unfertile of all the methods tested for this dataset. The main benefit that this paradigm has over the other methods is the human understanding of the results, the trees themselves are not applicable in this domain, since the input variables are machine generated [45].

7.4.3. Logistic Regression

Even though logistic regression is a linear-in-parameters algorithm that can only implement linear dichotomize hyper-planes between data points, it is nonetheless extensively used in medicine applications. The two main benefit this method has over other algorithms is its lack of difficulty for use and its variable-selection ability [45]. In all classification tasks (shown in Table 7.2, Table 7.3 and Table 7.4), logistic regression accomplish on about the same level as artificial neural networks and support vector machines, which are both adequate of implementing nonlinear dichotomize surfaces.

7.4.4. Artificial Neural Networks

Artificial neural networks is one of the most advantageous technology in the last two decades which has been widely used in a large variety of applications in differing areas. The early implementations depend upon a compelling amount of parameter tuning to achieve satisfactory results, a process that needed too much time and expertise for a unprofessional. Over the past few years, artificial neural networks and implementations of faster learning algorithms have allowed unprofessional, the use of cosmopolitan methods that needs little to no parameter-tuning [11]. For the experiments in this work, we used a multilayer perceptron algorithm

Table 7.2.: Performance comparison of *k*-nearest neighbors, logistic regression, artificial neural networks, decision trees, and support vector machines for the task of distinguishing common nevi from dysplastic nevi and melanoma

					SVM	
	Decision trees	*k*-NN	Log regression	ANN	Polynomial	RBF
Avg AUC	0.6657	0.7647	0.8408	0.8427	0.8340	0.8531
Std dev	0.0304	0.0301	0.0245	0.0157	0.0195	0.0198
Min AUC	0.6147	0.7285	0.7973	0.7994	0.7827	0.8139
Max AUC	0.7167	0.8010	0.8843	0.8861	0.8854	0.8923
Avg sens	0.7137	0.7389	0.7868	0.7714	0.7459	0.7967
Avg spec	0.6981	0.7397	0.7897	0.7417	0.7081	0.7539

Note. For nearest neighbors, $k = 40$. For SVM, the optimal polynomial kernel was linear, with $C = 100$, and the optimal RBF kernel had inverse variance $\gamma = 10^{-4}$ and $C = 100$.

Table 7.3.: Performance comparison of *k*-nearest neighbors, logistic regression, artificial neural networks, decision trees, and support vector machines for the task of distinguishing melanoma from common nevi and dysplastic nevi

					SVM	
	Decision trees	*k*-NN	Log regression	ANN	Polynomial	RBF
Avg AUC	0.7907	0.9014	0.9405	0.9542	0.9117	0.9601
Std dev	0.0565	0.0244	0.0273	0.0122	0.0376	0.0132
Min AUC	0.6951	0.8672	0.9011	0.9193	0.8537	0.9312
Max AUC	0.8863	0.9356	0.9790	0.9892	0.9698	0.9891
Avg sens	0.8051	0.8535	0.9432	0.9234	0.8684	0.9134
Avg spec	0.8362	0.9283	0.9452	0.9425	0.8953	0.9581

Note. For nearest neighbors, $k = 40$. For SVM, the optimal polynomial kernel was linear, with $C = 100$, and the optimal RBF kernel had inverse variance $\gamma = 10^{-4}$ and $C = 100$.

Table 7.4.: Performance comparison of *k*-nearest neighbors, logistic regression, artificial neural networks, decision trees, and support vector machines for the task of distinguishing dysplastic nevi from melanoma and common nevi

					SVM	
	Decision trees	*k*-NN	Log regression	ANN	Polynomial	RBF
Avg AUC	0.7055	0.8564	0.8984	0.9278	0.9229	0.9379
Std dev	0.0492	0.0368	0.0164	0.0184	0.0273	0.0132
Min AUC	0.6165	0.8146	0.8783	0.9031	0.8976	0.9147
Max AUC	0.7946	0.8983	0.9185	0.9525	0.9483	0.9612
Avg sens	0.7526	0.7953	0.8929	0.9073	0.9834	0.9398
Avg spec	0.8827	0.8723	0.9582	0.9261	0.8523	0.9141

Note. For nearest neighbors, $k = 40$. For SVM, the optimal polynomial kernel was linear, with $C = 100$, and the optimal RBF kernel had inverse variance $\gamma = 10^{-4}$ and $C = 100$.

[123] that required no additional parameters to be set. We used 20 nodes in the hidden layer; sample runs with 30 nodes showed similar results [45]. The results obtained by neural networks were in the same range as those of logistic regression and support vector machines. The training times were comparable to most of the other methods as well, with only a few seconds for each of the 10 dataset.

7.4.5. Support Vector Machine

As SVM only implement dichotomize hyperplanes, they can effectively construct nonlinear decision boundaries by mapping the data into a higher-dimensional space in a nonlinear manner by using kernel functions. Since it is impractical to identify in advance which kernel function works best for which dataset, enormous time is spent on trying different kernel functions. The popular kernel functions are polynomials and radial basis functions (RBF) [45].

The adjustable parameter for polynomial kernels, is the degree of the polynomial; for RBF kernels, it is the inverse variance. For any kernel function, it is also important to specify a cost factor C that determines the importance of misclassification on the training set.

Table 7.5.: Performance comparison of different SVM models for the task of distinguishing common nevi from dysplastic nevi and melanoma

	Polynomial kernel			RBF kernel		
	d=1	d=2	d=3	$\gamma = 10^{-6}$	$\gamma = 10^{-4}$	$\gamma = 10^{-2}$
Avg AUC	0.8340	0.7536	0.8086	0.8182	0.8531	0.8308
Std dev	0.0195	0.0352	0.0334	0.0253	0.0198	0.0153
Min AUC	0.7827	0.7163	0.7597	0.7785	0.8139	0.7954
Max AUC	0.8854	0.7909	0.8575	0.8579	0.8923	0.8663
Avg sens	0.7459	0.7482	0.7642	0.7692	0.7967	0.7735
Avg spec	0.7081	0.7195	0.7392	0.7354	0.7539	0.7627

Table 7.6.: Performance comparison of different SVM models for the task of distinguishing melanoma from common nevi and dysplastic nevi

	Polynomial kernel			RBF kernel		
	d=1	d=2	d=3	$\gamma = 10^{-6}$	$\gamma = 10^{-4}$	$\gamma = 10^{-2}$
Avg AUC	0.9117	0.8804	0.9051	0.9453	0.9601	0.9541
Std dev	0.0376	0.0362	0.0346	0.0274	0.0132	0.0262
Min AUC	0.8537	0.7854	0.7863	0.9453	0.9312	0.9328
Max AUC	0.9698	0.9754	0.9849	0.9742	0.9891	0.9752
Avg sens	0.8684	0.7462	0.8370	0.9284	0.9134	0.9003
Avg spec	0.8953	0.8723	0.8950	0.9627	0.9581	0.9283

We used polynomial kernels of degrees 1 to 4, RBF kernels with parameter $\gamma = 10^{-7}, 10^{-6}, ..., 10^{-1}$, and cost factor parameter values of *C* = 200 and *C* = 2000. The results for both these cost parameter settings were similar, with the *C* = 200 models performing inconsiderably better than the others. Therefore, we report only results for *C* = 200. Training times were about an order of magnitude unhurried than for the ANN models, but still in the range of only a few minutes. For the polynomial kernels, convergence times rely upon the degree of the kernel polynomial, with times for degree four kernels too slow to be included here [45]. RBF kernels are generally fast to converge and it also did not depend as heavily on the choice of precision parameter γ.

