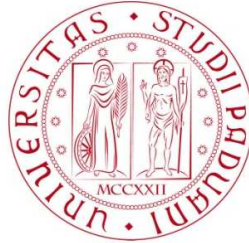


UNIVERSITÀ DEGLI STUDI DI PADOVA
DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE
CORSO DI LAUREA MAGISTRALE IN BIOINGEGNERIA



IDENTIFICATION AND ESTIMATION OF CLINICAL INDICES
USEFUL FOR THE DIAGNOSIS OF MELANOMA FROM
MACROSCOPIC IMAGES

RELATORE:
PROF. FABIO SCARPA

LAUREANDA:
SILVIA IACUMIN

ANNO ACCADEMICO 2022/2023

DATA DI LAUREA 12-12-2023

Abstract

Melanoma is an extremely aggressive form of skin cancer. When not promptly detected and treated, it can quickly metastasize, leading to unfavourable prognostic outcomes. Achieving early melanoma diagnosis relies heavily on accurate and thorough skin analysis, made by an expert dermatologist.

To address subjective judgments and time-expensive exams, a novel screening and diagnostic method utilising photogrammetry-derived images of skin lesions has been devised. This innovative approach is based on the acquisition of macroscopic images, depicting a large portion of the patient body, and enables the creation of a three-dimensional model of the patient, allowing for the extraction of corresponding images of each individual lesion.

This thesis aims to quantitatively assess the asymmetry, the irregularity of the border and the color of skin lesions through the analysis of segmented macroscopic images, contributing to the development of an automated diagnostic tool useful to the clinician for melanoma identification. The analysis was conducted on a dataset comprising images of healthy skin lesions and lesions reported as suspicious by dermatologists among which nine cases were confirmed as melanomas by biopsy.

By utilizing algorithms to objectively compute asymmetry and border irregularity parameters, coupled with an in-depth analysis of color features associated with melanocytic lesions, the investigation unveiled statistically significant differences in these attributes between benign and suspicious lesions. Indeed, statistical tests confirmed distinctive distributions of these parameters between the two skin lesion populations. These findings underscore the potential of automated diagnostic tools derived from macroscopic

images in effectively identifying suspicious lesions, thus contributing to early melanoma detection strategies.

Contents

Introduction	1
1 Melanoma and the current diagnostic tools	5
1.1 Melanoma	5
1.2 Epidemiology	5
1.3 Importance of early detection	6
1.3.1 Common benign pigmented lesions	7
1.3.2 Precursor lesions of malignant melanoma	9
1.3.3 ABCDE rule	10
1.4 Clinical instruments for the diagnosis of melanoma	12
1.4.1 Dermoscopy	12
1.4.2 Total body photography	13
1.4.3 Reflectance confocal microscopy	13
1.4.4 Computer-assisted diagnosis techniques	13
2 Automatic ABCD rule extraction for melanoma detection	15
2.1 Pre-processing	15
2.2 Segmentation	17
2.3 ABCD feature extraction	19
2.3.1 Asymmetry	19
2.3.2 Border Irregularity	20
2.3.3 Color	21
2.3.4 Diameter	22
2.4 Classification based on ABCD rule	23
2.5 Example of feature combination and classification	23
3 Device for macroscopic image acquisition, 3D modeling and cropping of skin lesions	27
3.1 Device description	28
3.2 3D Model and skin lesion crop	29
3.3 Dataset description	32
4 Quantitative evaluation of asymmetry, border irregularity and color	35
4.1 Dummy images	36
4.2 Asymmetry	39

4.2.1	Translation and rotation	39
4.2.2	Non-Overlapping area	42
4.2.3	Asymmetry scores	44
4.3	Border irregularity	45
4.3.1	Standard deviation of center-of-mass to border distances	45
4.3.2	Compactness Index	49
4.3.3	Convex Hull	50
4.3.4	Standard deviation of ellipse-to-border distances	53
4.4	Color	60
5	Results	65
5.1	Validation of methodologies on dummy images	65
5.1.1	Validation of methodologies for assessing asymmetry	66
5.1.2	Validation of methodologies for assessing border irregularity	67
5.1.3	Parameters chosen to assess border irregularity	77
5.2	Evaluation of asymmetry in benign and suspicious skin lesions	77
5.3	Evaluation of border irregularity in benign and suspicious skin lesions	86
5.4	Evaluation of color in benign and suspicious skin lesions	98
6	Conclusions	105
	Bibliography	109

List of Figures

1.1	Common benign pigmented lesions	8
1.2	Precursor lesions of malignant melanoma	10
1.3	ABCDE rule	11
2.1	Pre-processing steps	16
2.2	Segmentation	18
2.3	Types of border irregularities	20
3.1	Structure of the device	28
3.2	Device prototype present at the IOV	29
3.3	12 patient’s back images	30
3.4	Marked suspicious lesion	30
3.5	Moles identification	31
3.6	Crop images	31
4.1	Some reference images from HAM10000 dataset	37
4.2	Dummy images for border irregularity evaluation.	37
4.3	Dummy images for asymmetry evaluation.	38
4.4	Translation and rotation of dummy images.	41
4.5	Translation and rotation of suspicious lesion images.	41
4.6	Division of the object based on the horizontal axis.	42
4.7	Division of the object based on the vertical axis.	42
4.8	Non-Overlapping area over horizontal axis.	43
4.9	Non-Overlapping area over vertical axis.	43
4.10	Non-Overlapping area (NOA) defined in pixels.	44
4.11	Representation of the 8 binary masks.	46
4.12	Mask of suspicious lesion and its border.	46
4.13	Division of the border of the suspicious lesion in the 8 sectors.	47
4.14	Division of the suspicious lesion in the 8 sectors and respectively value of the standard deviation of the sector.	48
4.15	Distribution of the standard deviation values of the 8 sectors.	49
4.16	Compactness Index	50
4.17	Convex Hull on dummy images.	51
4.18	Convex Hull on suspicious lesion masks.	52
4.19	Elliptical approximation on a dummy image.	54
4.20	Elliptical approximation on a suspicious lesion mask.	55

4.21	Ellipse method application 1	57
4.22	Ellipse method application 2	58
4.23	Ellipse method application 3	59
4.24	Color Mask	60
4.25	Lesion with dark brown, light brown and blue-gray colors	63
4.26	Lesion with dark brown and light brown colors	63
4.27	Lesion with light brown and white colors	64
4.28	Lesion with black, dark brown, light brown and blue-gray colors	64
5.1	Comparison of PD	73
5.2	Asymmetry on low resolution image.	80
5.3	Number of clusters for asymmetry dataset.	81
5.4	Clusters of asymmetry dataset.	81
5.5	Asymmetry parameter distribution	83
5.6	Boxplot of Asymmetry parameter distribution	84
5.7	Distribution of clustering values for asymmetry	84
5.8	Elbow method used in border irregularity assessment.	88
5.9	Clusters of border irregularity.	89
5.10	Clusters visualization with PCA.	90
5.11	Sum parameter distribution	93
5.12	Boxplot of Sum parameter distribution	93
5.13	PD parameter distribution	94
5.14	Boxplot of PD parameter distribution	94
5.15	CI parameter distribution	95
5.16	Boxplot of CI parameter distribution	95
5.17	Scatter plot of Asymmetry-Sum	96
5.18	Scatter plot of Asymmetry-PD	96
5.19	Scatter plot of Asymmetry-CI	96
5.20	Distribution of clustering values for border irregularity	97
5.21	Color distribution in benign and suspicious lesions	101
5.22	Malignant color distribution in benign and suspicious lesions	103

List of Tables

2.1	Presentation of criteria of ABCD rules	25
4.1	Image categories of Figure 4.1 [34].	36
4.2	Percentage Discrepancy of images of Figures 4.17 and 4.18	52
4.3	Color Ranges	61
5.1	Asymmetry validation results	67
5.2	Standard deviation of center of mass-border distances validation results	69
5.3	Distribution Statistics for Std of Distances	70
5.4	Validation of CI and PD	72
5.5	Standard deviation of ellipse-border distances validation results	75
5.6	Distribution Statistics for Std of Distances	76
5.7	Asymmetry evaluation on suspicious and benign lesion masks	79
5.8	Cluster labels obtained from asymmetry assessment	82
5.9	Distribution of clustering values for asymmetry	85
5.10	Border irregularity evaluation on suspicious and benign lesion masks	87
5.11	Cluster labels obtained from border irregularity assessment	91
5.12	Distribution of clustering values for border irregularity	97
5.13	Results of color assessment in suspicious lesions	99
5.14	Results of color assessment in benign lesions	100
5.15	Numeric Values Table	101
5.16	Percentage values Table	101
5.17	Malignant colors numeric values	102
5.18	Malignant color percentage values Table	102

Introduction

This thesis aims to quantitatively assess the values of asymmetry, irregularity of the borders and color of skin lesions using macroscopic images. Asymmetry and border irregularity are two clinically significant parameters in the context of diagnosing skin lesions. Asymmetry refers to the difference in shape or structure between the two halves of a lesion. In a benign lesion, symmetry is often observed, while in a potentially malignant lesion, such as melanoma, asymmetry raises concern. Border irregularity indicates variations in the shape or definition of the margin of a lesion. While a benign lesion typically exhibits a regular and well-defined border, a lesion with irregular borders may require further attention. Asymmetry and border irregularity are pivotal parameters in assessing skin lesions, yet considering color variations adds further diagnostic insight. While benign lesions typically exhibit a single shade of brown, melanomas often present with multiple shades of brown, light brown, or black. Moreover, as a melanoma progresses, it might display additional colors such as red, white, or blue. The presence of multiple colors within a lesion serves as a warning sign, indicating the need for closer examination in the diagnostic process. Integrating color assessment alongside asymmetry and border irregularity analysis in macroscopic images contributes significantly to comprehensive lesion evaluation.

Dermatologists traditionally employ these parameters as an integral part of their assessment of skin lesions. For instance, pronounced asymmetry or irregular borders may raise suspicions of potential malignancy. In many cases, asymmetry, border irregularity and color are included in the ABCDE rule, an acronym denoting Asymmetry, Border irregularity, Uneven Color, Diameter greater than 6 mm, and Evolution over time. This rule provides a practical guide for lesion analysis and can help identify warning signs.

However, there are some critical issues in current clinical practice. Studies have shown that visual assessment, while fundamental, can be influenced by subjective factors and heavily relies on the experience and training of the dermatologist. Moreover, the time required for an accurate diagnosis can be prolonged, potentially affecting the timeliness of medical intervention. In some cases, subjective judgment can lead to erroneous diagnoses or delays in identifying potentially dangerous lesions.

Images obtained through dermatoscopy, pivotal for assessing skin lesions, are traditionally acquired using a dermatoscope. This optical instrument allows an enlarged view of skin lesions, placed directly on the skin and often employing a gel to reduce reflections and enhance light transmission. The resulting images possess high resolution, enabling a detailed evaluation of lesion characteristics.

The need to enhance objectivity and efficiency in evaluating skin lesions has driven the evolution of image acquisition methodologies. In this context, the innovative image acquisition method employed in this research represents a significant breakthrough in the field of dermatology. By using a system composed of 12 cameras, it is capable of simultaneously capturing a wide portion of the patient's skin surface, particularly in this study, focusing on the entire back area. From these photos, a 3D model is created, and individual images of skin lesions are extracted. However, it is important to note that these macroscopic images have a relatively low resolution, posing a significant challenge in the accurate evaluation of asymmetry and border irregularity.

The research carried out and the algorithms developed within this thesis primarily aim to provide an automated tool capable of objectively and reliably calculating values of asymmetry, border irregularity and color on macroscopic images of individual skin lesions. This approach aims to eliminate subjectivity in melanoma identification, enabling more timely and accurate diagnoses.

The thesis is divided into six chapters. In the first chapter, an overview of melanoma is provided, including its epidemiology and the significance of early diagnosis. The fundamental clinical distinctions between benign lesions and melanomas are emphasized. Additionally, the current diagnostic practices employed by dermatologists for melanoma detection are examined.

The second chapter reviews the methodologies currently available in the literature, with a specific focus on approaches based on the automatic estimation of ABCD parameters. These have been developed primarily for dermoscopic images, which represent the main area of research in the evaluation of skin lesions.

In the third chapter, the device used for the acquisition of the images employed in the research is presented in detail. Additionally, a comprehensive description of the image dataset utilized throughout the study is provided.

The fourth chapter provides a detailed description of the algorithms developed specifically for assessing the clinical parameters of asymmetry, border irregularity and color on macroscopic images containing individual skin lesions.

The fifth chapter, on the other hand, presents the results obtained from the implementation of these methodologies. Finally, the conclusions drawn from the research are reported.

Chapter 1

Melanoma and the current diagnostic tools

1.1 Melanoma

Melanoma, a form of skin cancer, originates from the malignant transformation of melanocytes. These cells are found in the epidermis, the most superficial layer of the skin, and are responsible for producing melanin, a pigment that protects against the damaging effects of sunlight. Under normal conditions, melanocytes can give rise to dark agglomerates visible on the surface of the skin and are known as moles or nevi in medical terms.

Melanoma can develop either on apparently healthy skin or from the modification of a pre-existing mole. While it accounts for only about 1% of skin cancers, it is responsible for most skin cancer-related deaths. This is because melanoma is particularly dangerous due to its higher propensity to metastasis to other parts of the body if not detected and treated in its early stages [1].

1.2 Epidemiology

In recent decades, there has been a significant rise in the incidence of melanoma. Globally, approximately 100,000 new cases are documented each year and this number has

doubled in the last decade.

Melanoma shows a significant demographic pattern, with a ten times higher prevalence among individuals of Caucasian origin compared to other ethnic groups. In fact, approximately 85% of annual cases of cutaneous melanoma worldwide are reported in populations of North America, Europe, and Oceania [1]. Particularly high incidence rates are observed in regions characterized by abundant sunshine and inhabited predominantly by light-skinned northern European communities.

In Italy, more than 14,000 cases are estimated to have occurred in 2020, according to the AIRTUM 2020 data [2]. Over the past five years, the deaths attributed to cutaneous melanoma have been 4.000 among males and exceeded 3.000 among females. This translates to an average mortality incidence rate of 5 cases per 100.000 inhabitants annually for males and 6 cases for females. The incidence rate shows peak values above 10 per 100.000 for both sexes in the population of Trieste and above 6-7 per 100.000 in Genova, Veneto and Romagna [2]. The Intergruppo Melanoma Italiano indicated that Italy exhibits an age-standardized 5-year net survival rate of 87.8% for men and even more encouraging 90.8% for women, based on incident cases from 2010 to 2014, with follow-up data available until 2018. Considering the different Italian geographic areas, there are differences in 5-year net survival rates: from 88% in men and 91% in women in the Northeast to 85% in men and 89% in women in the Northwest and Central Italy, to 80% in men and 84% in women in Southern Italy, as reported in the AIOM Guidelines of 2021 [3].

Over the past decade, from 2011 to 2020, advancements in treatment have led to a significant decline in melanoma mortality rates. Among adults younger than 50 years, rates decreased by approximately 5% annually, while in individuals aged 50 and older, the decline averaged 3% annually [1].

1.3 Importance of early detection

Early diagnosis of malignant melanoma is the most effective method of treating it. The prognosis of cure after surgical removal is excellent in these cases. The 10-year survival rate stands at 75%-95% for malignant melanoma measuring less than 0.76 mm in thickness (Breslow's depth), indicating an absence of metastases [4]. In contrast, the rate

drops to 48% for melanomas measuring 3 mm or more in thickness [5]. While exceptions exist, the earlier the diagnosis, meaning the thinner the malignant melanoma, and the earlier the surgical treatment, the higher the likelihood of survival.

In 1985, Robert J. Friedman et al. [5] emphasized the crucial role of physicians in enhancing the chances of early diagnosis and prompt eradication of malignant melanoma. This can be achieved through patient education on self-examination of the skin, conducting accurate, periodic, complete skin examinations, and using diagnostic techniques such as biopsy.

To accurately assess whether the mole under examination is highly suspicious for melanoma, the physician must possess a thorough understanding of the clinical characteristics of common pigmented skin lesions, as well as the features of dysplastic nevi and of the congenital melanocytic nevi, which can give rise to melanoma. Familiarity with the clinical traits of early malignant melanomas and knowledge of risk factors, such as family history, the presence of many melanocytic nevi, fair skin, history of sunburn or excessive sun exposure, and the age of the patient, further aids in this diagnostic process [5].

1.3.1 Common benign pigmented lesions

The most frequently encountered pigmented lesions are simple lentigines, common moles (i.e., melanocytic nevi junctional, compound, intradermal), solar lentigines, and seborrheic keratoses. The lesions are described briefly below [5].

- Simple lentigo (Figure 1.1 (a)) is a sharply defined, round, brown to black, macular lesion with smooth or jagged edges that may appear anywhere on the surface of the skin.
- Junctional nevus (Figure 1.1 (b)) is generally a small, less than 6 mm, macular, well-circumscribed, pigmented lesion with a smooth surface and relatively uniform pigmentation that ranges from light brown to very dark brown to black.

