Classification of melanoma lesions using non-conventional screening techniques

Mojdeh Rastgoo Dastjerdi

Universitat de Girona

Vicorob Institute

Université de Bourgogne

LE2I Laboratory

A Thesis Submitted for the Degree of Doctor of Philosophy in computer vision and robotics

· 2015 ·

Abstract

List of Figures

List of Tables

Abbreviations

ACC Accuracy

AdB AdaBoost

ANN Artificial Neural Network

AUC area under the curve

AWB Auto-white-balance

BD Barrel Deformation

BoW Bag of Words

CAD computer-aided diagnosis

CDS clinical diagnosis support

CLBP Completed Local Binary Pattern

CLSM confocal laser scanning microscopy

CUS Clustering

CWB Custom-white-balance

DoG Difference of Gaussian

DOS Data Space Over Sampling

ECOC Error Correcting Output Coding

ENN Editted Nearest Neighbor

FCM Fuzzy-C-*means*

GB Gradient Boosting

GLCM Grey-Level Co-Occurrence Matrix

GVF gradient vector flow

HOG Histogram of Oriented Gradients

JPEG joint photographic experts group

k-CV k fold cross-validation

LBP Local Binary Pattern

LDA Linear Discriminant Analysis

LLE Locally Linear Embedding

LOOCV leave-one-out-cross-validation

MRI magnetic resonance imaging

mRMR Minimum Redundancy Maximum Relevance

MSD multispectral dermoscopy

MSI Multispectral Imaging

MVU Maximum Variance Unfolding

NB Naive Bayes

NCR Neighborhood Cleaning Rule

NM1 NearMiss-1

NM2 NearMiss-2

NM3 NearMiss-3

NN Nearest Neighbor

NPD nonpolarized dermoscope

OCT optical coherence tomography

OS Over-Sampling

PCA principal component analysis

PD polarized dermoscope

pdf probablity density function

PI polarized imaging

PPV Precision (positive predictive value)

PSLs pigmented skin lesions

RDGM Random Deformation using Gaussian Motion

RF Random Forests

ROC Recever Operating Characteristics

ROS Random Over-Sampling

RUS Random Under-Sampling

SBFS Sequential Backward Feature Selection

SCF Sparse Coded Features

SE Sensitivity

SFFS Sequential Forward Feature Selection

SIFT Scale-Invariant Feature Transform

SMOTE Synthetic Minority Over-Sampling

SP Specificity

SVD Singular Value Decomposition

SVM Support Vector Machine

TL Tomek-Link

US Under-Sampling

US ultra-sound

Contents

Acknowledgments

Chapter 1

Introduction

1.1 The Human Skin

Skin is the largest organ of human body and consists of three main layers, epidermis, dermis and subcutaneous fat [12] (see Fig.1.1).

Epidermis is the outer layer and surface of the skin. This layer is divided into four sub-layers from top to bottom: stratum corneum, stratum granulosum, stratum spinosum and stratum basal [112].

The aforementioned layers contain four types of cells including keratinocytes, melanocytes, Langerhans' cells and Merkel cells. The majority of the cells in epidermis belong to keratinocytes, which are the main force for continuous renewal of the skin [112]. These cells contain two main attributes of division and differentiation which enable them to renew the outer layer of the epidermis within thirty days. During this journey, the keratinocytes, which are produced by division of the basal cells (keratinocytes in basal layer are called basal cells), will move to the next layer while they go through a differentiation process. Differentiation refers to morphology and biochemistry transformation of the cells. At the end of this journey, the keratinocyte cells will lose their nuclie and will be transformed to flattened cells which are filled with keratine. These cells form the outer most layer of the epidermis (stratum corneum). At the end of differentiation process the corneocytes lose their cohension and separate from the surface in the desquamation process, resulting in the renewed skin.