To examine SVM kernels we present results of the SVM classification tasks in more detail in Table 7.6 and Table 7.5 . For the polynomial kernels, it is substantial to note that the linear kernel function performs good than the polynomial kernels of degrees 3 and 4. In light of the good performance of the logistic regression model, it is not astonishing that a linear model should do well. It is astonishing, however, that the higher-degree polynomial kernels did not perform at the same level. For the RBF kernels, the excellent results were obtained for $= 10^{-5}$. The classification performance decreases for smaller and larger values of γ . The results for $\gamma = 10^{-4}$ and $\gamma = 10^{-2}$ are not listed in the tables because they are less than the best results, but better than those for $\gamma = 10^{-7}$ and $\gamma = 10^{-8}$ respectively [45].

7.5. Discussion

Selecting a best model for a given classification task rely upon not only on discriminatory power, but also on other factors such as feature extraction, segmentation, cost of model construction and model interpretability. In this chapter, we aim completely on identifying the classification performance. In addition, five methods were check into thoroughly in this chapter, the top three (logistic regression, artificial neural networks, and support vector machine) give almost same results, whereas the other two (*k*-nearest neighbors and decision trees) drop off significantly on some of the classification tasks. Even the unsatisfactory of the five (decision trees) achieves sensitivity and specificity values that are comparable to human experts visual assessment [61] [45].

In this testing setup for the classification of PSLs, we use 10 different datasets for training and testing, it is not possible to check for statistically significant dif-

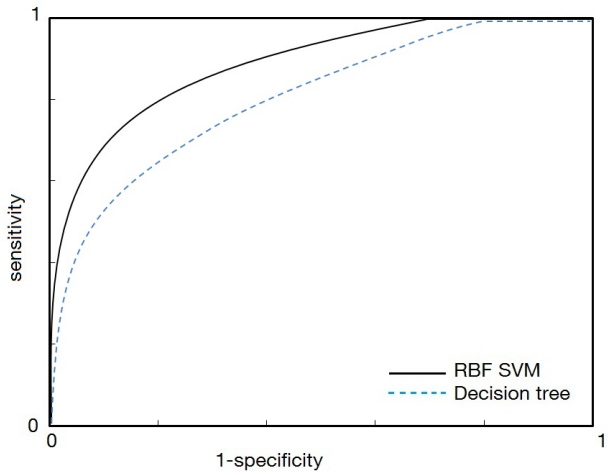


Figure 7.1.: Averaged ROC curves for support vector machines with RBF kernels and decision trees on the task of distinguishing common nevi from dysplastic nevi and melanoma. The AUC value is 0.8531 for the SVM and 0.6657 for the decision trees.

ferences in classification performance. This is due to the fact that the 10 different sets are highly correlated, and thus the results obtained from these runs are not independent. What can be said about the results of the runs is that the data set was large enough (or well-behaved enough) so that for almost all methods on all the tasks, there were no outliers in the results.

Support vector machines produce good results which shows that this method is going to be tested and used more frequently in medical domains. It seems to be a worth alternative to logistic regression and neural networks, especially since there are theoretical bounds on the generalization error in SVM models [154] [45].

7.6. Conclusion

We test different methods for modeling the visual assessment of the dermatologist using machine-learning paradigms on the problem of classifying pigmented skin lesions such as common mole, dysplastic mole, or melanoma. While the decision tree is not good for this problem domain (most of the input variables are continuous), the other methods performed good (*k*-nearest neighbors) to the best (logistic regression, artificial neural networks, and support vector machines) on

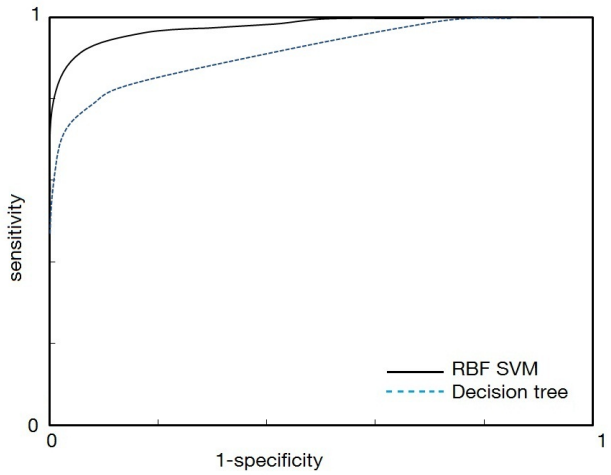


Figure 7.2.: Averaged ROC curves for support vector machines with RBF kernels and decision trees on the task of distinguishing melanoma from common and dysplastic nevi. The AUC value is 0.9601 for the SVM and 0.7907 for the decision trees.

the datasets.

Even though it is not the aim to replace dermatologists in the diagnostic procedure, the results of this work shows that decision support tools could be used to increase the performance of human experts. One possible area is in intelligent training tools. Such tools could be designed as tutoring systems for dermatologists, with large dataset of lesion images and gold standard diagnoses for these images. Trained models could then provide reference probability assessments and, for a given lesion from the repository, present lesions with similar degrees of malignancy. Similarity matching on the lesion features could also be used to present features that are not only similar in diagnosis, but also similar in appearance. Generation of a gold standard database is one of the important aspects for future work [45]. There is a need of further work to investigate these ideas in detail.

Part IV.

Final Conclusions

CHAPTER 8

Conclusions

THE present chapter summarizes the major contributions and findings of the research conducted in this thesis. Finally, it is concluded by discussing the remaining challenges, future directions and ideas worth investigating.

8.1. Summary

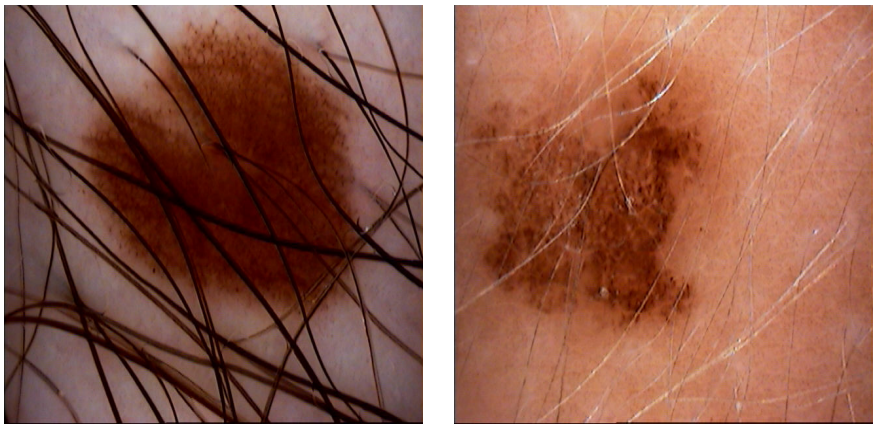
According to the literature [102], [75], it is often difficult to differentiate early melan-oma from other benign skin lesions. This task is not trivial even for experienced dermatologists, but it is even more difficult for primary care physicians and general practitioners [126]. On the other hand, the early diagnosis of skin cancer is of severe importance for the outcome of the therapeutic procedure and the basis for reducing mortality rates. The aim of this dissertation has been to gain a better understanding of optical spectroscopy and dermoscopic images for clinical use. In order to do so, the presented thesis has been divided in two successive parts.