- Compound nevus (Figure 1.1 (c)) is generally a well-circumscribed, small, less than 6 mm, raised papule that is uniformly pigmented with a range of colors from skin-colored to tan to shades of brown and a smooth or rough surface.
- Intradermal nevus (Figure 1.1 (d)) is generally a small, up to 6 mm, well-circumscribed papule that is usually uniform in pigmentation from skin-colored to tan or various shades of brown. The surface may be smooth or rough.
- Solar lentigo (Figure 1.1 (e)) is generally a uniform tan to brown macule, known to the lay public as a liver spot.
- Seborrheic keratosis (Figure 1.1 (f)) is generally a verrucous, round or ovoid, variably raised, light brown to black, sharply demarcated papule or plaque that ranges in diameter from a few millimeters to several centimeters.

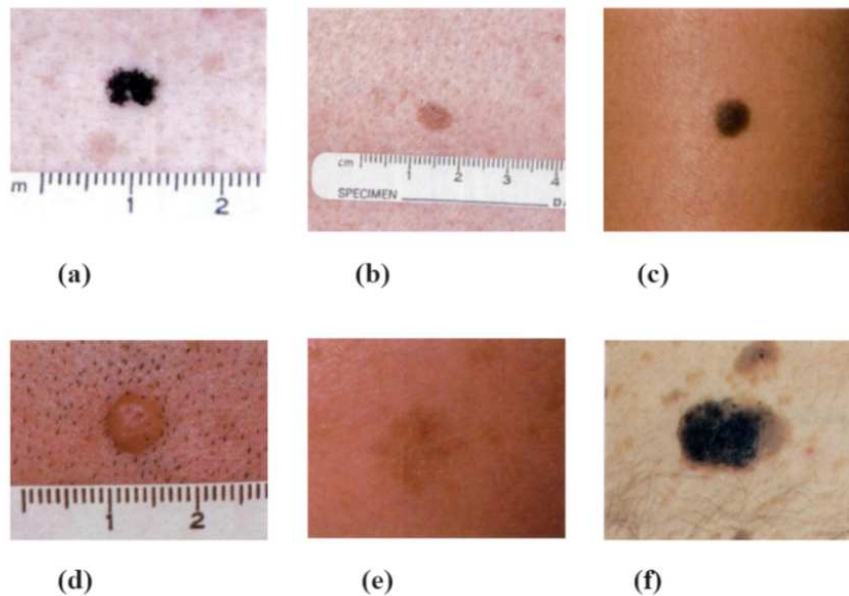


FIGURE 1.1: Common Benign Pigmented Lesions: (a) Simple lentigo; (b) Junction nevus; (c) Compound nevus; (d) Intradermal nevus; (e) Solari Lentigo; (f) Seborrheic keratoses [5].

1.3.2 Precursor lesions of malignant melanoma

Certain pigmented cutaneous lesions may give rise to malignant melanomas. These lesions include acquired abnormal melanocytic nevi known as dysplastic nevi and certain congenital melanocytic nevi.

- Dysplastic nevi are typically larger than ordinary nevi, generally ranging from 6 to 12 mm or more in diameter. They usually have both macular and elevated components. The borders of dysplastic nevi, unlike those of common nevi, are often irregular and frequently so ill-defined that they fade imperceptibly into the surrounding skin. Dysplastic nevi usually are variegated in color, ranging from tan to dark brown, at times with a prominent pink component. Some, however, cannot be distinguished clinically from malignant melanomas. In Figure 1.2 some examples of dysplastic nevi are presented.

Common melanocytic nevi usually do not develop after the age of 40, whereas dysplastic nevi generally begin in adolescence and continue to appear throughout life.

Overall, patients with dysplastic nevi have a reported lifetime risk for malignant melanoma of approximately 5 to 10 %, compared with the risk of about 0.7 % for the general population [5].

- Congenital melanocytic nevus is a lesion that contains nevus cells and is present at birth, can be categorized by size as follows:
 - Small: less than 1.5 cm in diameter.
 - Medium: 1.5 to 19.9 cm in diameter.
 - Large: 20 cm or more in diameter.

The lifetime risk of malignant melanoma in patients with large congenital melanocytic nevi has been estimated to be about 6 %, compared with a risk of 0.7 % for the general population. An example of giant congenital melanocytic nevi is reported in Figure 1.2 [5].

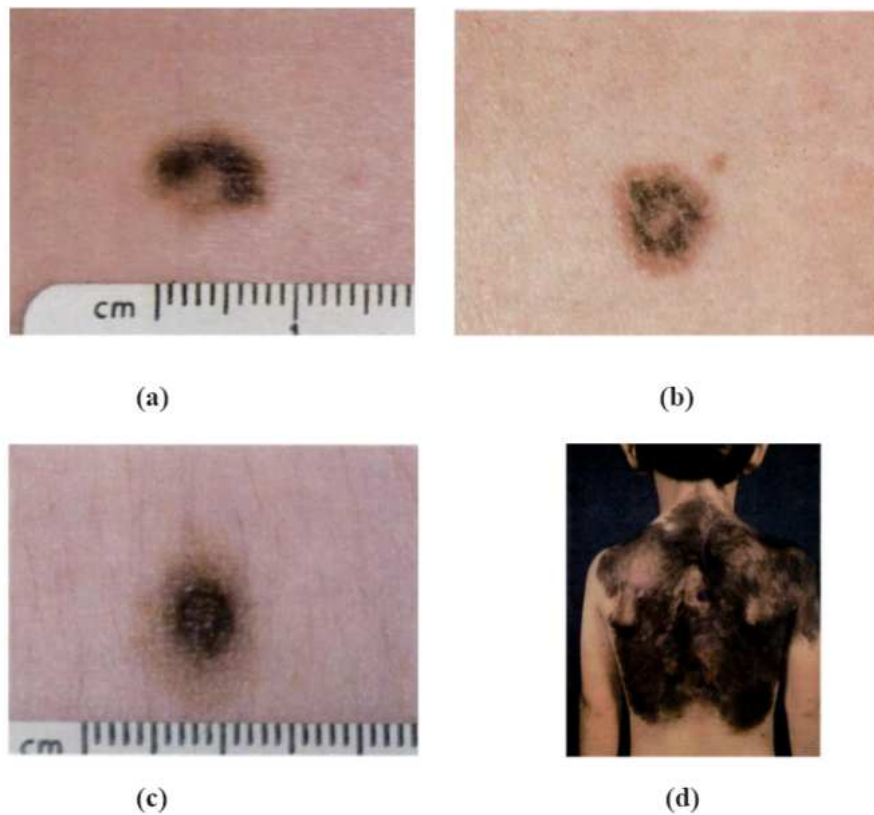


FIGURE 1.2: Precursor Lesions of Malignant Melanoma: (a) Dysplastic nevus measuring 9 mm in diameter; (b) Dysplastic nevus with color variegation from tan to brown; (c) Dysplastic nevus with features of a dark-target variant; (d) Giant congenital melanocytic nevi [5].

1.3.3 ABCDE rule

The ABCD acronym for melanoma screening was devised in 1985 [5] to provide the lay public and primary health care professionals with a useful and memorable mnemonic to aid in the early recognition of potentially curable cutaneous malignant melanoma. The now well-known parameters of Asymmetry, Border irregularity, Color variegation, and Diameter greater than 6 mm are used globally in medical education and in the lay press to provide simple parameters for appraisal of pigmented cutaneous lesions that may need to be further examined by a specialist [6].

Unlike benign pigmented lesions, which are generally round and symmetrical, with regular margins, uniform color and diameter less than 6 mm, early malignant melanomas are usually asymmetrical. Their contours are usually irregular, their color is variegated, ranging from various hues of tan and brown to black, and sometimes intermingled with red and white. In addition, cutaneous melanomas, when first detected, often have a

diameter of more than 6 mm [5].

Not all melanomas have all 4 ABCD features. It is the combination of features (e.g., ABC, A+C, and the like) that render cutaneous lesions most suspicious for early melanoma [6].

The diagnosis of malignant melanoma is based not only on clinical appearance but also on history and symptomatology. A change in a pre-existing melanocytic nevus or the development of a new pigmented lesion, particularly after the age of 40, is important in alerting the physician to the possibility of a malignant melanoma [5].

The studies conducted in 2004 [6] strongly support the inclusion of an 'E' for 'Evolving' in the ABCDE criteria for diagnosing cutaneous melanoma. This addition significantly enhances both physicians' and the general public's ability to identify melanomas at earlier, more treatable stages. The 'E' acknowledges the dynamic nature of this skin malignancy. It refers to 'evolving lesions', which are those observed to have changed in terms of size, shape, symptoms (such as itching or tenderness), surface characteristics (such as bleeding), or variations in shades of color. These characteristic clinical features of early malignant melanoma are reported in Figure 1.3.

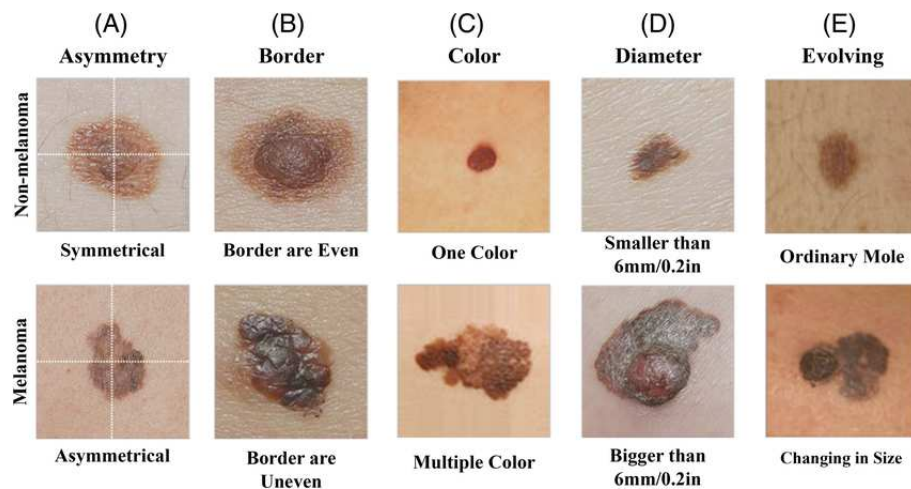


FIGURE 1.3: Example of ABCDE comparison between non-melanoma nevi and melanoma [7].

1.4 Clinical instruments for the diagnosis of melanoma

The clinical diagnosis of melanoma is based on [4]:

- total body visual skin examination for the detection of lesions displaying one or more of the ABCDE criteria;
- intra-individual comparative analysis, which is searching for the lesion that is not like the others in the same patient (ugly duckling sign);
- assessment of the evolution of lesions in case there is available documentation.

1.4.1 Dermoscopy

Dermoscopy or dermatoscopy refers to the examination of the skin using skin surface microscopy, which requires a high-quality magnifying lens and a powerful lighting system (a dermatoscope). Computer software can be used to archive dermoscopy images and allow expert diagnosis and reporting (mole mapping). Smart programs may aid in diagnosis by comparing the new image with stored cases with typical features of benign and malignant pigmented skin lesions [8].

Dermatoscopy should always be used in the clinical assessment of skin tumors, it should be applied on all lesions and not only on clinically suspicious ones. This is because dermatoscopy has the potential to uncover the morphologic asymmetry of melanoma before it becomes clinically recognizable and reveal clues that are strongly suggestive of melanoma [4].

Studies evaluating the diagnostic accuracy of clinical examination alone have shown that dermatologists are able to identify melanoma in 65-80% of cases. A recent systematic literature review has demonstrated that dermoscopy can increase the diagnostic sensitivity for melanoma by up to 35% compared to clinical observation alone. It has also been reported that this diagnostic improvement can only be achieved if the observer has a good level of experience in using the technique, whereas the accuracy of dermoscopic diagnosis may be even worse than clinical diagnosis for those who are not experienced. Therefore, adequate preparation is crucial for a truly effective diagnostic application [9].

1.4.2 Total body photography

Total body photography (TBP) is useful for following patients who are at high risk for melanoma. One of the goals is to detect subtle changes over time in lesions that may otherwise go unnoticed by pure clinical examination; this is especially useful in a patient with 100s of lesions. A patient initially undergoes a series of baseline images. These photographs are then compared with the patient's examination at follow-up visits. Lesions that remain the same are presumed benign, barring specific clinical or dermoscopic signs of melanoma. New and changing lesions can be further evaluated, through a biopsy, or followed for additional change [10].

1.4.3 Reflectance confocal microscopy

Reflectance confocal microscopy is an optical imaging technology that offers noninvasive visualization of skin lesions in vivo at nearly histologic resolution [11]. It may have a potential role in clinical practice, particularly for the assessment of lesions that are difficult to diagnose using visual inspection and dermatoscopy alone. Evidence suggests that reflectance confocal microscopy may be both more sensitive and specific in comparison to dermatoscopy. This technology also allows the diagnosis of amelanotic melanoma and helps to better distinguish the limits of the tumor [4].

1.4.4 Computer-assisted diagnosis techniques

Computer-assisted diagnosis (CAD) describes a range of artificial intelligence-based techniques that automate the diagnosis of skin cancer by using a computer to analyse lesion images, and determine the likelihood of malignancy, or the need for excision. Each CAD system has a data collection component, which collects imaging or non-visual data (e.g. electrical impedance measurements) from the suspicious lesion and feeds it to the data processing component, which then performs a series of analyses to arrive at a diagnostic classification. Images are acquired using several different techniques, although most commonly by digital dermoscopy (Derm-CAD) which creates digital subsurface images of the skin using a computer coupled with a dermatoscope, a videocamera and a digital television [12].

Other systems use spectroscopy (Spectro-CAD), whereby information on cell characteristics (such as cell shape or size) is gathered by measuring how electromagnetic waves pass through skin lesions. This information is most acquired using multispectral imaging (MSI-CAD) that enables computer-generated graphic representations of lesion morphology to be produced from detecting light reflected at several wavelengths across the lesion [12].

Artificial intelligence-based algorithms have been tested in multiple reader studies for the classification of skin tumors. In the experimental setting, they showed a remarkable accuracy for melanoma diagnosis, comparable to that of experienced dermatologists. However, although numerous Artificial intelligence-based apps are available, there is no evidence on their use in the clinical practice [4].

Chapter 2

Automatic ABCD rule extraction for melanoma detection

In recent years, there has been an increasing interest in computer-aided systems for the clinical diagnosis of melanoma as support for dermatologists in different analysis steps, such as lesion boundary detection, extraction of the ABCD parameters, and classification into different types of lesions [13].

Notably, CAD systems can obtain a higher level of sensitivity for melanoma detection compared with inexperienced dermatologists [14]. Additionally, dermoscopic images, which are the most utilized in CAD, provide a more detailed insight into the morphological structures and patterns compared to the normal magnified images of skin lesions [15].

Typically, the conventional approach to medical image analysis involves a series of low-level pixel processing methods. The pipeline for melanoma detection and diagnosis comprises key processing techniques, including image pre-processing, segmentation, feature extraction, and lesion classification [16].

2.1 Pre-processing

Pre-processing addresses issues such as noise, artifacts, low contrast, and color illumination. Noise and artifacts, introduced during image acquisition, can hinder the accurate

identification of skin lesions. Examples include hair artifacts, bubbles, and blood vessels. Low contrast between the lesion and surrounding skin presents an additional challenge for accurate segmentation. Moreover, variations in color, texture, and lighting conditions in dermoscopic images can lead to multiresolution images [16].

An example of the pre-processing phase is reported in the Nadia Smaoui Zghal and Nabil Derbel [15] work, who introduced a technique consisting of three steps: filtering, morphological closing, and contrast enhancement.

The first step involves applying a denoising technique like the median filter, which effectively diminishes noise by sorting all the entries in the window and then replacing the central pixel with the middle value. Thus, noise is reduced to a certain extent. The second step is the morphological closing, which removes artifacts such as remaining hairs post-median filtering. It involves dilation to expand the image based on a specified geometry form called 'structuring element', followed by erosion to shrink the image using the same element. This process effectively enhances the image quality. The last step is contrast enhancement to have better visibility of the image details without unrealistic visual apparitions and unwanted artifacts. The authors used the 'imadjust' Matlab function to map intensity values of the image in such a way that 1% of data are saturated to high and low intensities [15].

Figure 2.1 illustrates an example of the application of these techniques.

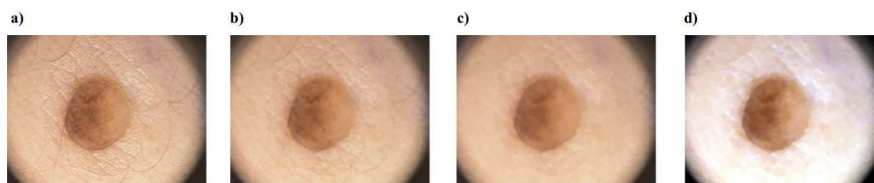


FIGURE 2.1: Pre-processing steps: (a) Original image; (b) Application of median filter; (c) Application of morphological closing; (d) Contrast enhancement [15].