Langerhans' cells are responsible for the detection of foreign bodies (antigens) which invade the epidermis, transporting them to the local lymph nodes, while Merkel cells act

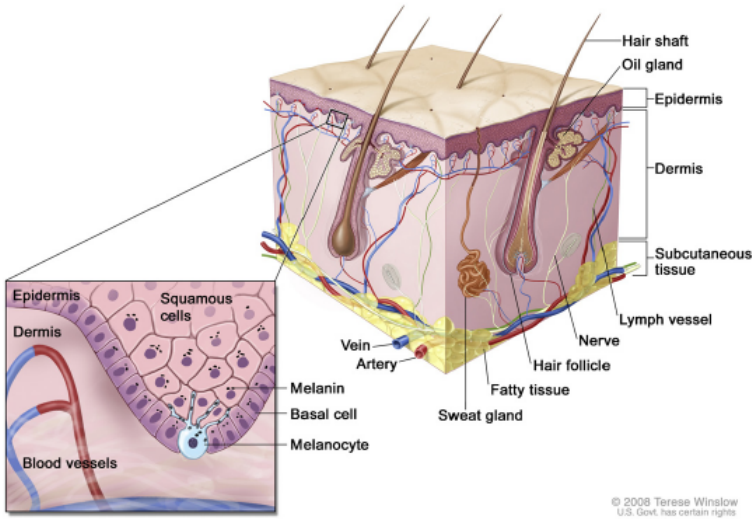


Figure 1.1: Anatomy of the skin, showing the three structure layer of epidermis, dermis and subcutaneous tissue [95]Get permission

as mechanosensory receptors in response to touch, forming close connections with sensory nerve endings [112].

Melanocyte cells are found in the basal layer of epidermis [112]. These cells are responsible for distributing packages of melanin pigment to the surrounding keratinocytes, which lead to individual skin and hair colors. This chromophore mainly protects the subcutaneous tissue from being damaged by UV radiation. Whenever the levels of UV radiation increases, melanocytes start producing more melanin, resulting in our tanning reaction to sun exposure. Melanin is the major chromophore of the epidermis which occupies the top 50-100 µm, with the exception to superior layers of epidermis.

In most light propagation models through skin, the sublayers of epidermis are considered as one layer [**?**, 91]. Epidermis thickness could vary depending on different body parts, however in average it is mentioned to be 0.1 mm. The most external sublayer of epidermis is stratum corneum. It is composed of dry dead cells without organelles and is filled with keratin fibres. The light is mainly absorbed in this layer, due to epidermis major component melanin.

Dermis is the middle layer between epidermis and the hypodermic layer. This layer is thicker than epidermis, with an average thickness of 0.6 to 3 mm [12]. The thermoregulation, mechanical resistance and nourishing of the epidermis are the main functions of this layer. The dermis is composed of elastic collagen fibres, blood vessels, nerves, lymph

vessels, hair follicles and sweat glands but its principal molecules, with relevant optical properties, are haemoglobin, carotene and bilirubin. Haemoglobin is a chromophore of red colour found in the microvascular network of the dermis, typically 100-500 µm below the skin surface. This chromophore carries oxygen through vessels and capillaries and accordingly is called oxy-haemoglobin since it contains oxygen and deoxy-haemoglobin otherwise. The dermis is divided into two sublayers, papillary dermis with the principal function of thermoregulation and reticular dermis which gives the skin its strength and elasticity.

Subcutaneous fat or hypodermic fat is the deepest layer of the skin, with an average thickness of 4 to 9 mm. This layer is composed of connective tissues, fat cells and blood vessels.

Cancer can develop from almost any cell in the body. However certain cells are more cancer prone compared to others and the skin is not an exception. The three most common malignant skin cancers are called basal cell carcinoma, squamous cell carcinoma and melanoma which develop from basal cells, squamous keratinocytes, and melanocytes, respectively. Melanocytes cells and their transformation, due to their malignant transformation potential, are our main concern in this thesis. Melanoma is less common in comparison to basal cell carcinoma and squamous cell carcinoma, it is the far more deadly and aggressive type of skin cancer. The characteristics and treatment of this cancer, as well as some skin lesions more prone to malignancy are described in the following sections.

1.2 Pigmented Skin Lesions

Pigmented skin lesions or melanocytic nevi appear on the surface of the skin [93], where the melanocytes cells grow in clusters beside the normal skin cells. This is a natural transformation of skin cells and creates benign pigmented skin lesions (PSLs) such as:

Freckle or ephelis are pale-brown macular lesions which are usually of less than 3 mm diameter with a poorly defined lateral margin [120].

Common nevi which are typical flat melanocytic nevi or moles.