In the first part of our thesis we presented a feasibility study of optical spectroscopy for the classification of skin lesions. We first explained the hardware setup of the optical spectroscopy system. Then we presented the method for combining optical spectroscopy with tracking for using it as a guidance to acquisition of spectral measurements. Following the data acquisition protocols we defined in the first part and using the SVM-based classification approach, we showed that

optical spectroscopy is a promising method for noninvasive diagnosis of skin lesions. Furthermore we saw that it has the potential to differentiate between benign and malignant skin lesions. Spectroscopy could form the basis of a clinical method to diagnose skin lesions due to the accuracy and reproducibility of its measurements. Acquisition of spectroscopic data causes little or no patient discomfort, does not alter the basic physiology of the skin, poses no hazard to the patient and does not interfere with any other standard clinical diagnostic practices. The scan could be performed by a non-specialist and therefore might be a useful tool for prescreening of skin lesions. Combining optical spectroscopy with other imaging technologies, e.g. dermoscopy imaging, multispectral imaging and hyperspectral imaging, can improve the diagnosis, since the optical spectroscopy provides complementary information to these techniques.

In the second part we present a machine learning approach to classify melanocytic lesions into malignant and benign from dermoscopic images by modeling visual assessments of dermatologist. The dermoscopic image database is composed of 6940 common nevi, 528 dysplastic nevi and 532 malignant melanoma. For segmentation we have used multiphase soft segmentation with total variation and H^1 regularization. Then, each lesion is characterized by a feature vector that contains shape, color and texture information, as well as local and global parameters which reflect structures used by the dermatologist for diagnosis. The learning and classification stage is performed using SVM with polynomial kernels. The classification delivered accuracy of 98.57% with a true positive rate of 0.981 and a false positive rate of 0.019.

We did performance comparison among different models for computer-aided diagnosis system in skin lesions classification. Concerning the classification, the SVM algorithm performed better than the compared techniques. However, we have observed that the number of used feature vectors is also crucial for the classification accuracy. This highlights the importance of the feature selection. In the literature, there has not yet been a clear agreement on which feature sets are the most suitable for this task. Thus, feature selection methods not only improves the classification complexity through minimizing the utilized number of features but also the classification accuracy.



(a) (b)

Figure 8.1.: Image:(a) Skin lesion image covered by dark thick hairs (b) A skin image covered by light colored hairs

8.2. Future Work

The feasibility study of optical spectroscopy to be integrated into the clinical workflow for diagnosis lesions, the challenges addressed in section 5.7 needs to be investigated further.

In some cases, the lesion in the dermatoscopic images is covered by hairs, as shown in Figure 8.1(a). These hairs, especially the dark thick ones with a similar color hue to the lesion, occlude the lesion and may mislead the segmentation process. Shaving the hairs before imaging sessions is one of the solutions. However, shaving not only adds extra costs and time to the imaging session, but also is uncomfortable and impractical especially for multiple lesions or total-body nevus imaging. Hence, a software approach for dark thick hair removal from skin images is needed. Dull-Razor-algorithm by Lee et al [88] and inpainting technique by Wighton et al [156] provide suitable approaches to address this issue. However, light-colored hairs such as the one shown in Figure 8.1(b) interfere with the segmentation and the image analyses. In further extension of the system, these methods need to be explored and extended to account for light-colored hairs.

In conclusion, the system presented in this thesis has achieved a classification accuracy similar to the one of the dermatology specialists. Also including the above discussed extensions for future work, our system has great potential to assist the dermatologists in the diagnosis of skin lesions in the clinical routine.

CHAPTER 9

Glossary and Acronyms

Glossary

Breslow's depth. In medicine, Breslow's depth is used as a prognostic factor in melanoma of the skin. It is a description of how deeply tumor cells have invaded.

Crust. Crusting is the result of the drying of plasma or exudate on the skin.

Cutaneous T-Cell Lymphoma. A cancer in immune system. The malignant T cells in the body initially migrate to the skin, causing various lesions to appear.

Erythema. Erythema is redness of the skin, caused by hyperemia of the capillaries in the lower layers of the skin. It occurs with any skin injury, infection, or inflammation.

In vivo. In vivo is experimentation using a whole, living organism as opposed to a partial or dead organism, or an in vitro (within the glass, i.e., in a test tube or petri dish) controlled environment.

Nevus. Nevus (plural nevi) is the medical term for chronic lesions of the skin. These lesions are commonly named birthmarks and moles. Nevi are benign by

definition. Using the term nevus and nevi loosely, most physicians and dermatologists are actually referring to a variant of nevus called the M̈elanocytic nevus, which are composed of melanocytes.

Noninvasive. Noninvasive procedures do not involve tools that break the skin or physically enter the body.

Nosocomial Infections. A nosocomial infection, also known as a hospital acquired infection or HAI, is an infection whose development is favoured by a hospital environment, such as one acquired by a patient during a hospital visit or one developing among hospital staff.

Relief features. Relief features are those features which are related to landscape of those areas, e.g. mountains, altitude, valleys, types of land and heights of mountains they are the opposite of drainage pattern as it includes water channels while relief does not.

Scale. The outermost layer of skin resembling a fish scale. They represent a heaping up of the outermost layer of the skin and can be due to a variety of skin conditions, most frequently excessive dryness.

Seborrheic keratosis. It is a noncancerous benign skin growth that originates in keratinocytes. Like liver spots, seborrheic keratoses are seen more often as people age.

Telemedicine. Telemedicine is the ability to provide interactive health-care utilizing modern technology and telecommunications.

Ulcer. An ulcer is a sore on the skin or a mucous membrane, accompanied by the disintegration of tissue.

Acronyms

1D one-dimensional

2D two-dimensional

3D three-dimensional

um Micrometer

ABCD Asymmetry, Border irregularity, Color variation and Diameter

ANN Artificial Neural Networks

AR Augmented Reality

ART Advanced Realtime Tracking

AUC Area Under Curve

AVG Average

CAD Computer Aided Diagnosis

CCD Charge Coupled Device

CDL Clinically Doubtful Lesions

CFS Correlation-based Feature Selection

CGI Computer-Generated Image

cm centimeter

CT Computed tomography

CTCL Cutaneous T-Cell Lymphoma

deg Degree

FDG Fluorodeoxyglucose

ELM Epiluminence Microscopy

GB Gigabyte

GHz Gigahertz

HOG Histogram of Oriented Gradients

Hz Hertz

IR Infrared

Isomap Isometric Feature Mapping

kNN k-nearest neighbors

MDS Multidimensional Scaling

mm millimeter

MRI Magnetic Resonance Imaging

ms millisecond

NIR Near-infrared

nm nanometer

OpenGL Open Graphics Library

PC Personal Computer

PCA Principal Components Analysis

PET Positron emission tomography

PSLs Pigmented Skin Lesions

px pixel

RBF Radial Basis Functions

RGB Red Green Blue

ROC Receiver operating characteristic

SCF Skin Cancer Foundation

SD Standard Deviation

SIA Spectrophotometric Intracutaneous Analysis

SVM Support Vector Machine

UV Ultraviolet

WHO World Health Organization

Bibliography

[1] M. Amico, M. Ferri, and I. Stanganelli. Qualitative asymmetry measure for melanoma detection. In *in Proc. IEEE Int. Symp. Biomed. Imag. Nano Macro*, pages 1155–1158, 4 2004.