The choice of parameters in image pre-processing is crucial, as it heavily influences the ultimate results, offering a more accurate assessment of various conditions or characteristics. Therefore, selecting suitable parameters is a critical aspect for any related study or analysis [15].

2.2 Segmentation

Segmentation aims to identify pixel-level and fine-grained regions of the lesion and also provides information on its location and contour. Understanding the shape of the lesion area is crucial for accurately discerning various skin diseases [17].

The segmentation process yields precise masks for each lesion. Dermatoscopic image, once converted to grayscale, inherently consists of three parts: the edges (the darkest), the lesion, and the skin (the lightest). To address segmetation, the Multi-thresholding Otsu method can be employed, allowing for the generation of multiple classes and thresholds from a single image [15].

Nadia Smaoui Zghal and Nabil Derbel [15] presented a method where the objective was to create a mask of the original image (Figure 2.2(a)), that exclusively considers the lesion. To achieve this, two binary masks were introduced: the first based on the first threshold determined by the Multi-Otsu thresholding algorithm (Figure 2.2(b)), and the second based on the second threshold (Figure 2.2(c)). Consequently, the first mask encompasses the dark part of the image along with the edges, while the second mask encompasses the lighter part. By subtracting the two masks, a substantial portion of the edges is eliminated, retaining the contour of the lesion, as depicted in Figure 2.2(d). Subsequently, the flood-fill operation was performed, starting from the center pixel. The resulting mask is presented in Figure 2.2(e). Although the lesion is successfully detected, remnants of the edges persist. Applying a simple erosion operation completely removes the edges, as demonstrated in Figure 2.2(f).

The resulting mask can be evaluated by comparing it to manually segmented regions by expert dermatologists. The comparison involves overlaying the two masks, and five metrics are then employed to assess the segmentation results: sensitivity (Equation 2.1), specificity (Equation 2.2), accuracy (Equation 2.3), Dice coefficient (Equation 2.4) and Jacard coefficient (Equation 2.5) [18].

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2.1)$$

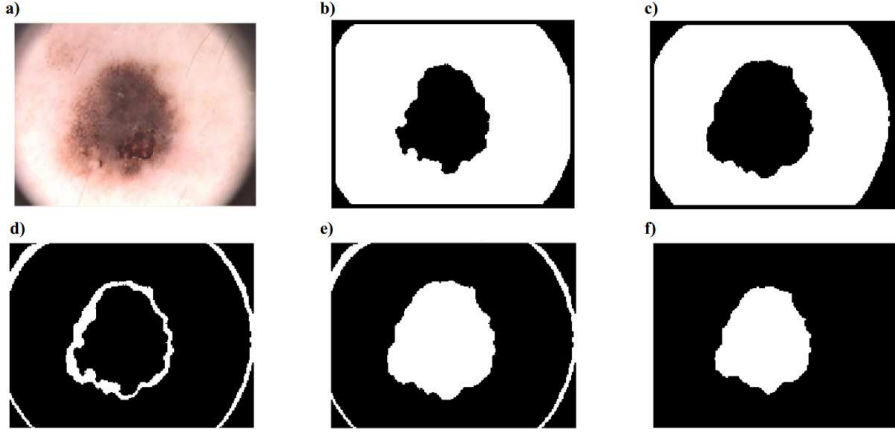


FIGURE 2.2: Definition of the mask: (a) Original image; (b) Mask1 (with the first threshold); (c) Mask2 (with the second threshold); (d) Mask2- Mask1; (e) Application of filling operator; (f) Final Mask [15].

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (2.2)$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (2.3)$$

$$\text{Dice} = \frac{\text{TP}}{\text{TP} + (\text{FP} + \text{FN})/2} \quad (2.4)$$

$$\text{Jacard} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \quad (2.5)$$

In these formulas TP is the true positive and TN is the true negative. FP is the false positive and corresponds to the pixels that have been segmented by the algorithm but not by the dermatologist. FN is the false negative and corresponds to missed pixels, i.e. those that are considered as skin lesion by the dermatologists but not from the automatic segmentation [18]. The mask resulting from segmentation is then used to extract the features of the skin lesion.

The method outlined above serves as just one example among various segmentation techniques documented in the literature. Over the years, alternative approaches based

on edge-detection, region-based methods, clustering, active contours, and artificial intelligence have been developed [19].

2.3 ABCD feature extraction

In the pursuit of an accurate diagnosis of melanoma, several methodologies have been devised to automate the ABCD rule, with the aim of providing objective assessments of these crucial features. The following section outlines the main methodologies developed for each of these features. Collectively, these studies contribute to the advancement of skin lesion analysis by offering a range of specialised methods adapted to different lesion characteristics.

2.3.1 Asymmetry

Asymmetry refers to the fact that when drawing a line across the middle of the mole, the two halves will not match, that is, the shape of one half does not match the other half, providing a warning sign of melanoma [14]. Numerous methodologies have emerged to assess this parameter, focusing on different aspects of lesion characteristics and shape. Researchers have explored geometric measures such as Compactness Index, also known as Circularity (calculated as $4\pi\frac{A}{P^2}$, with A representing the lesion's area and P its perimeter), and Symmetry Distance, which quantifies the minimum movement needed to transform the shape into its closest symmetric form [20]. A simple asymmetry index can be calculated from the smallest difference between the image area of the lesion and the image of the lesion reflected from the principal axis [21]. Additionally, some methodologies utilize the lesion's centroid for partitioning and assessment, while others rely on distance-based analysis to calculate distances between selected features. Some approaches integrate both shape and texture for asymmetry estimation. For instance, shape asymmetry involves computing symmetry axes through the Hough transform algorithm, assessing asymmetry based on overlapping pixels after rotating the lesion around these axes. Texture asymmetry, on the other hand, computes average quadratic errors between overlapping pixel intensities after such rotations [22]. Introducing the 'bi-fold method', which involves folding the lesion outline and measuring the non-overlapping

region, offers another perspective on asymmetry [23]. Furthermore, certain methodologies delve into color, brightness, and shape asymmetry based on specific criteria [24].

These different approaches represent only a selection of methods found in the literature. Collectively, they contribute to an improved analysis of the asymmetry of skin lesions [14].

2.3.2 Border Irregularity

Border irregularities can be divided into texture irregularities and structure irregularities. Texture irregularities are the fine variations along the border of the lesion; detection and measurement may be subject to the noise resulting from the image acquisition method. In contrast, structure irregularities, which are general undulations of the perimeter, may indicate abnormal histological signs and have a higher correlation with melanomas. Therefore, the accurate measurement of structure irregularities is important for the diagnosis of melanomas [25].

Figure 2.3 shows both types of irregularities, illustrated in three boundary contours extracted from pigmented skin lesions. The border of lesion A presents an absence of structural protrusions and indentations but shows a multitude of texture irregularities. The border of lesion B has a structural protrusion at the top of the border but shows fewer texture irregularities than the other two borders. Finally, lesion border C has a prominent structural protrusion and indentation at the lower part of the border.

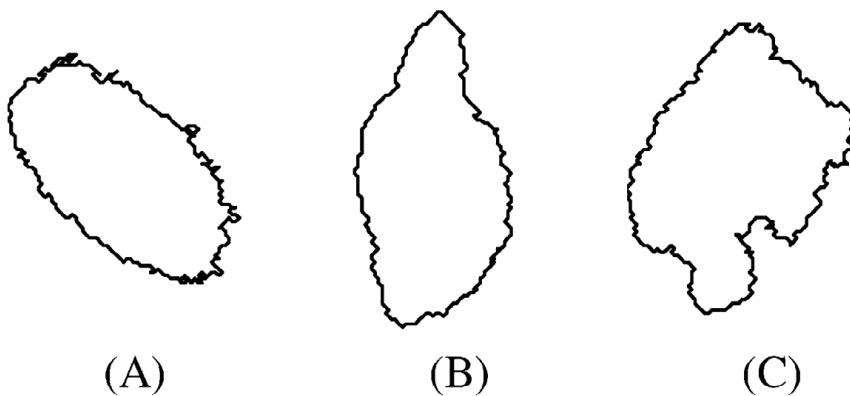


FIGURE 2.3: Types of border irregularities [25].

Various methodologies have been proposed to assess irregularities in skin lesion borders. An example is the combination of Circularity with Radial Distance, Fractal dimension and extraction of small changes in the contour. In lesions with irregular borders, there is a high variance in the radial distance distribution, which justifies the evaluation of the radial distance between the center of the mass and the border [13]. Additionally, a border function, determined by the distance between the lesion border and the image edge, can be employed to assess irregularities by detecting turning points in the function [26]. Fractal dimension has also been employed since it can be used to quantify the roughness or smoothness of a curve in a given space, such as the boundary of a mole in a picture [27]. Furthermore, an algorithm employing a measure called 'sigma-ratio', which is based on the number of Gaussian smoothing iterations required for eliminating all concavities along the lesion border, has been proposed [25]. Other methods explore wavelet decomposition and neural network-based classifiers to classify melanocytic lesions based on border irregularity features [28]. In addition to these methodologies, another approach has been proposed and it involves dividing the lesion into eight equal slices, approximating each slice's sub-contour with a third-order spline function; irregularity is then determined by comparing the fitting error to the length of the sub-contour [24].

These diverse approaches reflect ongoing efforts to provide reliable and accurate methods for border irregularity assessment in dermatology [14].

2.3.3 Color

Color variegation in skin lesions refers to the presence of multiple pigment shades or colors within the lesion's border. Melanoma lesions often exhibit more than two colors, while benign lesions tend to have a uniform coloration. Suspicious shades of color in melanoma typically include white, red, light brown, dark brown, blue-gray, and black. Several techniques have been proposed to evaluate color variegation in skin tumors. These methods employ various approaches, including color segmentation algorithms, unsupervised learning techniques, and analyses of color components [14].

For instance, one unsupervised learning method uses K-means to classify the pixels of the skin lesion, with the number of clusters set to 4, based on the number of colors

typically detected by dermatologists. Setting the number of clusters equal to 4 is supported by the observation that dermatologists rarely identify more than 4 colors in most cases [22].

Another method involves measuring the Euclidean distance between each pixel of the skin lesion and the six suspicious colors of melanoma (white, black, light brown, dark brown, red, and blue-gray). A pixel is considered to belong to a specific color if the distance is less than a threshold. A lesion is classified as containing a suspicious color if the pixels belonging to that color represent more than 5% of the skin lesion pixels [24].

2.3.4 Diameter

Melanomas typically exhibit a larger diameter compared to benign moles, often exceeding 6 mm. Various methods have been proposed to measure the diameter of skin lesions.

One approach involves measuring the horizontal and vertical dimensions of the lesion and then determining the diameter as the maximum of these two dimensions [29]. Another method employs the dimensions of the minimum area rectangle enclosing the lesion, from which the diameter is calculated using a conversion factor from pixels to millimeters and the factor is found using the imaging system parameters [30]. Other techniques rely on specialized measurements like Feret's diameter, involving the distance between parallel tangents at the lesion's contour [31], or the semi-major axis of a best-fit ellipse [32].

Generally, the diameter value is initially obtained in pixels and then converted to millimeters using known parameters of the imaging system. It's crucial to note that accurate segmentation of the lesion is a prerequisite for precise diameter measurement. Additionally, a common challenge lies in determining the correct conversion factor to translate measurements from pixels to millimeters, which hinges on the original size of the real-world image information not always available, particularly in online image datasets [14].

2.4 Classification based on ABCD rule

The classification of lesions as symmetric or asymmetric primarily employs threshold or machine learning approaches, including Support Vector Machine (SVM) and decision trees. These classifiers are trained using extracted measures, and the learned models are then applied to classify new or test lesions. When it comes to classifying skin lesions with irregular borders, some studies rely on threshold-like measures, while others employ machine learning approaches. Neural networks, CNNs, Gaussian naive Bayes, and fuzzy neural networks are also frequently used in these cases [19].

Combining ABCD features, rather than using them individually, enhances the accuracy, sensitivity, and specificity of melanoma detection. This principle extends to combining ABCD features with other characteristics derived from the skin lesion. Recent literature has introduced data-driven features obtained from deep learning methods. However, the main drawback is that these techniques may struggle to detect fine structures, such as those outlined in the ABCD rule [14].

Automated methods for extracting ABCD features offer a crucial advantage in providing an unbiased secondary assessment, reducing potential subjectivity compared to physician evaluation, especially in capturing the finer details of skin lesions. However, a significant drawback lies in their dependency on accurate segmentation, which may be compromised by the presence of artifacts or noise, impacting the feature extraction process. Additionally, machine learning-based approaches face challenges in data availability, particularly for individual features like A and B. Furthermore, these methods may fall short in representing real-world measures, as seen in diameter measurement [14].

2.5 Example of feature combination and classification

A straightforward method for detecting and classifying skin lesions using dermoscopy images based on ABCD rules has been presented by Nadia Smaoui Zghal et al. [15]. They considered asymmetry (A) in terms of form and color. To assess form asymmetry,

the lesion is centered and aligned along its major axis. The image is then divided based on the minor and major axes. Asymmetry is measured by folding the lesion's outline along these axes and calculating non-overlapping areas. The asymmetry score is calculated as:

$$\delta A = \begin{cases} 0 & \text{if } OVL(L, R) \leq T_0 \text{ and } OVL(U, d) \leq T_0 \\ 1 & \text{if } OVL(L, R) \leq T_0 \text{ or } OVL(U, d) \leq T_0 \\ 2 & \text{otherwise} \end{cases} \quad (2.6)$$

where T_0 is an overlapping threshold value set to 5%, $OVL(L,R)$ is the overlapping area between the left and right half of the image, $OVL(U,d)$ is the overlapping area between the upper and lower (down) half of the image. Form-based asymmetry is then determined by dividing the asymmetry score by the total lesion area. To assess color-based asymmetry the histograms are created for the RGB components of each part of the lesion, which are then normalized. Subsequently, the distance between the normalized histograms is calculated using the Chi-Square distance. This value represents the variation in color within the image. If the lesion is of a single color, the Chi-Square will be 0. Finally, the asymmetry score (A) is determined as the sum of form-based asymmetry and color-based asymmetry.

To calculate border irregularity (B), lesions are divided into eight sectors. Subsequently, the standard deviation of the pixel value is computed for each sector. If the standard deviation of a sector exceeds a predefined threshold, it is assigned a border score of 1. Consequently, the maximum irregular border score can reach up to 8, while the minimum is 0.

The color (C) parameter is defined by checking for the presence of the colors: white, red, light brown, dark brown, blue-gray, and black on the image showing the lesion. To do this, the difference between the color of the pixel in the lesion and the reference for each color is calculated. If this difference is equal to or less than a certain predetermined value, then the color score increases by 1.

The diameter (D) is calculated as the number of pixels in the length of the major axis (the largest diameter) of the lesion and transformed to a millimeter scale. The diameter

score is then calculated as:

$$\left\{ \begin{array}{l} D = 0.5 \quad \text{if } M \leq 2 \text{ mm} \\ D = 1 \quad \text{if } M \leq 3 \text{ mm} \\ \vdots \\ D = 4.5 \quad \text{if } M \leq 10 \text{ mm} \\ D = 5 \quad \text{otherwise} \end{array} \right. \quad (2.7)$$

Each of these ABCD parameters is multiplied by a given weight factor, summarized in Table 2.1, to calculate the Total Dermatoscopy Value (TDV), reported in the Equation 2.8, and gives a clear idea about the lesion state [15].

Criteria	Score	Weight Factor
Asymmetry	0-2	1.3
Border	0-8	0.1
Color	0-6	0.5
Diameter	0-5	0.5

TABLE 2.1: Presentation of criteria of ABCD rules

$$\text{TVD} = 1.3 \cdot A + 0.1 \cdot B + 0.5 \cdot C + 0.5 \cdot D \quad (2.8)$$

The result of TDV shows whether the lesion is benign or malignant. A TDV less than 4.75 indicates a benign lesion, a TDV between 4.8 and 5.45 indicates a suspicious lesion and a TDV larger than 5.45 indicates a malignant melanocytic lesion. The proposed method classifies images with a specificity of 92% and a sensitivity of 87%, which reflects its reliability [15].

Chapter 3

Device for macroscopic image acquisition, 3D modeling and cropping of skin lesions

The manual examination of dermatologists is often monotonous, time-consuming, and is prone to subjectivity. Furthermore, the accuracy of manual assessments can be influenced by the reviewer's level of experience and workload, as discussed in Chapter 1. Despite attempts to mitigate these limitations through CAD systems, the reliance on manually acquired images of individual moles persists. This procedure necessitates framing the mole with a dermatoscope, employing a lens-fitted camera to capture high-resolution images in direct contact with the skin, and utilizing specialized lighting, such as polarized light, to discern intricate details of lesions while applying a gel to the skin surface to minimize reflections and enhance light transmission.