Congenital nevi which are moles that appear at birth, also known as “birth marks”.

Atypical or dysplastic nevus is a common nevus with inconsistent coloration, irregular edges, blurry borders, scale-like texture and a diameter greater than 5 mm [93]. Atypical mole syndrome or dysplastic nevus syndrome, describes individuals with large quantities of dysplastic nevi. Such individuals face a higher risk of developing melanoma (6 to 10

times than the other people with few nevi) [120]. However, only a small number of these dysplastic nevi might develop into melanoma and most dysplastic nevi will never become cancer.

Blue nevus is a melanocytic nevus comprised of abnormal collections of benign pigment melanocytes which are located in the dermis rather than at the dermoepidermal junction [120]. The blue or blue-black appearance of the lesion is caused due to the light reflection of melanin from the dermis.

Pigmented Spitz nevus is an uncommon benign nevus that shares a similar physical characteristics to melanoma and is usually seen in children [120].

From the aforementioned lesions, congenital and dysplastic nevi are most likely to develop to malignant melanoma [72].

1.3 Melignant Melanoma

Although malignant melanoma accounts for less than 2% of all skin cancer cases, it is the deadliest type and causes the vast majority of deaths [149]. The incidence of melanoma has increased in the past decades to currently reaching 132,000 melanoma cases per year, according to the World Health Organization. The American Cancer Society also reported the estimated deaths of melanoma in 2014 as 9710 individuals and new cases as $76,100$ individuals.

Melanoma cancer is incurable in its advanced stages and the patient should go through surgery, possibly immunotherapy, chemotherapy, and/or radiation therapy. However, if it is diagnosed at its early stage, it is the most treatable kind of cancer [149, **?**]. In fact the patient survival rate has increased significantly, in the past decades, thanks to early diagnosis and treatment of melanoma in its early stages.

The stage of melanoma is measured based on how the lesion is expanded, including its invasion depth though the skin to nearby lymph nodes or other organs. This factor is measured through physical exam, biopsies, and different imaging tests such as computed tomography or magnetic resonance imaging (MRI) [149]. Depending on the obtained measurements, melanoma skin cancers are divided into four types.

The first three types begin *in situ*, meaning that they spread in the top layers of the skin and in the last stages they become invasive, while the fourth one is invasive from the start. Invasive melanoma could be more dangerous since they are in the deeper layers of the skin and they can spread faster to other body parts. The four types of cutaneous melanoma are listed in the following:

Superficial spreading melanoma is the most common type and is the leading cause of cancer death in young adults. Approximately 70% of all melanoma diagnosis is counted as superficial spreading melanoma. This type grows along the top layer of the skin and often occurs in a previously benign mole. The location of this melanoma can be anywhere in the body, however it is usually located on the trunk and back in male patients and on the legs and back in females patients.

Acral lentiginous melanoma accounts for less than 5% of all melanoma's diagnosis and is the most common type in dark skin individuals. This disease usually is located on the palms, soles of the feet and under the finger nails and often look like a bruise or injuries on the body. For this reason, melanoma may be discovered later than other forms.

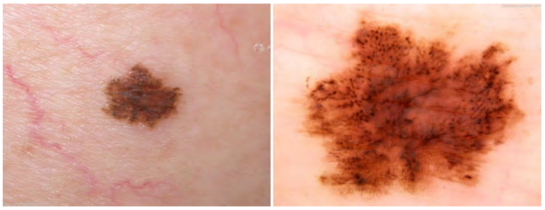
Lentigo maligna melanoma accounts for 5-10 % of the melanoma diagnosis. Letigo maligna melanoma arises from a pre existing lentigo rather than a mole and mostly occurs on the face of middle-aged to elderly individuals as a result of sun damage. If this cancer stay undiagnosed, being mistaken with sun spot, it can spread to the deep layers of skin and danger the patients life. The cancer lesions of this type, usually have a very irregular border and vary in shades of brown or black however like other types of melanoma, they can be blue, red, gray or white.

Nodular melanoma accounts for 15-30% of all melanoma diagnosis. This is the most aggressive type due to the fact that it spreads more rapidly in depth and it is difficult to visualize the progression of the cancer. Nodular melanoma is more common in males compared to females. The lesion is usually darkly pigmented (blue-black) and often is found in pink or red.