[2] M. Anantha, R. H. Moss, and W. V. Stoecker. Detection of pigmented network in dermatoscopy images using texture analysis. *Comput. Med. Imag. Graph.*, 28:225–234, 2004.

[3] R Rox Anderson and John A Parrish. The optics of human skin. *Journal of Investigative Dermatology*, 77(1):13–19, 1981.

[4] G. Argenziano, G. Fabbrocini, P. Carli, V. De Giorgi, E. Sammarco, and M. Delfino. Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions. comparison of the abcd rule of dermatoscopy and a new 7-point checklist based on pattern analysis. *Arch. Dermatol.*, 134(12):1563–1570, 1998.

[5] Bubacarr Bah. Diffusion maps: Analysis and applications. dissertation 66, Mathematical Modeling and Scientific Computing, Wolfson College, University of Oxford, UK, September 2008.

[6] M. Belkin and P. Niyogi. Laplacian eigenmaps and spectral techniques for embedding and clustering. in: Advances in neural information processing systems. *MIT Press, Cambridge*, 14(1):585–591, 2002.

[7] Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Comput*, 15(6):1373–1396, 2003.

[8] P. A. Berg, I. Nicander, J. Hansson, P. Geladi, U. Holmgren, and S. Ollmar. Skin cancer identification using multifrequency electrical impedance-a potential screening tool. In *IEEE Trans. Biomed. Eng.*, pages 2097–2102, 12 2004.

[9] G. Betta, G. Di Leo, G. Fabbrocini, A. Paolillo, and M. Scalvenzi. Automated application of the 7 point checklist diagnosis method for skin lesions: Estimation of chromatic and shape parameters. In *in Proc. Instrum. Meas. Technol. Conf. (IMTC) vol. 3*, pages 1818–1822, 5 2005.

[10] M. Binder, M. Schwarz, A. Winkler, A. Steiner, A. Kaider, K. Wolff K, and H. Pehamberger. Epiluminescence microscopy. a useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists. *Arch Dermatol*, 131:286–291, 1995.

[11] C. Bishop. *Neural networks for pattern recognition*. London: Oxford Univ. Press, 1995.

[12] A. Blum, H. Luedtke, U. Ellwanger, R. Schwabe, G. Rassner, and C. Garbe. Digital image analysis for diagnosis of cutaneous melanoma. development of a highly effective computer algorithm based on analysis of 837 melanocytic lesions. *Br. J. Dermatol.*, 151(5):1029–1038, 2004.

[13] J. Boldrick, C. Layton, J. Ngyuen, and S. Swtter. Evaluation of digital dermoscopy in a pigmented lesion clinic: Clinician versus computer assessment of malignancy risk. *J. Amer. Acad. Dermatol.*, 56(3):417–421, 2007.

[14] A. Bono, S. Tomatis, and C. Bartoli. The invisible colors of melanoma. a telespectrophotometric diagnostic approach on pigmented skin lesions. *Eur. J. Cancer*, 34(10):727–729, 1996.

[15] A. Bono, S. Tomatis, C. Bartoli, G. Tragni, G. Radaelli, A. Maurichi, and R. Marchesini. The abcd system of melanoma detection: a spectrophotometric analysis of the asymmetry, border, color, and dimension. *Cancer*, 85:7–72, 1999.

[16] E. Bredensteiner and K. Bennett. Multicategory classification by support vector machines. *Comput Opt Appl*, 12:53–79, 1999.

[17] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone. *Classification and Regression Trees*. New York/London, U.K.: Chapman & Hall, 1993.

[18] C. Burges. A tutorial on support vector machines for pattern recognition. [Online], 2001. Available: `http://www.kernel-machines.org/`.

[19] Donald A. Burns and Emil W. Ciurczak. *Handbook of Near-infrared Analysis*. CRC press Taylor and Francis Group, 35 edition, 2008.

[20] M. Burroni, P. Sbano, G. Cevenini, M. Risulo, G. Delleva, P. Barbini, C. Miracco, M. Fimiani, L. Andreassi, and P. Rubegni. Dysplastic naevus vs. in situ melanoma: Digital dermoscopy analysis. *Br. J.Dermatol.*, 152(4):679–684, 2005.

[21] G. Capdehourat, A. Corez, A. Bazzano, and P. Muse. Pigmented skin lesions classification using dermatoscopic images. *Springer Lecture Notes in Computer Science 5856*, pages 537–544, 2009.

[22] P. Carli, V. De Giorgi, E. Crocetti, F. Mannone, D. Massi, A. Chiarugi, and B. Giannotti. Improvement of malignant/benign ratio in excised melanocytic lesions in the dermoscopy era : A retrospective study 1997-2001. *The British Journal of Dermatology*, 137:150, 2004.

[23] N. Cascinelli, M. Ferrario, T. Tonelli, and E. Leo. A possible new tool for clinical diagnosis of melanoma. *The computer. Journal of the American Academy of Dermatology*, 16:361–367, 1987.

[24] Lawrence Cayton. Algorithms for manifold learning. TR 2008–0923, University of California, Franklin St., Oakland,, June 15 2005.

[25] M.E. Celebi, H.A. Kingravi, B. Uddin, H. Iyatomi, Y.A. Aslandogan, W.V. Stoecker, and R.H. Moss. A methodological approach to the classification of dermoscopy images. *Comput. Med. Imaging Graph*, 31:362–373, 2007.

[26] Chih-Chung Chang and Chih-Jen Lin. Libsvm: A library for support vector machines. Open source, 2001. `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.

[27] N. Christianini and J. Shawe-Taylor. *An Introduction to Support Vector Machines*. Cambridge, U.K.: Cambridge Univ. Press, second edition, 2000.

[28] B. W. Chwirot, S. Chwirot, J. Redziski, and Z. Michniewicz. Detection of melanomas by digital imaging of spectrally resolved ultraviolet light-induced autofluorescence of human skin. *Eur. J. Cancer*, 34(10):1730–1734, 1998.

[29] E. Claridge, S. Cotton, P. Hall, and M. Moncrieff. From colour to tissue histology: Physics based interpretation of images of pigmented skin lesions. *Medical Image Analysis*, 7:489–502, 2003.

[30] D. Clark. Computational methods for probabilistic decision trees. *Comput. Biomed. Res.*, 30:19–33, 1997.

[31] Mayo Clinic. Webpage:. `http://www.mayoclinic.com/health/medical/IM02400`.

[32] Mayo Clinic. Skin cancer. Information on web link, 2010. `http://www.mayoclinic.com/health-information/`.

[33] R.R. Coifman and S. Lafo. Diffusion maps: Applied and computational harmonic analysis. *Special Issue: Diffusion Maps and Wavelets*, 21(1):5–30, 2006.

[34] O. Colot, R. Devinoy, A. Sombo, and D. de Brucq. A color image processing method for melanoma detection. *Medical Image Computing and Computer-Assisted Interventation MICCAI*, 1496:562, 1998.

[35] C. Cortes and M. Mohri. Confidence intervals for the area under the roc curve. *MIT Press*, 17:305–312, 2005.

[36] T. F. Cox and M. A. A. Cox. *Multidimensional scaling*. Chapman & Hall, London, 2 edition, 1994.