It's noteworthy that the images commonly utilized in dermatology literature stem from this methodology, where dermatoscopic acquisition ensures the highest resolution possible, allowing for a meticulous examination of skin lesions. However, this thesis introduces a groundbreaking imaging device for skin assessment, representing a departure from conventional methods, that can be combined with traditional dermatoscopic acquisition with benefit especially in the melanoma screening.

The device employed allows a 3D mapping of the patient body, potentially the total

body, combined with the automated identification and analysis of every single mole in the skin surface.

3.1 Device description

The device used for image acquisition is protected by Italian Patent n° 102016000132357 [33] and a first prototype is present at the 'Istituto Oncologico Veneto' (IOV) in Padua. It consists of a vertical structure with a horizontal movable arm containing 12 cameras, one thermal camera and 4 polarized led lights. All 12 cameras shoot simultaneously so that the movements of the patient did not influence the acquisition of images. Figure 3.1 shows the structure of the device. The black dots indicate the 12 cameras. The red dots indicate the position of the thermal imaging camera. Finally, the yellow dots indicate the polarized LED lights.

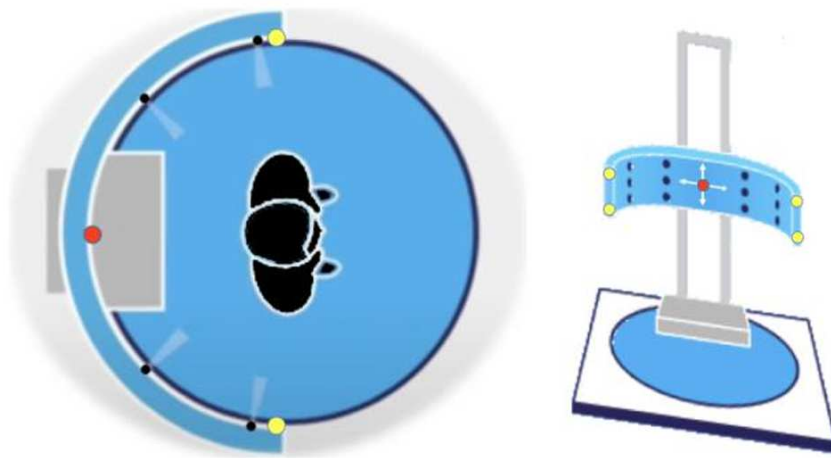


FIGURE 3.1: Structure of the device.

In more detail (Figure 3.2 (a)), the structure has three half-moon shaped arms that are fixed in the central part on the vertical support. The 12 cameras are Nikon D3500 and are placed four on each arm. The thermal camera is located in the center of the structure. In addition, the prototype has a manual control that allows the vertical sliding of the three arms, so that the frame can be adapted to the height of the patient. The implemented prototype enables the acquisition of images of the patient's entire

back, with the patient positioned in the correct posture depicted in Figure 3.2(b), and their back uncovered.

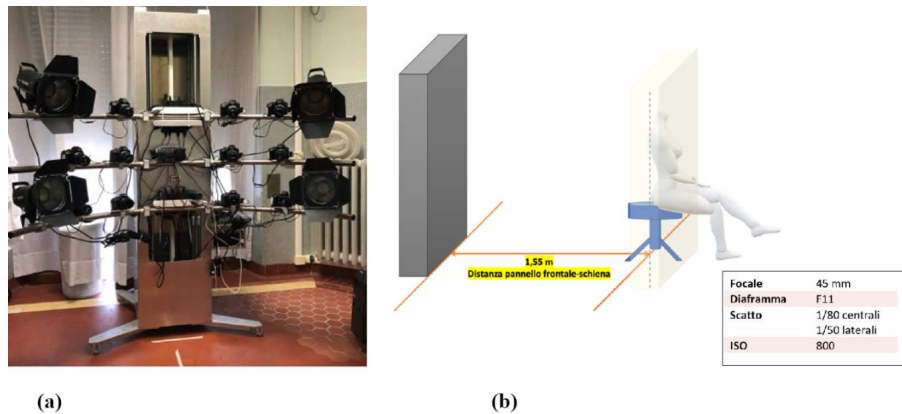


FIGURE 3.2: Device prototype present at the IOV: (a) Device structure; (b) Correct patient position.

3.2 3D Model and skin lesion crop

In order to capture the images, the patient is asked to expose their back and position themselves facing away from the medical device. When the patient has a suspicious lesion in a part of the body different from the back, an image of that specific area will be acquired.

Image acquisition is initiated through the DigiCamControl software which allows simultaneous shooting. The images captured by the 12 cameras display the patient's back (or other parts of the body) from various angles, as depicted in Figure 3.3.

For a subsequent validation of the automatic melanoma recognition algorithm, it is necessary to indicate the suspicious lesions on the patient. To do so, an additional acquisition of the back or the specific area of concern is performed with the suspicious lesion marked by an arrow drawn on a label placed near this area. Two images, in which the suspicious lesion is marked by an arrow, can be observed in Figure 3.4.

From the 12 images of the back a 3D model is reconstructed using the Structure from Motion (SfM) Photogrammetry technique. This procedure finds common key points between images, aligns the photos, creates a mesh, and then textures it.

One advantage of 3D modeling is that the wide representation of a remarkable body

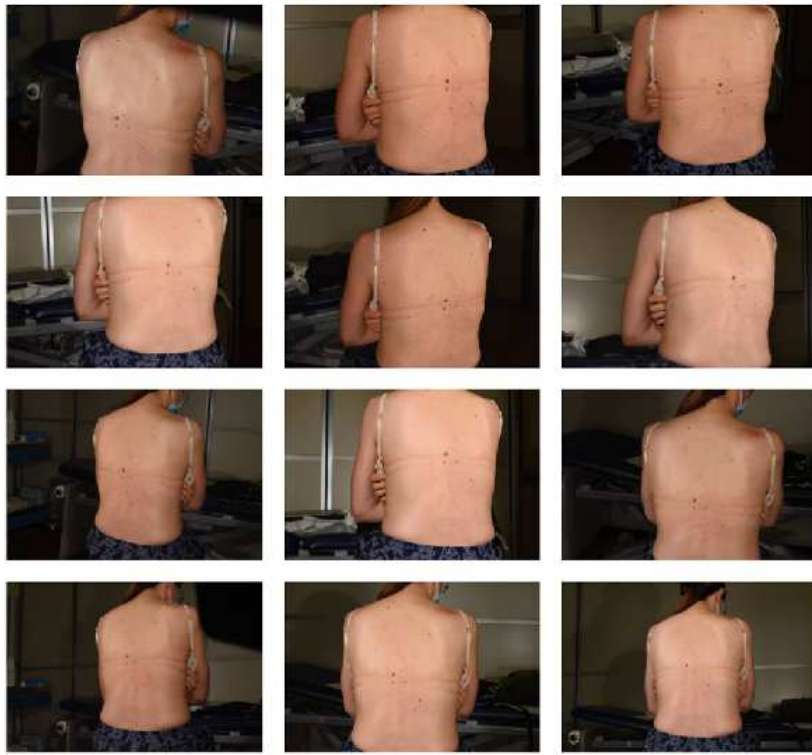


FIGURE 3.3: 12 patient's back images from various angles.

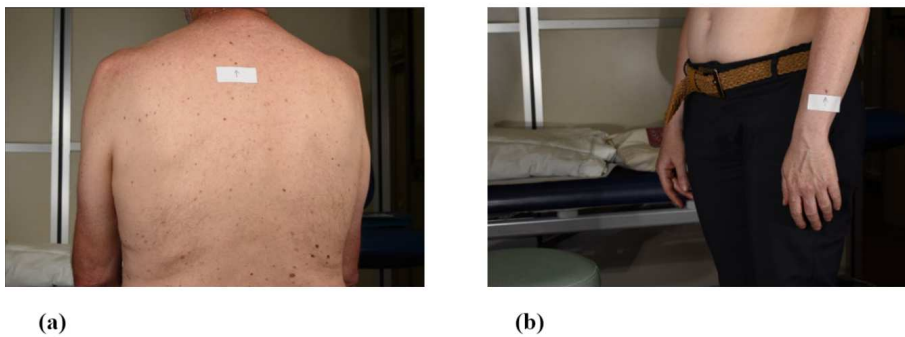


FIGURE 3.4: Suspicious lesion marked by an arrow: (a) suspicious lesion on the back; (b) suspicious lesion on the forearm.

area facilitates the identification of lesions that are distinguished from others, thus revealing one of the main clinical characteristics indicative of melanoma, commonly known as the ugly duckling sign.

From the obtained 3D model, images are extracted that allow for the best visualization of the lesions, specifically derived from the model's optimal orientation. These extracted images are referred to as 2.5D images. For each 2.5D image, the HSV (Hue, Saturation, Value) color space is used. The locations of the mole within the skin area are identified by examining the intensity of the S-channel. After identifying each mole on

the S-channel as the local maximum within the region of interest, a cropped image from the original is generated. This crop is done in such a way that each mole corresponds to an image in which about 30% of the pixels belong to the mole.

Figure 3.5 shows one of the 12 images of a patient's back and its representation with the S-channel. In the figure there are also examples of cropped images taken from the original image.

Three examples of cropped images in their actual dimensions are shown in Figure 3.6; Figures (a) and (b) represent benign moles while Figure (c) represents a lesion defined as suspicious by the dermatologist, which has been successively subjected to biopsy.

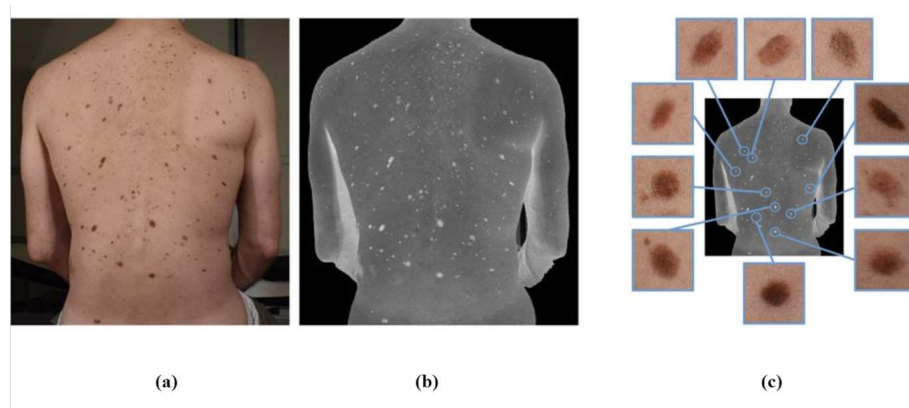


FIGURE 3.5: Moles identification: (a) One back image of the patient; (b) S-channel of the image; (c) Example of crop images.

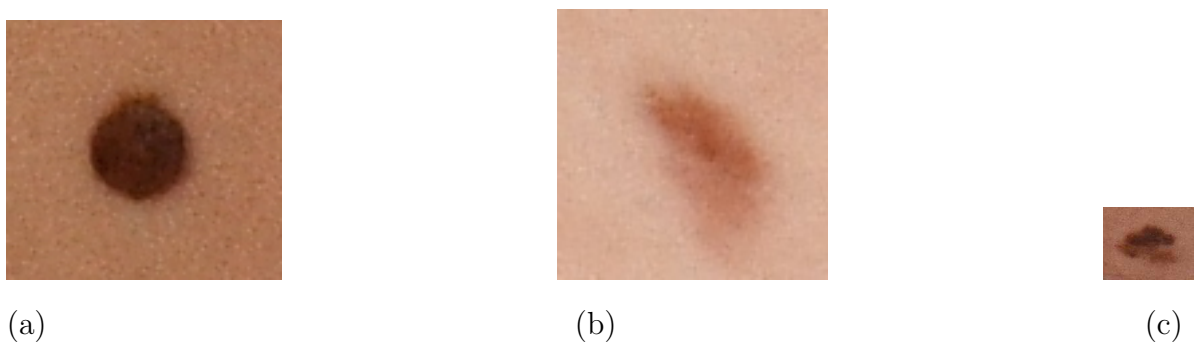


FIGURE 3.6: Crop images: (a) benign mole; (b) benign mole; (c) suspected mole.

3.3 Dataset description

All patients, who have participated in the study, after being properly informed about the purpose, methods and implications of the study, have signed the informed consent, approved by the local Ethics Committee.

The images used in this thesis study were acquired, using the described medical device, from 78 patients who underwent a check-up at the IOV institute in Padua. From these patients 86 lesions were defined as suspicious by the dermatologists and were consequently biopsied. The histological results reported the following data:

- 9 melanomas;
- 3 nevi with low-grade and focal high-grade dysplasia;
- 27 nevi with low-grade dysplasia;
- 4 basal cell carcinoma
- 43 nevi or other lesions.

Nevi with dysplasia cannot be classified as false positives as their removal is highly recommended due to their potential risk of giving rise to melanomas.

Therefore, 39 lesions i.e. the 9 melanomas and 30 nevi with dysplasia are considered true positives. The 47 lesions, on the other hand, represent false positives as far as melanoma identification is concerned. The histological results confirm that several issues arise, for dermatologists, in discriminating between benign and malignant lesions. In particular, the percentage of false positives is relatively high.

This thesis incorporates three diverse sets of images.

One set of images pertains to the manual segmentation of the 86 suspicious lesion crops. The segmentation of these images was performed by a single dermatologist.

Another dataset consists of images of the crops of the 86 suspicious lesions subjected to automatic pre-processing and segmentation. These two procedures are not the focus of this thesis, but have been used to evaluate the application of algorithms for the quantitative determination of clinical parameters, detailed in Chapter 4, in a fully automated

context.

Similarly, an evaluation of clinical features was conducted on a dataset comprising 189 benign lesions, derived from the 78 patients and processed using the automatic algorithm. This additional analysis provided the opportunity to compare the quantitative values obtained for suspicious and benign lesions.

Chapter 4

Quantitative evaluation of asymmetry, border irregularity and color

The objective of this thesis is to conduct a quantitative assessment of asymmetry, border irregularity and color parameters. The analysis was conducted in several steps. Initially, two sets of dummy images were generated, one tailored for assessing asymmetry and the other for evaluating border irregularity. These sets were designed following a study of the HAM10000 dataset [34], which includes a collection of dermatoscopic images of crucial diagnostic categories within the domain of pigmented lesions.

Subsequently, a set of functions for analysing skin lesion images was developed using the Python programming language. These functions were implemented using specialized libraries such as 'OpenCV' and 'NumPy' for image manipulation and processing, and 'Matplotlib' for result visualization. They were then tested on the dummy images to ensure their accuracy and effectiveness.

Finally, the code was applied to cropped images sourced from the dataset outlined in Chapter 3, demonstrating its applicability and reliability in a real-world scenario.

Color parameter evaluation was conducted through the analysis of masks generated by combining the original lesion image with its respective segmentation mask. Subsequently, a targeted algorithm was developed to detect color based on these processed images.

Image ID	Typology	Diagnosis method
ISIC 0024516	Melanoma	Histopathology
ISIC 0024617	Melanocytic Nevi	Follow up
ISIC 0024929	Melanoma	Histopathology
ISIC 0024940	Melanoma	Histopathology
ISIC 0025439	Melanoma	Histopathology
ISIC 0026551	Melanocytic Nevi	Histopathology
ISIC 0027277	Melanoma	Histopathology
ISIC 0027300	Melanoma	Histopathology
ISIC 0027302	Melanoma	Histopathology

TABLE 4.1: Image categories of Figure 4.1 [34].

4.1 Dummy images

Dummy images were created to build two datasets in which a gradual increase in asymmetry or irregularity of the border of the represented object could be visually observed. The main objective was to obtain a series of images ranging from cases of perfectly round/oval lesions (typical of benign lesions) to cases of malignant lesions characterized by a high degree of irregularity at the border or strong asymmetry. These images are represented as black and white masks, where the object of interest is depicted in white (see Figures 4.2, 4.3).

In order to make them comparable to real cases of skin lesions that may exhibit marked asymmetry or irregularity at the border, the HAM10000 image dataset [34], which contains a wide collection of images of various types of skin lesions, was taken as a reference.

Figure 4.1 displays some examples of lesions characterized by high irregularity in the border and asymmetry, selected from the HAM10000 dataset. Additionally, Table 4.1 provides detailed information about the type of lesion depicted in the image and the method used to confirm the diagnosis.

The dummy images created for the quantitative assessment of border irregularity are presented in Figure 4.2, while the images used for the quantitative analysis of asymmetry are shown in Figure 4.3.