1.4 Melanoma Diagnosis and Screening

The clinical prognosis of early stage melanoma is commonly via visual inspection of the lesions based on set of rules or guidelines, such as "ABCDE" [9] or Glasgow 7-point checklist [9]. The "ABCDE" rule characterizes the lesion based its asymmetry (A), irregular borders (B), variegated colors (C), diameter $\geq$ 6 mm (D) and evolving stage over time (E) while the Glasgow 7-point checklist contains 7 criteria: 3 major (changes in size, shape and color) and 4 minor (diameter $\geq$ 7mm, inflammation, crusting or bleeding, and sensory change). The former rule has been extensively applied in clinical routine, mainly due to its simplicity, rather than the latter one.

The visual inspection of lesions are carried out using different non-invasive imaging techniques such as: clinical photography, dermoscopy, confocal laser scanning microscopy (CLSM),



(a) In situ melanoma (stage 0)

(b) Invasive melanoma (stage I/II)

Figure 1.2: Clinical and Dermoscopy images, right and left column, respectively. Images submitted to www.dermoscopyatlas.com by Dr. Alan Cameron (a), Dr. Jean-Yves Gourhant (b). Used with permission.

optical coherence tomography (OCT), multispectral dermoscopy (MSD), high frequency ultrasound (US), and magnetic resonance imaging (MRI) among other spectroscopic imaging. Within the aforementioned techniques, some are well-utilized by the clinicians and dermatologists. We will refer here and after to these techniques as "conventional" techniques. Clinical photography, dermoscopy, OCT, and CLSM belong to this category. While the rest, such as MRI, US and MSD are categorized as "non-conventional" techniques.

We focus our research on conventional techniques such as clinical and dermoscopy as well as non-conventional techniques such as MSD and polarized imaging (PI).

1.4.1 Clinical photography

Clinical photographs are referred to digital or non-digital images captured from the surface of the skin, showing one or multiple lesions. These images reproduce what a clinician sees with the naked eye [53] and commonly are tainted with unnecessary highlights and reflections from the outer layer of the epidermis. The left column of Figs 1.2 shows such images.

1.4.2 Dermoscopy

Dermoscopy (also known as dermatoscopy or epiluminescence microscopy) is a well-established and effective non-invasive technique for early recognition of melanoma. This technique, which

(a) Clinical (b) NPD (c) PD

Figure 1.3: Melanoma lesion shown by clinical photography, NPD, PD, respectively. Shiny-white streaklike (arrow in (c)) within the melanoma are only visible with NPD. The images are taken from Benvenuto et al. [25]

was introduced in 1971 [103, 102], uses a hand-held lighted magnifier to analyze the skin lesions.

Initially the device was used in conjunction with thin layer glass and oil or alcohol interface to reduce light reflection, refraction and diffraction. This material made the epidermis translucent and allowed in vivo visualization of subsurface anatomic structures of epidermis and papillary dermis which are not visible with naked eye [131, 159]. This type of dermoscopes are called NPD [159].

The polarized dermoscope (PD), second type, was introduced later and made this process much easier by using cross polarized light. This device is equipped with two polarized filters, one in front of the light source and one in front of the sensor. The two polarized filters are perpendicular to each other in order to capture the backscattered light from the deeper levels of the skin. The light reflected from the skin has the same angle as the polarized incident light, hence it will be eliminated with the cross polarized filter in front of the sensor. However, the backscattered light from the skin, due to the structural nature of the tissue, will become unpolarized and passes through the filter. This technique eliminates the need for fluid or direct contact with the skin.

The captured images using PD or NPD are relatively similar. However the surface dermoscopic structures (such as blue white veil) are better observed with NPD while the deep structures (such as vessels) are better seen with PD [159, 25]. The right column of Fig 1.2 demonstrate the dermoscopic images of the same lesions captured with NPD. Figure 1.3 shows difference of clinical, NPD, and PD imaging for a melanoma lesion.

A sample of commercially available dermoscopes (Dermlite®) with and without polarized light are shown in Fig 1.4. MoleMax (Derma Medical System, Vienna, Austria) is another commercially available system that includes a polarized dermoscope. MoleMax is a computer based device and providing its own software it allows live-video dermoscopy and total body

photography [110, 131].