[37] Nello Cristianini and John Shawe Taylor. *An Introduction to Support Vector Machines and Other Kernel based Learning Methods*. Cambridge University Press, first edition, 2000.

[38] R. K. Curley, M. G. Cook, M. E. Fallowfield, and R. A. Marsden. Accuracy in clinically evaluating pigmented lesions. *Br. Med. J.*, 299:16–18, 1989.

[39] I. K. Daftari, E. Aghaian, J.M. Brien, W.Dillon, and T. L. Phillips. 3d mri-based tumor delineation of ocular melanoma and its comparison with conventional techniques. *Med. Phys.*, 32(11):3355–3362, 2005.

[40] J. B. Dawson, D. J. Barker, D. J. Ellis, E. Grassam, J. A. Cotterill, G. W. Fisher, and J. W. Feather. A theoretical and experimental study of light absorption and scattering by in vivo skin. *Phys. Med. Biol.*, 25:695–709, 1980.

[41] J. de la Portey, B. M. Herbsty, W. Hereman, and S. J. van der Walty. An introduction to diffusion maps. In *In Proceedings of the nineteenth annual symposium of the Pattern Recognition Association of South Africa (PRASA)*, August 2008.

[42] C. Decaestecker, I. Salmon, O. Dewitte, I. Camby, P.V. Ham, J. Paseels, J. Brotchi, and R. Kiss. Nearest-neighbor classification for identification of aggressive versus non-aggressive low-grade astrocytic tumors by means of image cytometry-generated variables. *J Neurosurg*, 86:532–537, 1997.

[43] Skin Different Depths. Webpage:. `http://www.missionignition.net/bms/led_heal_clip_image008.jpg`.

[44] B. L. Diffey, R. J. Oliver, and P.M. Farr. A portable instrument for quantifying erythema induced by ultraviolet radiation. *Br. J. Dermatol.*, 111:663–72, 1984.

[45] S. Dreiseitl, L. Ohno Machado, H. Kittler, S. Vinterbo, H. Billhardt, and M. Binder. A comparison of machine learning methods for the diagnosis of pigmented skin lesions. *Journal of Biomedical Informatics*, 34:28–36, 2001.

[46] A. Durg, W. V. Stoecker, J. P. Vookson, S. E. Umbaugh, and R. H.Moss. Identification of variegated coloring in skin tumors. *IEEE Eng. Med. Biol. Mag.*, 12:71–75, 1993.

[47] B. Dasarathy editor. *Nearest neighbor (NN) norms: NN pattern classification techniques*. IEEE Computer Society Press, Los Alamitos, CA, 1991.

[48] D. Altman editor. *Practical statistics for medical research*. London/New York: Chapman & Hall, 1991.

[49] Edward A. Edwards and S. Quimby Duntley. The pigments and color of living human skin. *American Journal of Anatomy*, 65(3):1–33, 2005.

[50] E. el Kwae, J. Fishman, M. Bianchi, P. Pattany, and M. Kabuka. Detection of suspected malignant patterns in three-dimensional magnetic resonance breast images. *J Digit Imaging*, 11:83–93, 1998.

[51] M. Elbaum, A. W. Kopf, H. S. Rabinovitz, R. Langley, H. Kamino, M. Mihm, A. Sober, G. Peck, A. Bogdan, D. Krusin, M. Greenebaum, S. Keem, M. Oliviero, and S. Wang. Automatic differentiation of melanoma from melanocytic nevi with multispectral digital dermoscopy: A feasibility study. *J. Amer. Acad. Dermatol*, 44(2):207–218, 2001.

[52] F. Ercal, A. Chawla, W. V. Stoecker, H.C. Lee, and R. H. Moss. Neural network diagnosis of malignant melanoma from color images. *IEEE Trans. Biomed. Eng.*, 14(9):837–845, 1994.

[53] T. J. Farrell, M. S. Patterson, and B. Wilson. A diffusion theory model of spatially resolved, steady state diffuse reflectance for the non–invasive determination of tissue optical properties in vivo. *Med. Phys.*, 19:879–88, 1992.

[54] T. Fawcett. An introduction to roc analysis. *Pattern Recognition Letters*, 27:861–874, 2006.

[55] J. W. Feather, M. Hajizadeh-Saffar, G. Leslie, and J. B. Dawson. A portable scanning reflectance spectrophotometer using visible wavelengths for the rapid measurement of skin pigments. *Phys. Med. Biol.*, 34(1):807–20, 1989.

[56] Joshua B Fishkin, Olivier Coquoz, Eric R. Anderson, Matthew Brenner, and Bruce J. Tromberg. Frequency-domain photon migration measurements of normal and malignant tissue optical properties in a human subject. *Applied Optics*, 36(1):10–20, 1997.

[57] R. J. Friedman, D. S. Rigel, and A. W. Kopf. Early detection of malignant melanoma: the role of physicians examination and self examination of the skin. *CA Cancer Journal for Clinicians*, 35:130–151, 1985.

[58] H. Ganster, A. Pinz, R. Rhrer, E. Wildling, M. Binder, and H. Kittler. Automated melanoma recognition. *IEEE Transactions on Medical Imaging*, 20:233–239, 2001.

[59] H. Ganster, P. Pinz, R. Rohrer, E. Wildling, M. Binder, and H. Kittler. Automated melanoma recognition. *IEEE Trans. Med. Imag.*, 20(3):233–239, 2001.

[60] C. Grana, G. Pellacani, R. Cucchiara, and S. Seidenari. A new algorithm for border description of polarized light surface microscopic images of pigmented skin lesions. *IEEE Trans.Med. Imag.*, 22(8):959–964, 2003.

[61] C. Grin, A. Kopf, B. Welkovich, R. Bart, and M. Levenstein. Accuracy in the clinical diagnosis of malignant melanoma. *Arch Dermatol*, 126:763–766, 1990.

[62] P. Grzymala-Busse, J.W. Grzymala-Busse, and Z.S. Hippe. Melanoma prediction using data mining system lers. *25th Annual International Computer Software and Applications Conference*, pages 615–620, 2001.

[63] A. Gutenev, V. N. Skladnev, and D. Varvel. Acquisition-time image quality control in digital dermatoscopy of skin lesions. *Comput. Med. Imag. Graph.*, 25(1):495–499, 2001.

[64] G. Argenziano H, H. P. Soyer, S. Chimenti, R. Talamini, R. Corona, F. Sera, M. Binder, L. Cerroni, G. De Rosa, G. Ferrara, R. Hofmann-Wellenhof, M. Landthaler, S.W.Menzies, H. Pehamberger, D. Piccolo, H. S. Rabinovitz, R. Schiffner, S. Staibano, W. Stolz, I. Bartenjev, A. Blum, R. Braun, H. Cabo, P. Carli, V. De Giorgi, M. G. Fleming, J. M. Grichnik, C. M. Grin, A. C. Halpern, R. Johr, B. Katz, R. O. Kenet, H. Kittler, J. Kreusch, J. Malvehy, G. Mazzocchetti, M. Oliviero, F. Ozdemir, K. Peris, R. Perotti, A. Perusquia, M. A. Pizzichetta, S. Puig, B. Rao, P. Rubegni, T. Saida, M. Scalvenzi, S. Seidenari, I. Stanganelli, M. Tanaka, K. Westerhoff, I. H. Wolf, O. Braun-Falco, H. Kerl, T. Nishikawa, K. Wolff, and A.W. Kopf. Dermoscopy of pigmented skin lesions: Results of a consensus meeting via the internet. *J. Amer. Acad. Dermatol.*, 48(5):680–693, 2003.