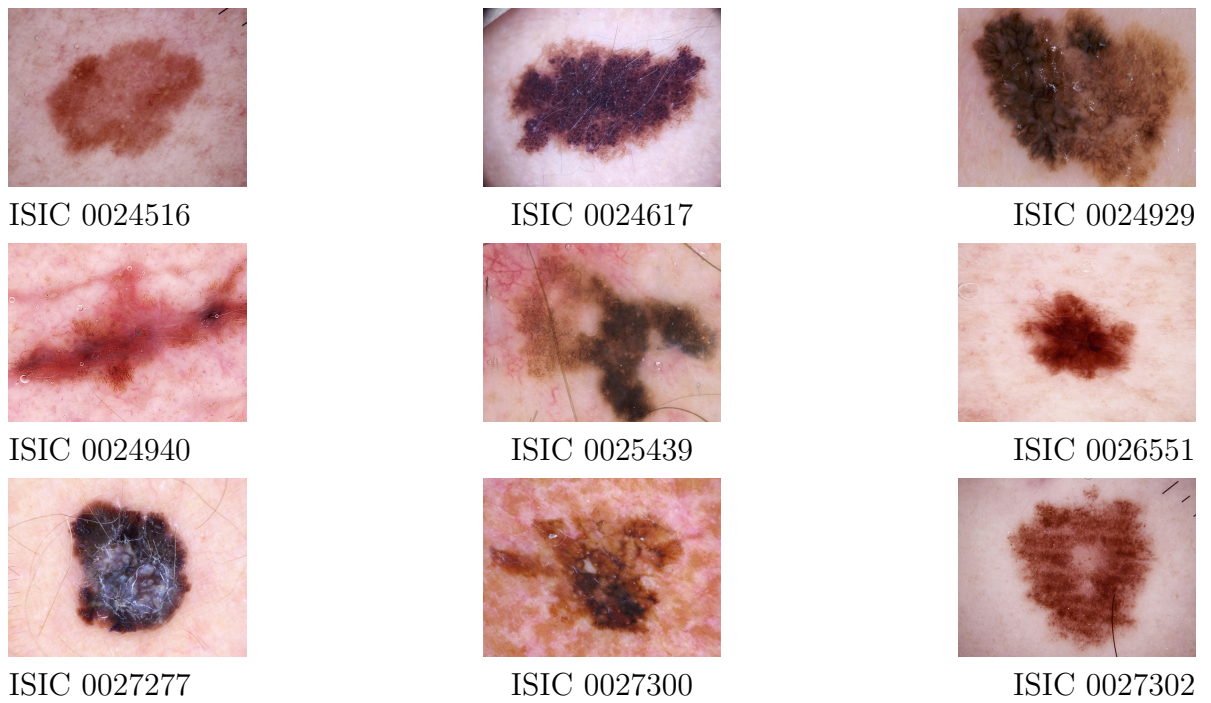


FIGURE 4.1: Some reference images from HAM10000 dataset [34].

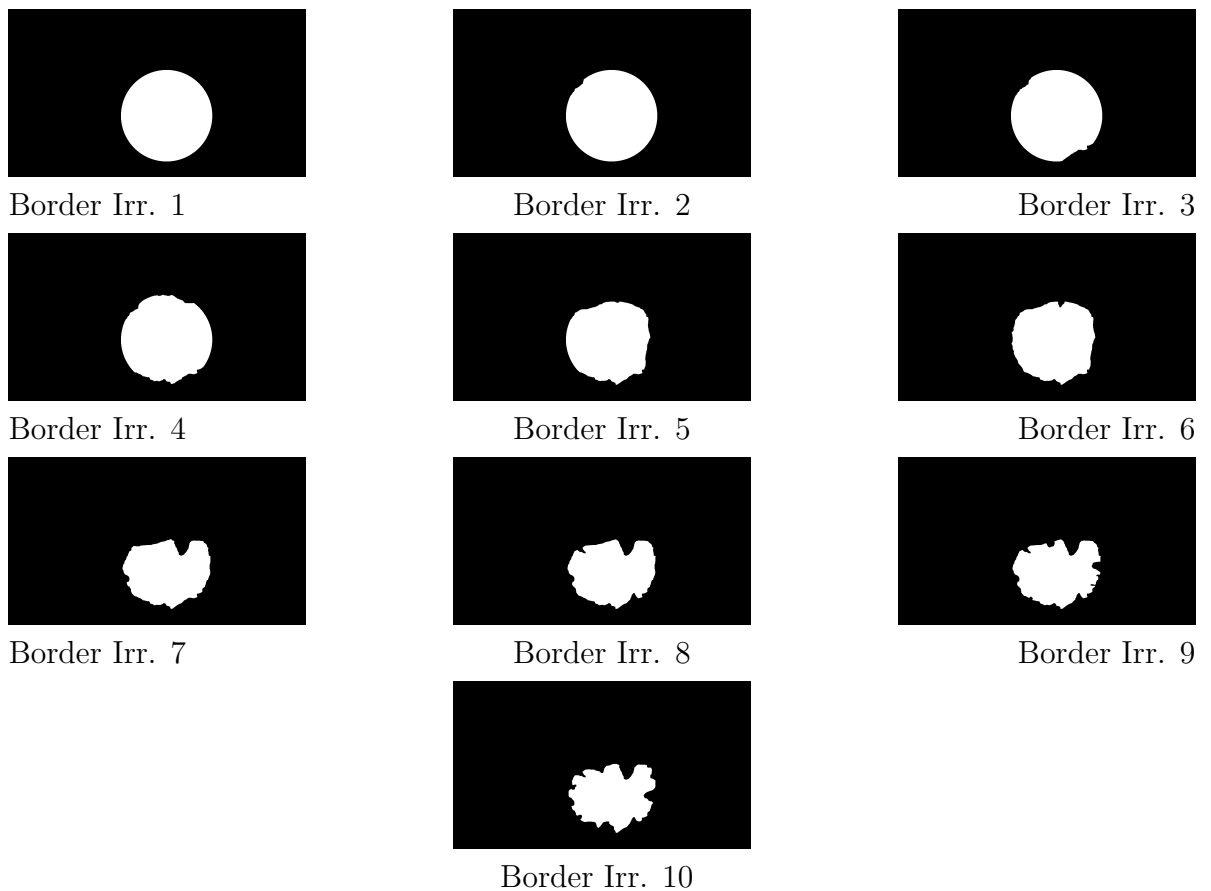


FIGURE 4.2: Dummy images for border irregularity evaluation.

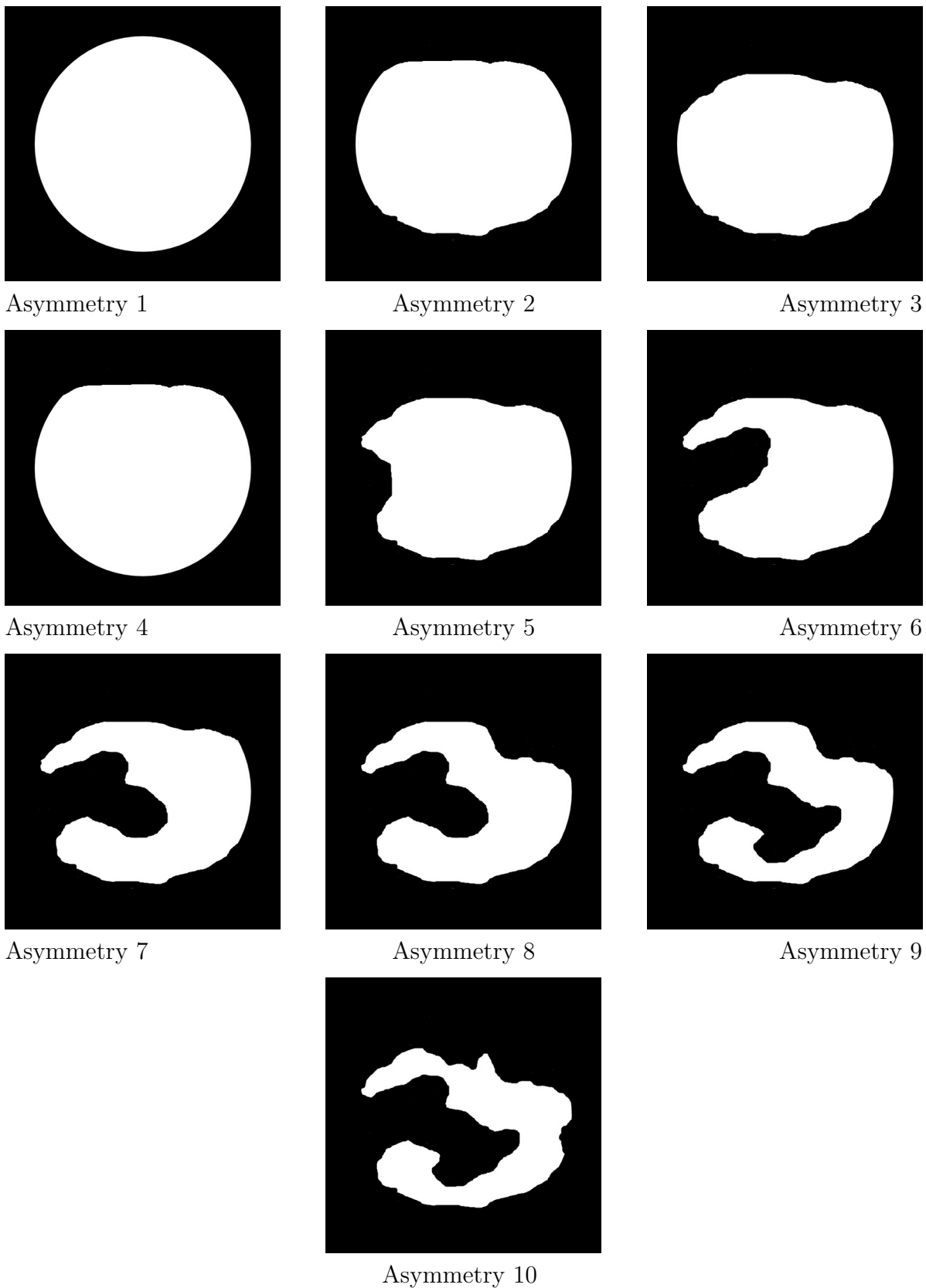


FIGURE 4.3: Dummy images for asymmetry evaluation.

4.2 Asymmetry

Throughout the course of this research, a set of functions has been developed to analyze the asymmetry of skin lesion images. The image processing pipeline begins with the application of geometric transformations to rectify the position and orientation of the object. Subsequently, the image is partitioned along the main axes to compute the overlapping area.

The asymmetry of the lesion is determined in terms of overlapping areas. Moreover, insights into the symmetry and shape of the lesion are provided through the calculation of the percentage of non-overlapping area relative to the total lesion area.

4.2.1 Translation and rotation

The translation and rotation of the object in the image aim to align the object's major axes with the x and y axes of the image, assuming that the origin of the image is at its center. Additionally, through these two operations, the centroid of the object will be precisely located at the center of the image. These operations are valuable as they significantly facilitate the subsequent calculation of the overlap area.

Initially, the contour of the object is determined and then the position of the centroid is calculated using the zeroth and first-order geometric moments. The geometric moment is defined as:

$$m_{pq} = \sum_x \sum_y x^p \cdot y^q \cdot f(x, y) \quad (4.1)$$

where p and q define the order of the geometric moment and $f(x, y)$ is a discrete function that corresponds to the object of the image.

The zeroth and first-order geometric moments provide information about how points are distributed within the object. The zeroth order moment corresponds to the area of the object while the first-order moments identify the coordinates of the centroid of the object.

The zeroth-order geometric moment, denoted as m_{00} , is calculated as:

$$m_{00} = \sum_x \sum_y x \cdot y \cdot f(x, y) \quad (4.2)$$

The first-order geometric moment with respect to the x-axis, denoted as m_{10} , is calculated as:

$$m_{10} = \sum_x \sum_y x \cdot f(x, y) \quad (4.3)$$

The first-order geometric moment with respect to the y-axis, denoted as m_{01} , is calculated as:

$$m_{01} = \sum_x \sum_y y \cdot f(x, y) \quad (4.4)$$

Dividing these moments by the zeroth-order geometric moment m_{00} provides the centroid coordinates x and y :

$$\text{Centroid}(x, y) = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (4.5)$$

Subsequently, the distance between the centroid and the center of the image is computed. This involves defining $shift_x$ as the pixel count along the x-axis and $shift_y$ along the y-axis. These measures determine the necessary translation along the x and y axes to align the centroid with the image center.

Following this, a 2x3 affine transformation matrix is created. The components $[1, 0, shift_x]$ indicate translation along the x-axis, while $[0, 1, shift_y]$ indicate translation along the y-axis.

The affine transformation is applied to the binarized image. This shift ensures that the object's centroid is precisely positioned at the center of the image.

Subsequently, the best-fitting ellipse for the specific contour of the translated object is determined. This process also yields the angle of rotation of the ellipse relative to the x-axis, signifying the orientation of the object.

A 2x3 rotation matrix is then generated, considering both the centroid and the rotation angle. This matrix will be utilized to rotate the image.

As a result of these operations, the object is now centred within the image, with its

major axis perfectly aligned with the image's x-axis.

Figure 4.4 displays two examples of this process applied to dummy images, while Figure 4.5 demonstrates its application on masks derived from images of suspicious lesions segmented by a dermatologist.

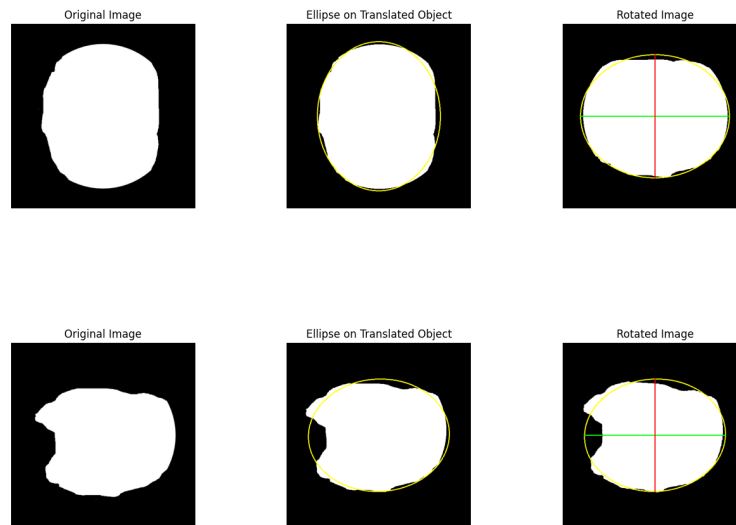


FIGURE 4.4: Translation and rotation of dummy images.

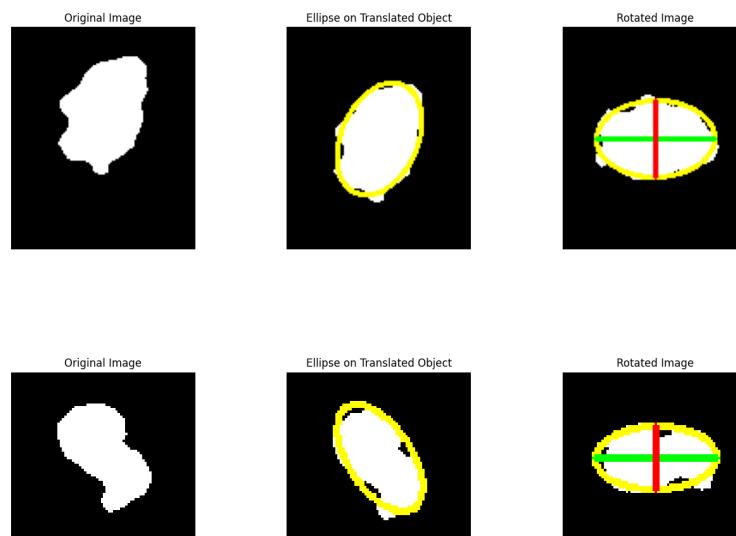


FIGURE 4.5: Translation and rotation of suspicious lesion images.

4.2.2 Non-Overlapping area

Asymmetry is determined by assessing the non-overlapping area between the upper and lower parts, as well as between the left and right parts of the object defined with respect to the main axes.

This process begins by dividing the image along the vertical axis, followed by mirroring the left part of the object. The two halves are then superimposed. The same procedure is performed by dividing the object along the horizontal axis. In this case, the lower part is mirrored and overlaid with the upper part.

Figure 4.6 presents a dummy image after the translation and rotation and two images corresponding to the two halves of the object, the upper one and the mirrored lower one, while Figure 4.7 presents the right half and the left mirrored one.

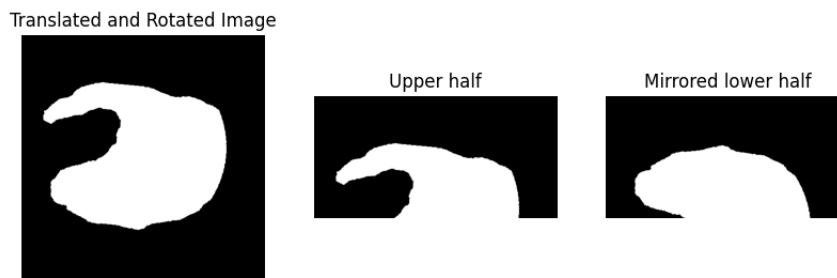


FIGURE 4.6: Division of the object based on the horizontal axis.