Beside NPD and PD, another dermoscopy technique, via transillumination, was introduced by Dhawan et.al. [55, 56]. In this technique, the light is directed to the skin in such a way that allows the backscattered light illuminate the lesion from within. Novescope is a patented device, developed for this technique.

(a) Dermlite® II Fluid (b) Dermlite® II Pro HR

Figure 1.4: Commercially available dermoscopes by Dermlite®: (a) immersion fluid dermoscope using non-polarized light; (b) cross polarized light dermoscope. Both devices can be attached to a digital camera.

1.4.3 Polarized Imaging

In PI or image polarimetry systems a polarizer state generator and analyzer are used to create a set of polarized images. These images define the polarization state of light beam and depolarization property of the tissue. The former is represented by four measurable quantities called stokes parameters and the later is represented by the Mueller matrix. The advantage and benefits of PI systems beside cross-polarized is not so evident in the field of skin imaging or tissue imaging, in general. In the last decades few studies are dedicated to find and explore the depolarization properties of the tissues. Anastasiadou et al. proposed a Mueller polarimetry system for diagnosing the cervical cancer [11] and Manhas et al. proposed a Mueller polarimetry system for measuring polarized diffuse reflectance on cancerous and non-cancerous regions of oral cavity and breast tissues [106]. In the field of skin tissue, Jacques et al. proposed stokes polarimetry to differentiate different pigmented lesions [106]. Another work recently was proposed by Tchvialeva et al. were polarized speckle imaging was suggested for in vivo screening of pigmented lesions [106].

The stokes polarimetry, Mueller matrix measurements, and the proposed methods by the research community are further explained in Chapt. 4.

1.4.4 Multispectral Imaging

In multispectral imaging, a set of wavelength dependent images are acquired in order to visualize different depths and layers of the skin. This is possible since depth of light penetration into the skin is directly related to wavelength. The wavelengths can range from ultra-violet to near infrared. This technique offers the advantage of analyzing the sub-layers which are not visible to the human eye, probing up to 2 mm below the surface of the skin [131] (see Fig. 1.5).

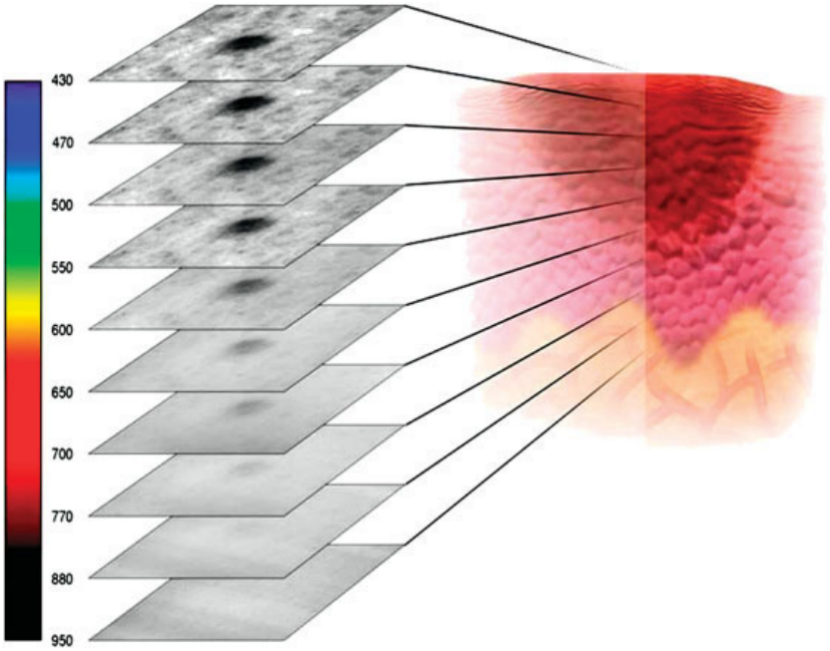


Figure 1.5: Multispectral analysis of pigmented skin lesion. Longer wavelength penetrates deeper into the skin, providing diagnostic features. The image is taken from Rigel et al. [131]