[65] D. Hagen. Test characteristics: How good is that test? *Primary Care*, 22:213–233, 1995.

[66] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. The weka data mining software: An update; sigkdd explorations. Issue 1, 2009. Volume 11, Issue 1.

[67] M.A. Hall. Correlation-based feature selection for discrete and numeric class machine learning. In *Proceedings of the Seventeenth International Conference on Machine Learning*, pages 359–366, 2000.

[68] P.N. Hall, E. Claridge, and J.D. Smith. Computer screening for early detection of melanoma: Is there a future? *British J. Dermatology*, 132:325–328, 1995.

[69] D. J. Hand and R. J. Till. A simple generalization of the area under the roc curve to multiple class classification problems. *Machine Learning*, 45:171–186, 2001.

[70] H. Handels, T. Ross, J. Kreusch, H. Wolff, and S. Poppl. Feature selection algorithm for optimized skin tumor recognition using genetic algorithms. *Artif Intell Med*, 16:283–297, 1999.

[71] J. Hanley and B. McNeil. The meaning and use of the area under the receiver operating characteristic (roc) curve. *Radiology*, 143:29–36, 1982.

[72] J. Hanley and B. McNeil. Method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology*, 148:839–843, 1982.

[73] G. Hansen, E. Sparrow, J. Kokate, K. Leland, and P. Iaizzo. Wound status evaluation using color image processing. *IEEE Trans. Med.Imag.*, 16:78–86, 1997.

[74] T. Hastie, R. Tibshirani, and J. H. Friedman. *The elements of statistical learning: Data Mining, Inference and Prediction*. Springer, 2001.

[75] M. F. Healsmith, J. F. Bourke, J. E. Osborne, , and R. A. C. Graham-Brown. An evaluation of the revised seven-point checklist for the early diagnosis of cutaneous malignant melanoma. *Br. J. Dermatol.*, 130:48–50, 1994.

[76] M. Herbin, F. Bon, A. Venot, F. Jeanlouis, M. Dubertret, L. Dubertret, and G. Strauch. Assessment of healing kinetics through true color image processing. *IEEE Trans. Med. Imag.*, 12(1):39–43, 1993.

[77] K. Hoffmann, T. Gambichler, A. Rick, M. Kreutz, M. Anschuetz, T. Grunendick, A. Orlikov, S. Gehlen, R. Perotti, L. Andreassi, J. Newton Bishop, J. P. Cesarini, T. Fischer, P. J. Frosch, R. Lindskov, R. Mackie, D. Nashan, A. Sommer, M. Neumann, J. P. Ortonne, P. Bahadoran, P. F. Penas, U. Zoras, and P. Altmeyer. Diagnostic and neural analysis of skin cancer (danaos). a multicentre study for collection and computer-aided analysis of data from pigmented skin lesions using digital dermoscopy. *Br. J. Dermatol.*, 149(10):801–809, 2003.

[78] Alexander Horsch. *Chapter title: Melanoma Diagnosis*. Springer: Biomedical Image Processing, Biomedical Engineering, ISSN 1618–7210, 2011.

[79] T. L. Hwang, W. R. Lee, S. C. Hua, and J.Y. Fang. Cisplatin encapsulated in phosphatidylethanolamine liposome enhances the in vitro cytotoxicity and in vivo intratumor drug accumulation against melanomas. *J. Dermatol. Sci.*, 46(1):11–20, 2007.

[80] A. K. Jain and W. G. Waller. On the optimal number of features in the classification of multivariate gaussian data. *Pattern Recognit*, 10:365–374, 1978.

[81] B. Jalil and F. Marzani. *Multispectral Image Processing Applied to Dermatology*. Le2i laboratory Universite de Bourgogne, 2008.

[82] R. H. Johr. Dermoscopy: alternative melanocytic algorithms - the abcd rule of dermatoscopy, menzies scoring method, and 7–point checklist. *Clinics in Dermatology*, 20:240–7, 2002.

[83] C.D. Kalemke, S. Goerdt, D. Schrama, and J.C. Becker. New insights into the molecular biology and targeted therapy of cutaneous t-cell lymphomas. *Journal of the Germany Society of Dermatology*, 4:395–405, 2006.

[84] A. Kienle, L. Lilge, M. S. Patterson, R. Hibst, R. Steiner, and B. C. Wilson. Spatially resolved absolute diffuse reflectance measurements for noninvasive determination of the optical scattering and absorption coefficients of biological tissue. *Appl. Opt.*, 35:2304–14, 1996.

[85] A. Kim, U. Kasthuri, B.C. Wilson, A. White, and A. L. Martel. Preliminary clinical results for the in vivo detection of breast cancer using interstitial diffuse optical spectroscopy. *MICCAI Workshop on Biophotonics Imaging for Diagnostics and Treatment*, 75:1601–2321, 2006.

[86] M. Kudo and J. Sklansky. Comparison of algorithms that select features for pattern classifiers. *Pattern Recognit.*, 33:25–41, 2000.

[87] H. E. Kuppenheim and R aymond Heer. Spectral reflectance of white and negro skin between 440 and $1000 m\mu$. *Journal of Applied Physiology*, 4(10):800–806, 1952.

[88] T. Lee, V. Ng, R. Gallagher, A. Coldman, and D. McLean. Dullrazor: A software approach to hair removal from images. *Computers in Biology and Medicine*, 27:533–543, 1997.

[89] T. K. Lee, M. S. Atkins, M.A.King, S. Lau, and D. I.McLean. Counting moles automatically from back images. *IEEE Trans. Biomed. Eng.*, 52(11):1966–1969, 2006.

[90] Fang Li, Chaomin Shen, and Chunming Li. Multiphase soft segmentation with total variation and h1 regularization. *Journal of Mathematical Imaging and Vision*, 37:98–111, 2010.

[91] B. Lindelof and M. Hedblad. Accuracy in the clinical diagnosis and pattern of malignant melanoma at a dermatological clinic. *Journal of Dermatology*, 21(7):461–464, 1994.

[92] D. Liu and S. Li. Accurate calibration of a stereo-vision system in image-guided radiotherapy. *Med Phys*, 33:4379–4383, 2006.

[93] M. Loane, H. Gore ans R. Corbet, and K. Steele. Effect of camera performance on diagnostic accuracy. *J. Telemed. Telecare*, 3(1):83–88, 1997.

[94] W. Lohman and E. Paul. In situ detection of melanomas by fluorescence measurements. *Naturewissenschaften*, 75:201–202, 1988.

[95] H. Lorentzen, K. Weismann, R. Kenet, L. Secher, and F. Larsen. Comparison of dermatoscopic abcd rule and risk stratification in the diagnosis of malignant melanoma. *Acta Derm Venereol*, 80:122–126, 2000.

[96] H. Lorentzen, K. Weismann, C.S. Petersen, F.G. Larsen, L. Secher, and V. Skodt. Clinical and dermatoscopic diagnosis of malignant melanoma. assessed by expert and non-expert groups. *Acta dermato–venereologica*, 79(4):301–304, 1999.