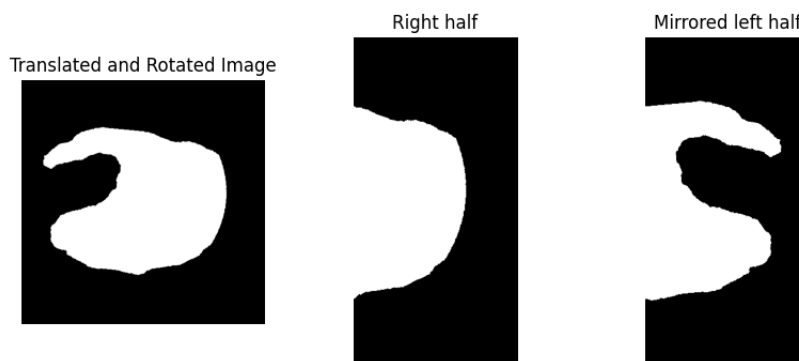


FIGURE 4.7: Division of the object based on the vertical axis.

To quantify the non-overlapping area, the algorithm employs bitwise operations. Specifically, it performs logical OR and AND operations on the two parts of the object. The OR operation combines the pixel values from both parts, while the AND operation isolates the overlapping regions. The non-overlapping area is obtained by taking the difference between the results of the OR and AND operations. This area represents the portion of the object where the two halves do not overlap. Figures 4.8 and 4.9 show an example of OR image, AND image and the result of the subtraction of these two images.



FIGURE 4.8: Non-Overlapping area over horizontal axis.

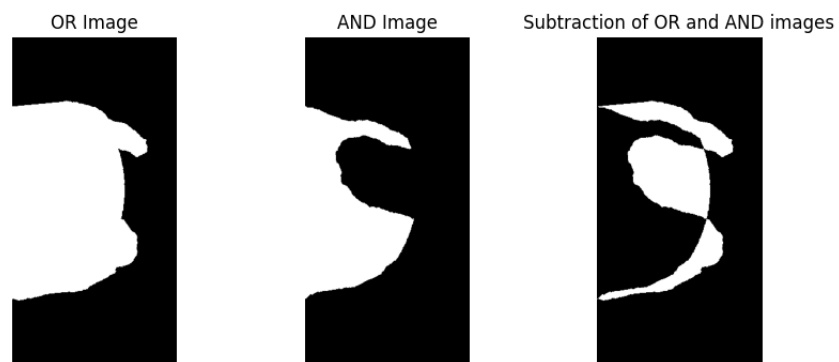


FIGURE 4.9: Non-Overlapping area over vertical axis.

The non-overlapping area is then calculated by counting the pixels of this area. Figure 4.10 shows the non-overlapping areas of a dummy image reported with units in pixels.

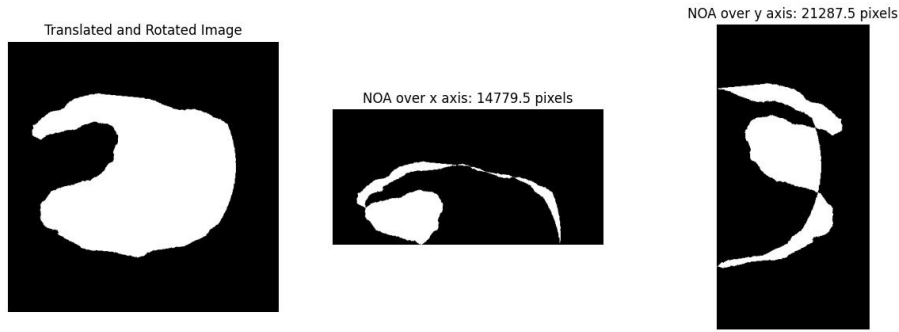


FIGURE 4.10: Non-Overlapping area (NOA) defined in pixels.

4.2.3 Asymmetry scores

The values considered to quantitatively assess the asymmetry of the lesions are the percentage of non-overlapping area and the definition of an asymmetry index that determines along how many axes the object is asymmetric. High values of these indexes denote high asymmetry of the skin lesion.

- Percentage of non-overlapping area (%NOA): the calculation of this parameter is based on the values of non-overlapping area obtained in the previous step, namely non-overlapping area along the x-axis (NOA_x) and along the y-axis (NOA_y). The sum of these two values is divided by 2, as otherwise the total area of the object would be considered twice. At this point, the percentage of this area relative to the total area of the lesions is calculated.
- Asymmetry Index (AI): this index indicates along how many axes the object is asymmetric. It is based on the threshold T_0 , which is 5% of the total area. The index is set to 0 if both NOA_x and NOA_y are less than or equal to this threshold. It is set to 1 if only one of the two is less than or equal to T_0 . If both are greater than T_0 , the index is set to 2.

$$AI = \begin{cases} 0 & \text{if } NOA_x \leq T_0 \text{ and } NOA_y \leq T_0 \\ 1 & \text{if } NOA_x \leq T_0 \text{ or } NOA_y \leq T_0 \\ 2 & \text{if } NOA_x > T_0 \text{ and } NOA_y > T_0 \end{cases} \quad (4.6)$$

4.3 Border irregularity

To quantitatively evaluate border irregularity, four distinct methods were explored. The first method relies on assessing the standard deviation of center-of-mass to border distances (see section 4.3.1). The second method makes use of the compactness index (detailed in section 4.3.2). The third method uses the Convex Hull to identify indentations or expansions of the border (see section 4.3.3), while the fourth method involves analyzing the standard deviation of ellipse-to-border distances (explained in section 4.3.4).

4.3.1 Standard deviation of center-of-mass to border distances

This method, developed for the quantitative assessment of the irregularity of the border of skin lesions is based on evaluating the standard deviation of distances between the center of mass of the lesion and the contour points of the lesion.

As an initial step, a translation (as described in the previous section) of the object is performed on the image to align its centroid with the center of the image. This step is crucial to ensure accurate and centered measurements. Once translated, the object is partitioned into eight sections, each corresponding to a segment of the complete circle. This partition is achieved by constructing eight binary masks, each delineating an angular sector around the centroid, using a series of predetermined angles spanning from 0 to 360 degrees. This division allows for a detailed and oriented analysis of the border irregularity. Figure 4.11 shows a representation of the eight binary masks used in the thesis.

Subsequently, an empty image with the same format as the translated one is created. This image is used to draw the contours of the lesion. A bitwise AND operation is then applied between the border image and the binary sector mask, isolating only the border belonging to that angular region. The contour points within this region are then identified through a process of extracting the white pixels.

An example of a binary mask of a suspicious lesion, obtained after manual segmentation by a dermatologist, and its border, is shown in Figure 4.12, while the subdivisions

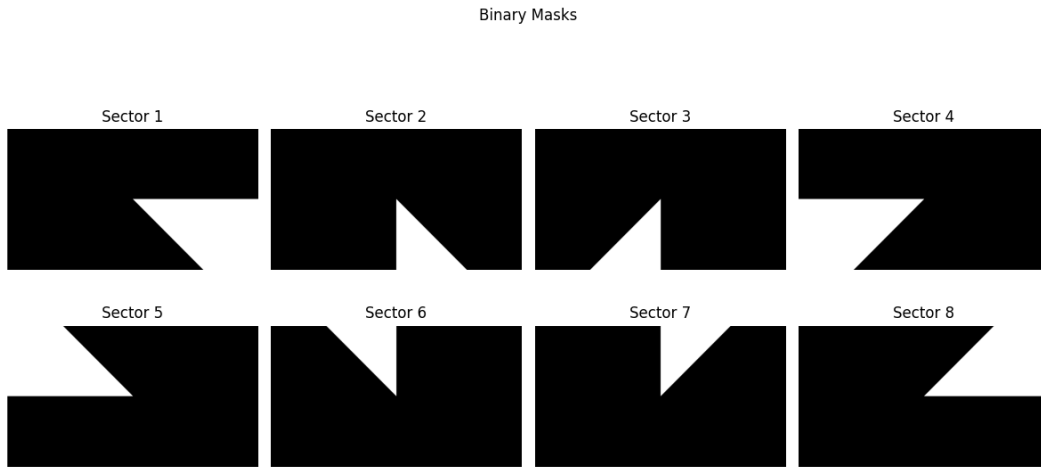


FIGURE 4.11: Representation of the 8 binary masks.

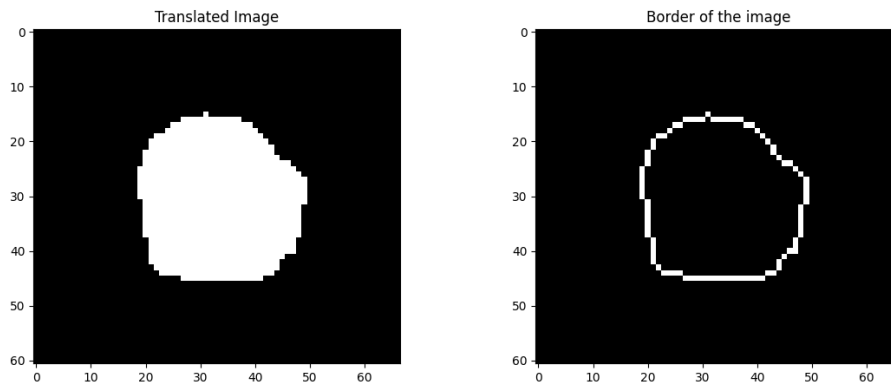


FIGURE 4.12: Mask of suspicious lesion and its border.

of the border with respect to the 8 sections (shown previously in Figure 4.11) are shown in Figure 4.13.

Finally, for each identified contour point, the Euclidean distance (Equation 4.7) from the centroid of the lesion is calculated. These distances serve as the basis for computing the standard deviation (Equation 4.8) within each sector, providing a quantitative measure of border irregularity in that specific angular region.

Equation (4.7) is employed to determine the Euclidean distance between the centroid of the lesion and each border point. Here, 'x' and 'y' denote the coordinates of the border point, while ' $center_x$ ' and ' $center_y$ ' represent the 'x' and 'y' coordinates of the centroid, respectively.

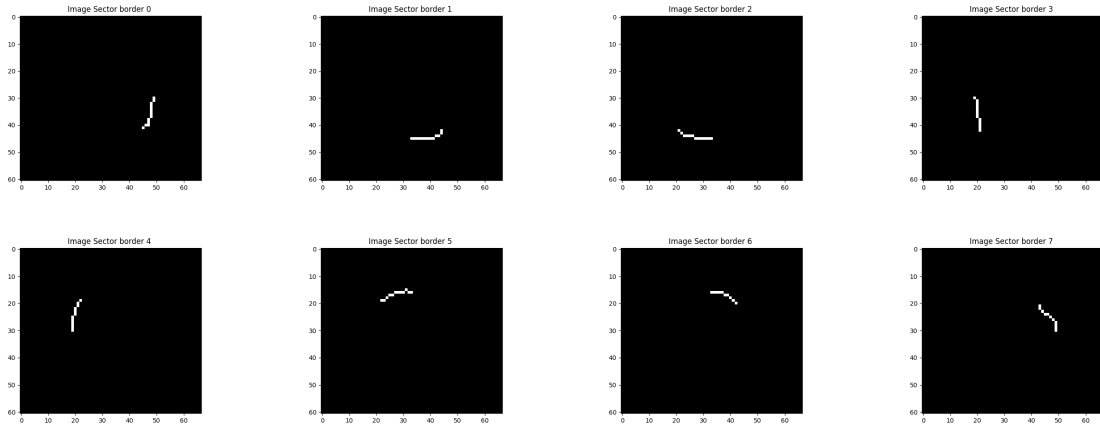


FIGURE 4.13: Division of the border of the suspicious lesion in the 8 sectors.

$$\text{Distance} = \sqrt{(x - center_x)^2 + (y - center_y)^2} \quad (4.7)$$

Equation (4.8) defines the standard deviation (σ) of a dataset of distances, where:

- n is the total number of data points in the set of distances of the section;
- x_i represents individual values in the distance dataset;
- μ is the mean of the distances of the section.

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2} \quad (4.8)$$

In essence, the standard deviation is the square root of the average of the squared differences between each data point and the mean of the dataset. It indicates how much the data points deviate from the mean. A larger standard deviation signifies greater variability within the points.

It is important to emphasize that the calculation of the standard deviation implicitly incorporates a form of 'normalization' of measurements. This is because it takes into account the dispersion of the measured distances from the center of the object, rather than evaluating absolute distances in isolation. As a result, objects of different sizes are also considered based on how the distances vary relative to their intrinsic dimensions. In other words, the standard deviation allows for the assessment of contour irregularity in a relative manner, focusing on the variability of distances from the center of the

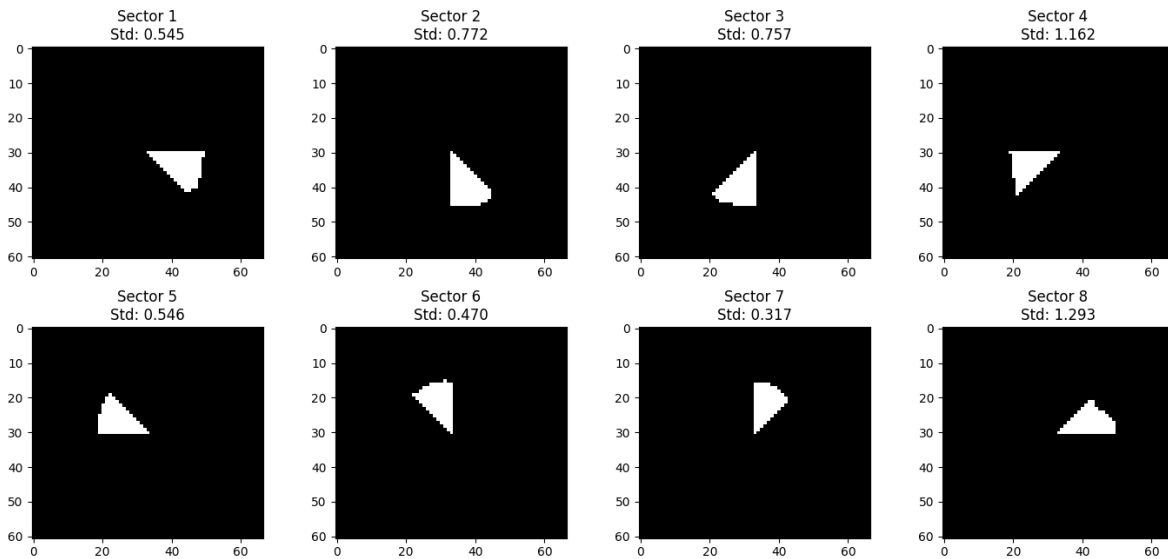


FIGURE 4.14: Division of the suspicious lesion in the 8 sectors and respectively value of the standard deviation of the sector.

object rather than the absolute dimensions of the object itself. This aspect is crucial for ensuring an accurate and meaningful comparison between lesions of different sizes and shapes.

Furthermore, to gain deeper insights into the distribution of standard deviations across all sectors, an informative histogram is generated. This visualization offers a clear representation of the frequency of different standard deviation values. The vertical lines on the histogram mark key statistical points such as the mean, the 75th percentile, the 90th percentile, the maximum, and the minimum values. These statistical values provide valuable insights into the variability and patterns of irregularities along the lesion's border.

Figure 4.14 illustrates an example of the division of the suspicious lesion into eight sectors, accompanied by the respective standard deviation values for each sector.

Figure 4.15 presents the distribution of standard deviation value across these eight sectors.

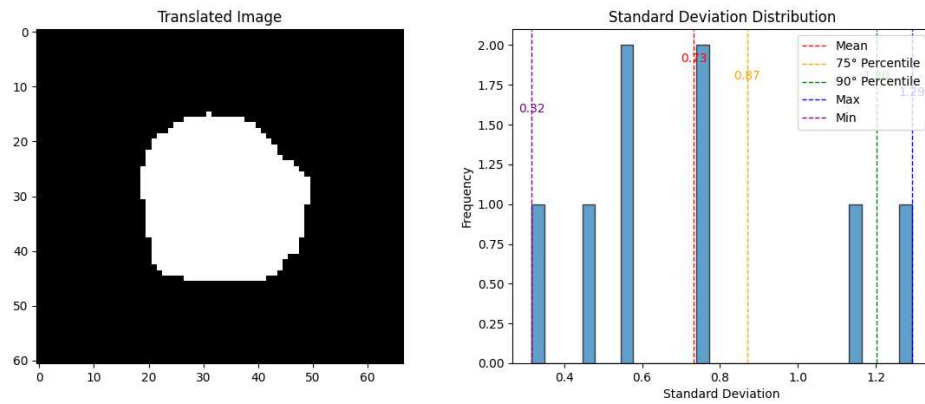


FIGURE 4.15: Distribution of the standard deviation values of the 8 sectors.

4.3.2 Compactness Index

The Compactness Index is a vital metric for assessing the shape characteristics of skin lesions, and provides valuable insights into the overall compactness and shape regularity of a lesion.