Opposed to standard color cameras which capture the skin image in three channels (red, green, blue), Multispectral Imaging (MSI) acquires a sequence of gray level images at different wavelengths. This extended spectral capacity provides a high advantage in skin imaging since it is reducing the effects of metameric mismatches, which may occur due to different illumination and variability of sensor spectral responses [91]. The spectral device of MSI can be placed either along the incident path, or in detection path. In either way, the main motivation is to go beyond the limited red, green and blue color information which are available by other imaging techniques. This technique combines the advantage of both a spectrophotometer and a digital camera. The MSI system acquires three-dimensional data, the spatial information in 2 dimensions and a spectral information in the third dimension, obtaining spectral information

for each pixel in the image.

Currently, several multispectral dermoscopes are commercially available: MelaFind [63, 83], SolarScan [115, 114], and SIAscope [118]. This topic is extensively discussed in Chapt. 5.

1.5 Automated diagnosis of melanoma

Computer-aided diagnosis (CAD) or clinical diagnosis support (CDS) systems are proposed to provide automated diagnosis of melanoma lesions. These systems are intended to reproduce the decision of the dermatologists when observing images of PSLs. Automated diagnosis of melanoma was proposed to assist the dermatologists and increase the sensitivity and specificity of melanoma recognition in early stages as well as to reduce the unnecessary excisions. From a computer vision and pattern recognition point of view, CAD systems for melanoma intend to classify and differentiate melanoma lesions from others. In general, image processing techniques are used to locate and delineate the lesions and extract the image parameters (features) which coincide with the dermatologist's point of view and dermatological features. The extracted parameters are further used with machine learning tools to perform the diagnosis (classification). Such systems are being developed for various imaging modalities [109, 157, 95]. However, dermoscopy, being the most conventional imaging technique, most CAD systems are dedicated to this modality. In our research, we considered automated classification of melanoma based on conventional dermoscopy, and non-conventional MSI and PI imaging. The general steps of a CAD system are extensively discussed in Chapt. 2.

Clinical impact

Numerous research projects are dedicated to the development of a CAD or CDS system and studies show that their performance is sufficient under experimental conditions [73]. The proposed systems can assist the dermatologists. However, their practical value is still unclear and they cannot be recommended as a sole determinant of malignancy of a lesion. Even though most patients would accept computerized analysis of melanoma, the proposed CAD systems cannot function alone due to their tendency to over diagnose benign melanocytic and non-melanocytic lesions [73]. Day and Barbour [52] listed two main shortcomings for general approaches which are adapted to develop a CAD system for recognition of melanoma:

1. A CAD system is expected to reproduce the decision of pathologists (a binary result like "melanoma/non-melanoma lesion") with only the input used by dermatologists: clinical or dermoscopic images;

2. Histopathological data are not available for all lesions, only for those considered suspicious

by dermatologists.

The listed items refer to the lack of sufficient information and interaction with dermatologists used in the proposed method. The listed items refer to the fact that the proposed methods lack in sufficient information and interaction with dermatologists. This fact was highlighted by Dreiseitl et al. [58] as well. The authors mentioned that current systems are designed to work "in parallel with and not in support of" physicians, thus only few systems are adapted in clinical routines. In this regard, an ideal CAD or CDS system for melanoma recognition should be able to reproduce the dematologists's decision with extensive information regarding the reason and ground of that decisions [58].

1.6 Research Motivation

To be written

1.7 Thesis outline

This thesis describes the research work that resulted in the development and validation of classification frameworks for differentiation of melanoma using conventional and non-conventional imaging techniques. Dermoscopy is considered as conventional technique and MSI and PI modalities are explored as non-conventional techniques.

Since classification is the base of our CAD systems, related machine learning and computer vision techniques are discussed in **Chapter 2**.

Chapter 3 is dedicated to dermoscopy modality. State of the art of CAD systems, our proposed framework, experiments and obtained results for dermoscopy modality are depicted in this chapter.

Chapter 4 presents the framework developed for PI modality. PI system and developed classification framework along obtained results are presented in this chapter.

Chapter 5 discusses the framework developed for MSI modality. This chapter contains the MSI system, proposed the framework, and the obtained results for this modality.

Finally **Chapter 6** concludes the thesis and presents avenues for future research.