[97] C. J. Lynn, I. S. Saidi, D. G. Oelberg, and S. L. Jacques. Gestational age correlates with skin reflectance in newborn infants of $24 - -42$ weeks gestation. *Biol. Neonate*, 75:9–64, 1993.

[98] M. Hintz Madsen, L. Hansen, J. Larsen, and K. Drzewiecki. A probabilistic neural network framework for detection of malignant melanoma, in arti-

ficial neural networks in cancer diagnosis prognosis and patient management. In *18th IEEE Symposium on Computer-Based Medical Systems*, pages 141–183, 6 2001.

[99] I Maglogiannis and Charalampos N. Doukas. Overview of advanced computer vision systems for skin lesions characterization. *IEEE Transactions On Information Technology In Biomedicine*, 13(5):721–733, 2009.

[100] I. Maglogiannis and D. Kosmopoulos. A system for the acquisition of reproducible digital skin lesion images. *Technol. Healthcare*, 11(3):425–441, 2003.

[101] I. Maglogiannis, S. Pavlopoulos, and D. Koutsouris. An integrated computer supported acquisition, handling and characterization system for pigmented skin lesions in dermatological images. *IEEE Trans. Inf. Technol. Biomed.*, 9:86–98, 2005.

[102] R. B. Mallet, M. E. Fallowfield, M. G. Cook, W. N. Landells, C. A. Holden, and R. A. Marsden. Are pigmented lesion clinics worthwhile? *Br. J. Dermatol.*, 129:689–693, 1993.

[103] A. G. Manousaki, A. G. Manios, E. I. Tsompanaki, and A. D. Tosca. Use of color texture in determining the nature of melanocytic skin lesions: A qualitative and quantitative approach. *Comput. Biol. Med.*, 36:419–427, 2006.

[104] R. Marchesini, M. Brambilla, C. Clemente, M. Maniezzo, A. E. Sichirollo, A. Testori, D. R. Venturoli, and N. Cascinelli. In vivo spectrophotometric evaluation of neoplastic and non-neoplastic skin pigmented lesions–i. reflectance measurements. *Photochem Photobiol*, 53:77–84, 1991.

[105] R. Marchesini, N. Cascinelli, M. Brambilla, C. Clemente L. Mascheroni E. Pignoli, A. Testori, and D. R. Venturoli. In vivo spectrophotometric evaluation of neoplastic and non-neoplastic skin pigmented lesions–2. discriminant analysis between nevus and melanoma. *Photochem Photobiol*, 55:515–22, 1992.

[106] A.A. Marghoob, L.D. Swindle, C.Z. Moricz, F.A. Sanchez Negron, B. Slue, A.C. Halpern, and A.W. Kopf. Instruments and new technologies for the in vivo diagnosis of melanoma. *J Am Acad Dermatol.*, 49(5):777–97, 2003.

[107] R. Marks. Epidemiology of melanoma, clinical excremental dermatology. *PMID: 11044179 (PubMed - indexed for MEDLINE)*, 25:395–406, 2000.

[108] L.M. McIntosh, R. Summers, M. Jackson, H.H. Mantsch, J.R. Mansfield, M. Howlett A.N. Crowson, and J.W. Toole. Towards non-invasive screening of skin lesions by near-infrared spectroscopy. *Journal Invest Dermatol*, 116:175–181, 2001.

[109] D. McKenzie, P. McGorry, C. Wallace, L. Low ans D. Copolov, and B. Singh. Constructing a minimal diagnostic decision tree. *Methods. Inf. Med.*, 32:161–166, 1993.

[110] MedicineNet. melanoma. Public Awareness through web, 2011. `http://www.medicinenet.com/script/main/hp.asp`.

[111] S.W. Menzies, L.Bischof, H. Talbot, A.Gutener, M.Avramidis, and L.Wong et al. The performance of solarscan: An automated dermoscopy image analysis instrument for the diagnosis of primary melanoma. *Arch. Dermatol.*, 141:1388–1396, 2009.

[112] M. Moncrieff, S. Cotton, E. Claridge, and P. Hall. Spectrophotometric intracutaneous analysis: a new technique for imaging pigmented skin lesions. *Br J Dermatol.*, 146(3):448–57, 2002.

[113] C.A. Morton and R.M. Mackie. Clinical accuracy of the diagnosis of cutaneous malignant melanoma. *Journal of Dermatology*, 7:138–283, 1998.

[114] H. Motoyama, T. Tanaka, M. Tanaka, and H. Oka. Feature of malignant melanoma based on color information. In *in Proc. SICE Annu. Conf., Sapporo, Japan vol. 1*, pages 230–233, 8 2004.

[115] B.W. Murphy, R.J. Webster, B.A. Turlach, C.J. Quirk, C.D. Clay, P.J. Heenan, and D.D. Sampson. Toward the discrimination of early melanoma from common and dysplastic nevus using fiber optic diffuse reflectance spectroscopy. *Journal Biomed Opt*, 10:064020, 2005.

[116] F. Nachbar, W. Stolz, T. Merkle, A. Cognetta, T Vogt, M. Landthaler, P. Bilek, O. Braun-Falco, and G. Plewig. The abcd rule of dermatoscopy: high prospective value in the diagnosis of doubtful melanocytic skin lesions. *Journal of the American Academy of Dermatology*, 30:551–559, 1994.

[117] Product name: MelaFind(R). Mela sciences, inc. Address: 50 South Buckhout Street, Suite 1, Irvington, NY 10533, 2004. http://www.melasciences.com/.

[118] Product name: SpectroShade(R). Mht s.p.a. headquarter. Production and Commercialization Via Milano 12, 37020 Arbizzano di Negrar (VR) Italy, 2007. www.mht.it.

[119] Kelly C. Nelson. Mole mapping. Website. http://dermatology.duke.edu/clinical-services/mole-mapping.

[120] J. A. Nimunkar, P. A. Dhawan, P. A. Relue, , and S. V. Patwardhan. Wavelet and statistical analysis for melanoma classification. *Progr. Biomed. Opt. Imag.*, 3:1346–1353, 2003.

[121] KC Nischal and Uday Khopkar. Dermoscope. Indian J Dermatol Venereol Leprol 300–303, Department of Dermatology, Seth GS Medical College & KEM Hospital, Parel, Mumbai–400 012, India, 2005.

[122] SM. Nischic and C. Forster. Analysis of skin erythema using true color images. *IEEE Trans. Med. Imag.*, 16(6):711–716, 1997.

[123] L. Noriega. *Multilayer Perceptron Tutorial*. school of Computing Staffordshire University, 2005.

[124] World Health Organization. Ultraviolet radiation and the intersun programme. Website, 2007. http://www.who.int/uv/faq/skincancer/en/.

[125] M. Osawa and S. Niwa. A portable diffuse reflectance spectrophotometer for rapid and automatic measurement of tissue. *Meas. Sci. Technol.*, 4:668–76, 1993.

[126] R.J. Pariser and D.M. Pariser. Primary care physicians errors in handling cutaneous disorders. *Journal of the American Academy of Dermatology*, 17:239–245, 1987.

[127] S. V. Patwardhan, S. Dai, and A. P. Dhawan. Multi-spectral image analysis and classification of melanoma using fuzzy membership based partitions. *Comput. Med. Imag. Graph.*, 29(1):287–296, 2005.