The Compactness Index is defined as CI and calculated as:

$$CI = \frac{4\pi \cdot \text{Area}}{\text{Perimeter}^2} \quad (4.9)$$

Here, 'Perimeter' refers to the boundary length of the lesion, while 'Area' represents the total surface area enclosed by the contour. The values of the CI range from 0 to 1, providing a standardized measure of the lesion's shape characteristics. A CI closer to 1 indicates a more compact and regular shape, while value closer to 0 suggest a more irregular and fragmented contour. Understanding the CI is crucial in distinguishing between different types of lesions and evaluating their morphological features.

Figure 4.16 reports four examples of compactness index evaluation: two referred to dummy images and two estimated on a suspicious lesion mask.

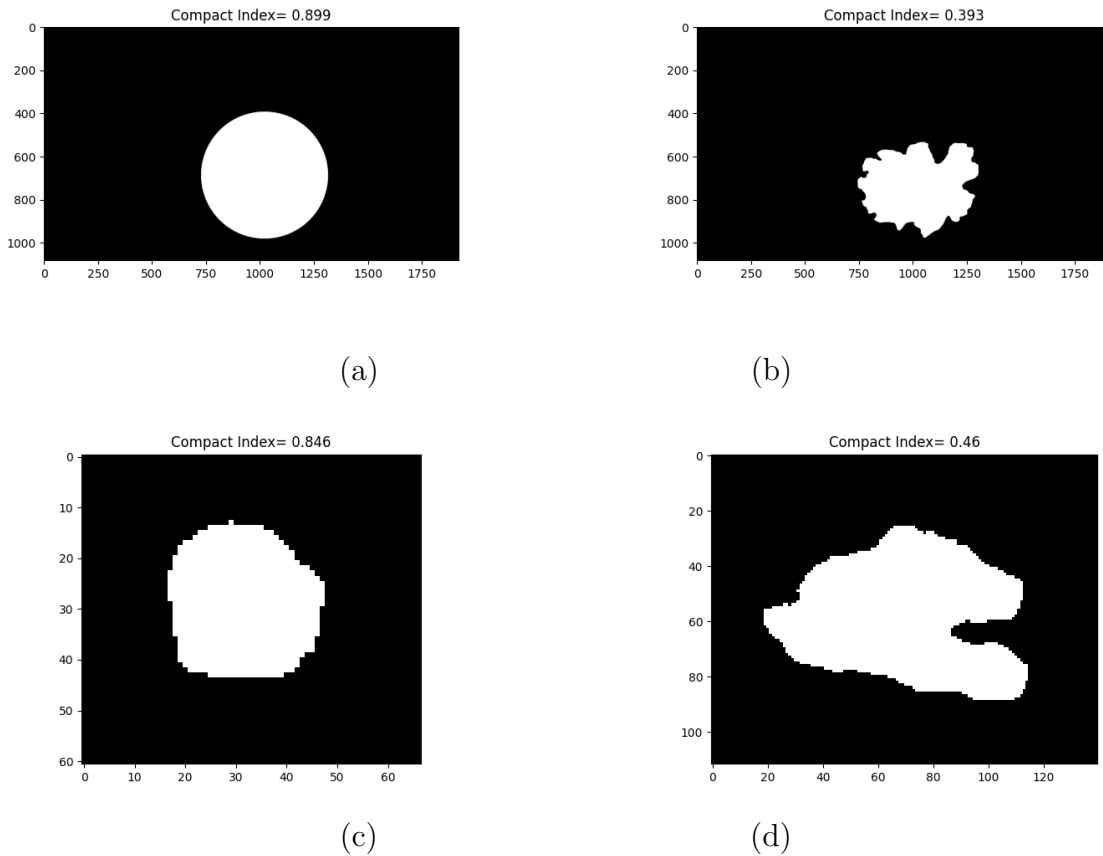


FIGURE 4.16: Example of Compactness Index: (a) Compact and regular dummy image; (b) Irregular dummy image; (c) Compact and regular suspicious lesion mask; (d) Irregular suspicious lesion mask.

4.3.3 Convex Hull

The Convex Hull, defined as the smallest convex polygon enclosing all points in a given set, is a key concept in the analysis of shapes and contours. Its function is to provide a concise and informative representation of the spatial distribution of points.

An important aspect is the relationship between the convex hull and the original contours. In particular, it is useful to analyse the variation of the perimeter length of the convex hull in relation to that of the original contour. This Percentage Discrepancy (PD) can offer important indications of local border features such as indentations or expansions.

To quantify this difference, the following equation is used:

$$\text{PD} = \frac{\text{Original Contour Perimeter} - \text{Convex Hull Perimeter}}{\text{Original Contour Perimeter}} \times 100 \quad (4.10)$$

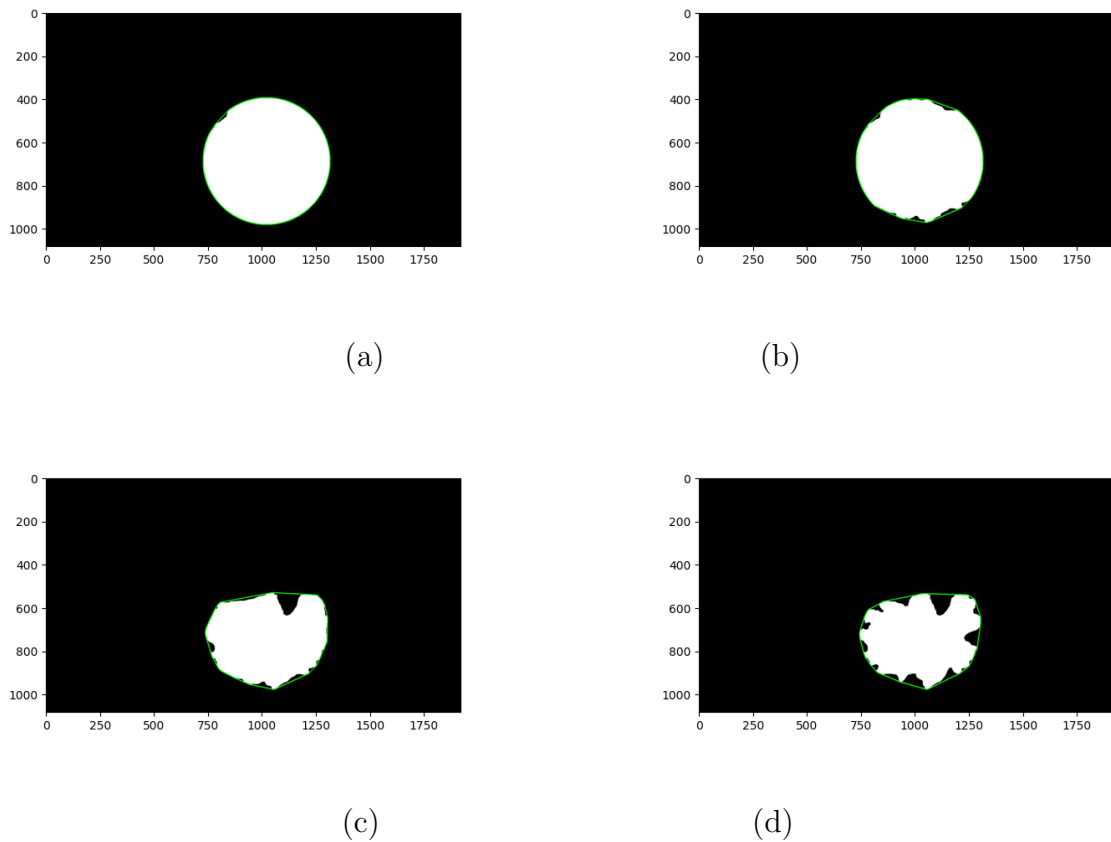
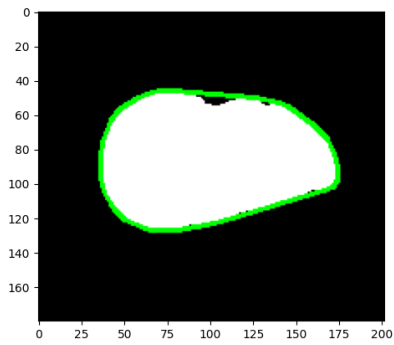


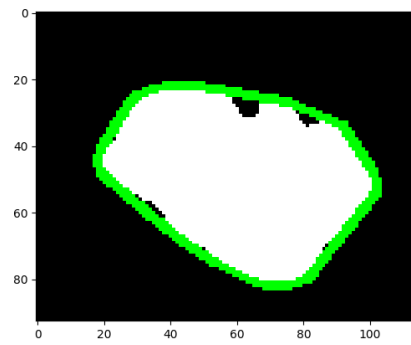
FIGURE 4.17: Convex Hull on dummy images.

This formula provides a normalized measure of the difference, highlighting whether the Convex Hull is shorter or longer compared to the original contour. A positive percentage indicates a reduction, suggesting the presence of indentations or recesses in the edge. Conversely, a negative percentage indicates an elongation of the Convex Hull compared to the original contour, suggesting the possible presence of expansions or protrusions.

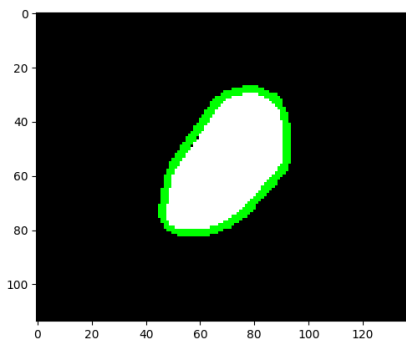
Figure 4.17 presents examples of Convex Hull (represented by the green contours) calculated on the dummy images, while Figure 4.18 shows results obtained on the masks of the suspect lesions. The Table 4.2 reports the percentage discrepancy values between the original contour perimeters and the Convex Hull perimeters, referring to the images in Figures 4.17 and 4.18.



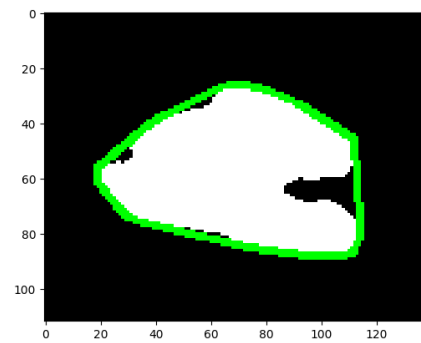
(e)



(f)



(g)



(h)

FIGURE 4.18: Convex Hull on suspicious lesion masks.

Image	Percentage Discrepancy
a	5,649 %
b	8,696 %
c	16,972 %
d	29,872 %
e	6,913 %
f	11,153 %
g	6,050 %
h	20,286 %

TABLE 4.2: Percentage Discrepancy of images of Figures 4.17 and 4.18

4.3.4 Standard deviation of ellipse-to-border distances

This method quantitatively assesses skin lesion border irregularity by computing the standard deviation of distances between points on the lesion's border and those on an elliptical approximation of the lesion.

The elliptical approximation is generated using the 'cv2.fitEllipse' function [35], which employs an error minimization algorithm to find the best-fitting ellipse based on the object's points in the provided input image. This process determines parameters such as the center, major and minor axes, and rotation angle of the ellipse, minimizing the sum of squared distances between the object's border points and the ellipse.

The resulting ellipse provides the most accurate geometric representation of the given points.

This type of elliptical approximation is employed to model the shapes of detected objects. In this context, the approximation achieved with the ellipse corresponds to the highest degree of shape and border regularity of the object. The computed distance between the points on the ellipse and the points on the object's border illustrates how much the object deviates from its regular shape.

Figures 4.19 and 4.20 show two examples of elliptical approximation, one performed on a dummy image and one performed on a suspicious lesion mask.

As a first step in implementing this code, the object within the image is translated to align its center of mass with the image center. Subsequently, an approximate ellipse of the object is defined. Then, using masks (defined as in Section 4.3.1), the object is divided into eight sections. For each section, only the pixels forming the object's border within that section are considered. For every border point, the Euclidean distance between this point and its corresponding point on the approximate ellipse is calculated. This distance metric measures the straight-line distance between two points in a Cartesian plane.

In the Equation (4.10) the variables x and y represent the coordinates of the border point. The angle is computed between the point (x, y) and the positive x -axis relative to the center of the ellipse $(center_x, center_y)$. This angle provides information about the orientation of the point with respect to the ellipse's centroid.

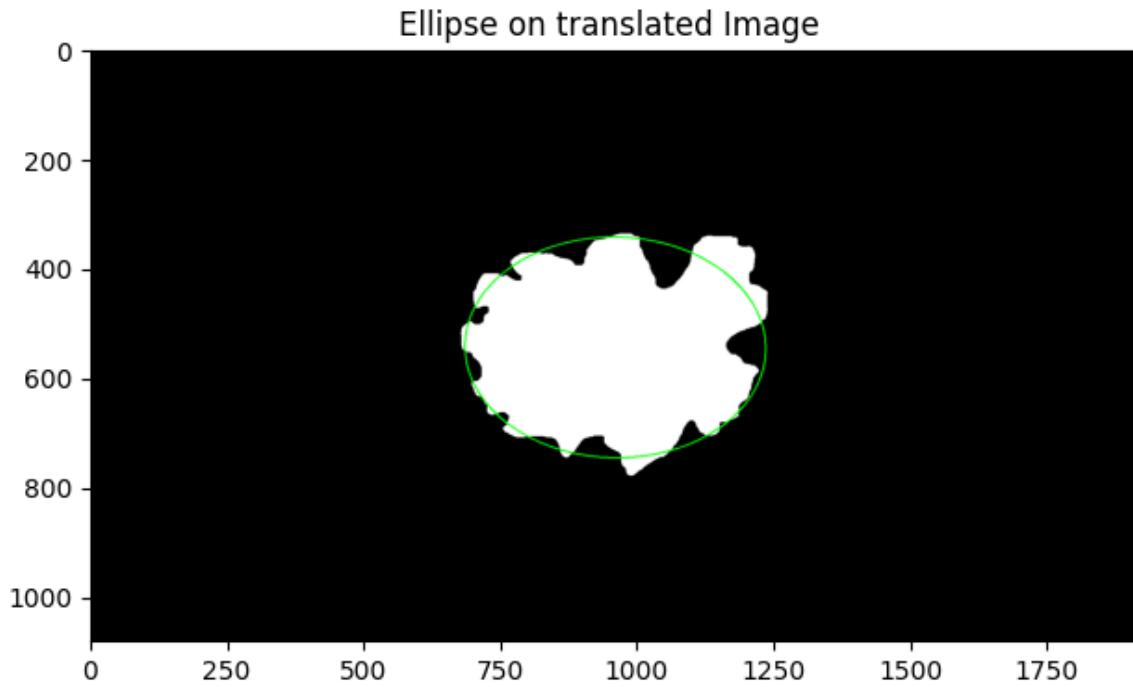


FIGURE 4.19: Elliptical approximation on a dummy image.

$$\text{angle} = \arctan 2(y - \text{center}_y, x - \text{center}_x) \quad (4.11)$$

Using trigonometric functions, coordinates of the corresponding point on the fitted ellipse, denoted as x_{ellipse} and y_{ellipse} are determined. The Equations (4.11) and (4.12) illustrate the functions used:

$$x_{\text{ellipse}} = \text{center}_x + MA \cdot \cos(\text{angle}) \quad (4.12)$$

$$y_{\text{ellipse}} = \text{center}_y + mA \cdot \sin(\text{angle}) \quad (4.13)$$

In these two equations, MA represents the major axis length of the ellipse while mA represents the minor axis length.

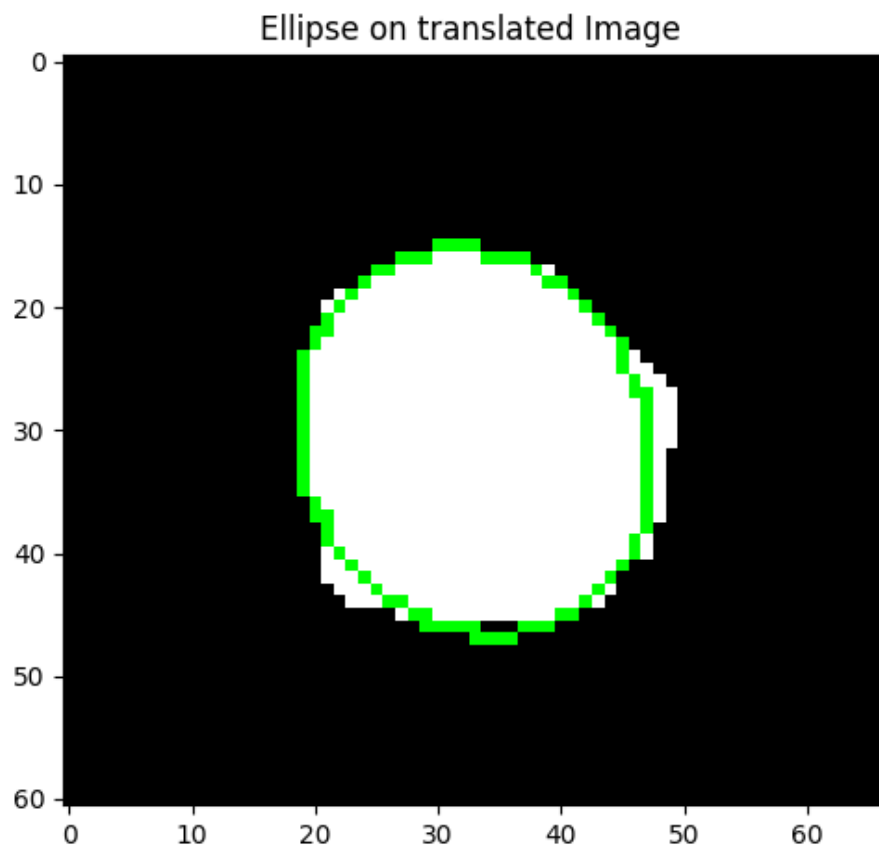


FIGURE 4.20: Elliptical approximation on a suspicious lesion mask.