[128] S. V. Patwardhan, A. P. Dhawan, and P. A. Relue. Classification of melanoma using tree structured wavelet transforms. *Comput. Methods Programs Biomed.*, 72(8):223–239, 2003.

[129] C. Pleiss, J. H. Risse, H. J. Biersack, and H. Bender. Role of fdg pet in the assessment of survival prognosis in melanoma. *Cancer Biother. Radiopharm.*, 22(6):740–747, 2007.

[130] B. Rosado, S. Menzies, A. Harbauer, H. Pehamberger, K. Wolff, M. Binder, and H. Kittler. Accuracy of computer diagnosis of melanoma: A quantitative meta-analysis. *Archives of Dermatology*, 139(3):361–367, 2003.

[131] P. Rubegni, G. Cevenini, M. Burroni, R. Perotti, G. Dell'eva, P. Sbano, C. Miracco, P. Luzi, P. Tosi, P. Barbini, and L. Andreassi. Automated diagnosis of pigmented skin lesions. *Publication of the International Union Against Cancer*, 101:576–580, 2002.

[132] J. Sanders, B. Goldstein, D. Leotta, and K. Richards. Image processing techniques for quantitative analysis of skin structures. *Comput. Methods Programs Biomed.*, 59(4):167–180, 1999.

[133] A. Sboner, C. Eccher, E. Blanzieri, P. Bauer, M. Cristofolini, G. Zumiani, and S. Forti. A multiple classifier system for early melanoma diagnosis. *Artif. Intell. Med.*, 27:29–44, 2003.

[134] Bernhard Schölkopf and Alexander J. Smola. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. The MIT Press, first edition, 2001.

[135] P. Schmid-Saugeon, J. Guillod, and J.P Thiran. Towards a computer–aided diagnosis system for pigmented skin lesions. *Comp. Med. Imag. Graphics*, pages 65–78, 2003.

[136] M. Schneider and C. Stevens. Development and testing of a new magnetic-tracking device for image guidance. In *Proceedings of SPIE Medical Imaging*, volume 7035, pages 65090I–65090I–11, 2007.

[137] S. Sigurdsson, P. A. Philipsen, L. K. Hansen, J. Larsen, M. Gniadecka, and H. C. Wulf. Detection of skin cancer by classification of raman spectra. In *IEEE Trans. Biomed. Eng.*, pages 1784–1793, 10 2004.

[138] M.R. Sikonja and I. Kononenko. Theoretical and empirical analysis of relieff and rrelieff. *Mach. Learn. Kluwer Academic Publishers*, 53:23–69, 2003.

[139] S. Singh, J. H. Stevenson, and D. McGurty. An evaluation of polaroid photographic imaging for cutaneous-lesion referrals to an outpatient clinic: A pilot study. *Br. J. Plastic Surg.*, 54(1):140–143, 2001.

[140] Richard Socher and Matthias Hein. *Manifold Learning and Dimensionality Reduction with Diffusion Maps*. Tech. report, 2008.

[141] J. Solomon, S. Mavinkurve, D. Cox, and R. M. Summers. Computer assisted detection of subcutaneous melanomas. *Acad. Radiol.*, 11(6):678–85, 2004.

[142] H. P. Soyer, J. Smolle, H. Kerl, and H. Stettnre. Early diagnosis of malignant melanoma by surface microscopy. *Lancet*, 2(10):803, 1987.

[143] D. Spiegelhalter. Probabilistic prediction in patient management and clinical trials. *Stat. Med.*, 5:421–433, 1986.

[144] R. J. Stanley, R. H. Moss, W. Van Stoecker, and C. Aggarwal. A fuzzy based histogram analysis technique for skin lesion discrimination in dermatology clinical images. *Comput. Med. Imag. Graph.*, 27:387–396, 2003.

[145] W. V. Stoecker, K. Gupta, R. J. Stanley, R. H. Moss, and B. Shrestha. Detection of asymmetric blotches in dermoscopy images of malignant melanoma using relative color. *Skin Res. Technol.*, 11(3):179–184, 2005.

[146] G. Surowka and K. Grzesiak-Kopec. Different learning paradigms for the classification of melanoid skin lesions using wavelets. In *in Proc. 29th Annu. Int. Conf. IEEE EMBS*, pages 3136–3139, 8 2007.

[147] J. A. Swets. Measuring the accuracy of diagnostic systems. *Science*, 4857:1285–93, 1988.

[148] J. B. Tenenbaum, V. de Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000.

[149] S. Tomatis, C. Bartol, G. Tragni, B. Farina, and R. Marchesini. Image analysis in the rgb and hs colour planes for a computer assisted diagnosis of cutaneous pigmented lesions. *Tumori*, 84(5):29–32, 1998.

[150] TransLite. 8410 highway 90a, suite 150, sugar land. Website. `http://www.tlite.com/illum.html`.

[151] L. Tsap, D. Goldgof, S. Sarkar, and P. Powers. Vision based technique for objective assessment of burn scars. *IEEE Trans.Med. Imag.*, 17(4):620–633, 1998.

[152] S. Umbaugh, Y. Wei, and M. Zuke. Feature extraction in image analysis. *IEEE Eng. Med. Biol.*, 16(4):62–73, 1997.

[153] S. E. Umbaugh, R. H. Moss, and W. V. Stoecker. Applying artificial intelligence to the identification of variegated coloring in skin tumors. *IEEE Eng. Med. Biol. Mag.*, 10(4):57–62, 1991.

[154] V. Vapnik. *Statistical learning theory*. New York Wiley, 1998.

[155] Ulrike von Luxburg. A tutorial on spectral clustering. *Stat Comput, Springer Science and Business Media*, 17(5):395–416, 2007.

[156] P. Wightona, Tim K. Leea, and M. Stella Atkinsa. Dermascopic hair disocclusion using inpainting. *SPIE Medical Imaging*, 6914:691427–1, 2008.

[157] Klaus-Peter Wilhelm, Peter Elsner, Enzo Berardesca, and Howard I. Maibach. *Bioengineering of the Skin: Skin Imaging and Analysis, (Dermatology: Clinical and Basic Science)*. Informa Healthcare, second edition, 2006.

[158] M. Wiltgen, A. Gerger, and J. Smolle. Tissue counter analysis of benign common nevi and malignant melanoma. *Int. J. Med. Inf.*, 69:17–28, 2003.

[159] X. Yuan, Z. Yang, G. Zouridakis, and N. Mullani. Svm based texture classification and application to early melanoma detection. *28th IEEE EMBS Annu. Int. Conf. Proc.*, 3(8):4775–4778, 2006.

[160] Matteo Zanotto. *Visual Description of Skin Lesions*. School of Informatics University of Edinburgh, 2010.

[161] Z. Zhang, R. H. Moss, and W. V. Stoecker. Neural networks skin tumor diagnostic system. In *in Proc. IEEE Int. Conf. Neural Netw. Signal Process.*, pages 191–192, 12 2003.

[162] M. Zorman, M. Stiglic, P. Kokol, and I. Malcic. The limitations of decision trees and automatic learning in real world medical decision making. *J. Med. Syst.*, 21:403–415, 1997.

[163] G. Zouridakis, M. Doshi, and N. Mullani. Early diagnosis of skin cancer based on segmentation and measurement of vascularization and pigmentation in nevoscope images. In *in Proc. 26th Annu. Int. Conf. IEEE EMBS,San Francisco, CA*, pages 1593–1596, 9 2004.