Then the Euclidean distance between the border point and its corresponding point on the ellipse using the distance formula is computed:

$$\text{Distance} = \sqrt{(x - x_{\text{ellipse}})^2 + (y - y_{\text{ellipse}})^2} \quad (4.14)$$

This distance is appended to the list of distances. Finally, the standard deviation (Equation 4.8) of these distances within the sector is calculated, providing a quantitative measure of the border irregularity in that specific angular region.

The use of the approximate ellipse is crucial in analyzing irregularities in the lesion's contour. In addition to capturing geometry, this approach takes into account the object's intrinsic shape, ensuring a precise representation of irregularities. Furthermore, it implicitly normalizes the data, making the analysis independent of the object's absolute dimensions.

This allows for a relative assessment of contour irregularities, focusing on deviation

from regular shape rather than absolute measurements. This feature is crucial when comparing lesions of different sizes, ensuring accurate and reliable analysis regardless of variations in object dimensions.

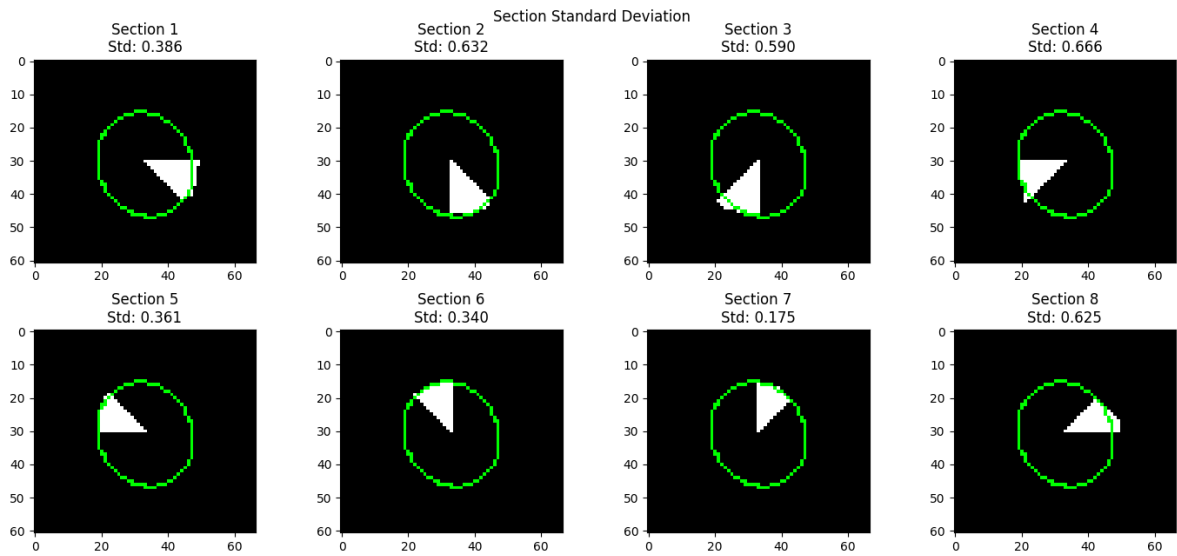
This methodology includes the computation of a histogram accompanied by vertical lines outlining key statistical parameters such as the mean, 75th percentile, 90th percentile, maximum and minimum. This addition aims to provide a more accurate representation of the distribution of standard deviations across sectors. In addition, it adds further statistical measures to assess irregularities along the lesion border, helping to improve the comprehensiveness of the evaluation process.

Figures 4.21, 4.22, 4.23 present illustrative examples of the application of this method. The first image in these figures is a representation of the subdivision of the object in the eight sections and the overlapping of the approximation ellipse of the object: for each section is reported the standard deviation of the distances between the border's points of the object and the corresponding ellipse points, while the second image reports the histogram of the distribution of the standard deviation of each section.

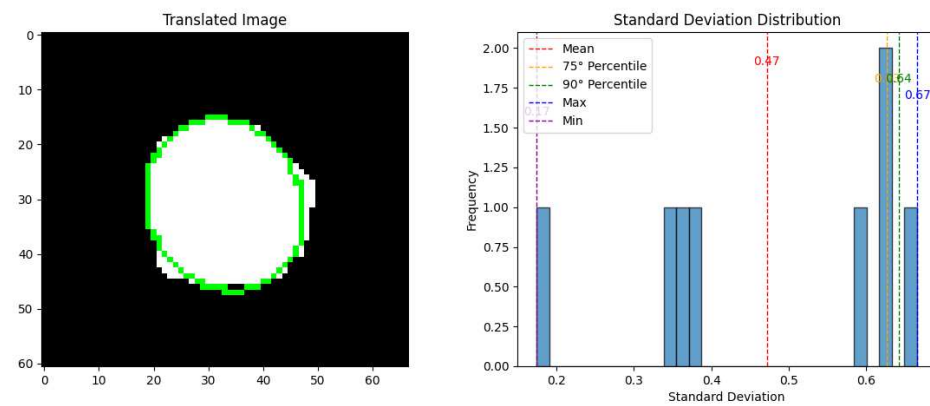
In Figure 4.21, the mask of a suspicious lesion is depicted, whose shape closely resembles an approximating ellipse. As a result, the standard deviation values in the various sections are low.

Figure 4.22 displays a mask of an irregular suspicious lesion with many indentations. In this case, the differences between the approximating ellipse and the actual lesion are evident, as reflected in the significantly higher standard deviation values.

Finally, Figure 4.23 presents a highly irregular dummy image. Once again, a substantial deviation between the object's border and the approximating ellipse is observed, which is also reflected in the standard deviation values across the different sections.

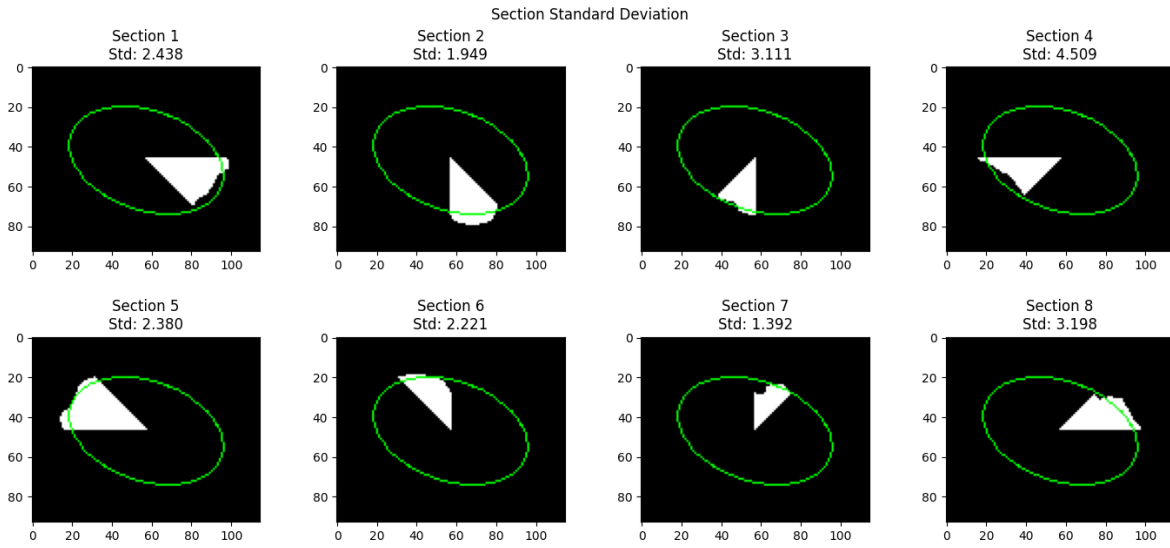


(a)

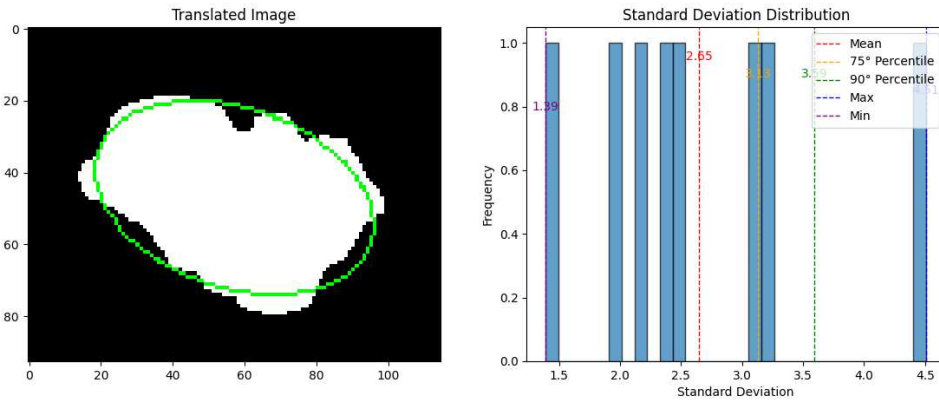


(b)

FIGURE 4.21: Results of the method applied on suspicious lesion mask: (a) Division of the object in eight sections and corresponding standard deviation; (b) Distribution of standard deviation values.

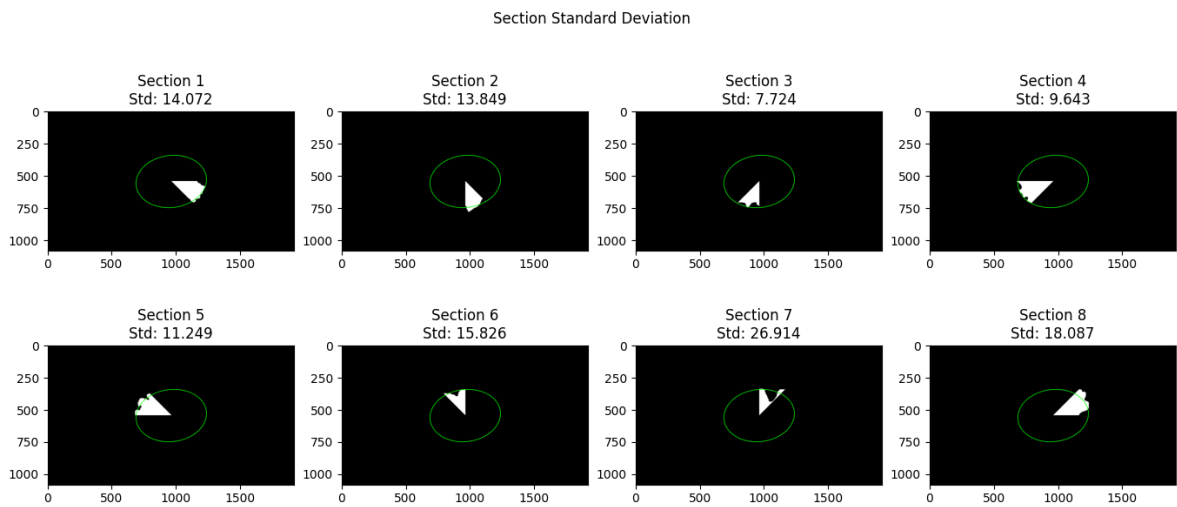


(a)

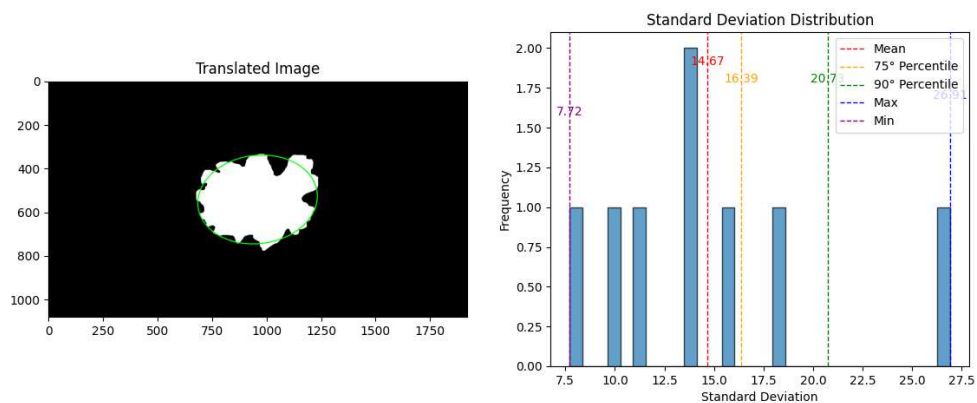


(b)

FIGURE 4.22: Results of the method applied on suspicious lesion mask: (a) Division of the object in eight sections and corresponding standard deviation; (b) Distribution of standard deviation values.



(a)



(b)

FIGURE 4.23: Results of the method applied on dummy image: (a) Division of the object in eight sections and corresponding standard deviation; (b) Distribution of standard deviation values.

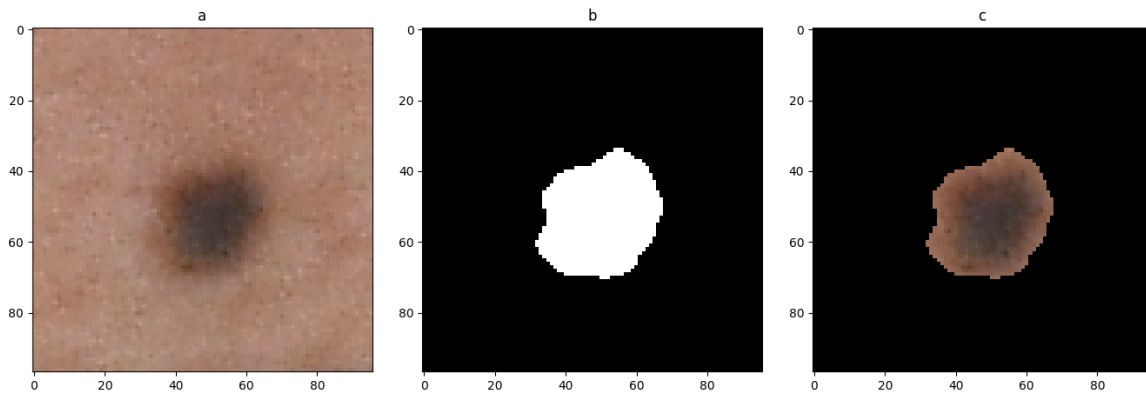


FIGURE 4.24: Color mask creation: (a) color cropped image; (b) binary mask; (c) result of bitwise AND operation;

4.4 Color

The variation in color within skin lesions denotes the existence of various pigment shades or tones visible within the lesion's borders. While melanoma lesions frequently display multiple colors, often surpassing two shades, benign lesions typically demonstrate a more consistent and uniform coloration.

To analyze the color composition within the lesions, a tailored methodology was developed. This involved creating an image conducive to color analysis derived from the original cropped image of the lesion in full color. Through a bitwise AND operation, this image was superimposed with his respective black-and-white segmented mask, resulting in an image isolating the lesion against a black background.

In Figure 4.24, an example of the cropped image (a) and the binary mask of the segmented lesion (b) are depicted. Image (c) showcases the result obtained through the bitwise AND operation applied to images (a) and (b), revealing the lesion isolated against a black background.

Subsequently, this processed image was subjected to an assessment of typical color characteristics found within skin lesions, including white, red, light brown, dark brown, gray-blue, and black hues. This assessment was preceded by an analysis of color ranges present within the dataset, considering the nuances and variations in each color across

the RGB channels. For each color, a specific range was defined within the RGB channels to encapsulate various shades and nuances, as outlined in Table 4.3:

TABLE 4.3: Color Ranges

Color Range	Range Values
Background	[(0,0,0)]
Black	[(1, 1, 1) , (50, 50, 50)]
Red	[(150, 50, 50), (255, 0, 0)]
Dark Brown	[(50, 0, 0), (150, 100, 100)]
Light Brown	[(150, 50, 6) , (200, 150, 100)]
Blue-Gray	[(0, 0, 51) , (100, 100, 200)]
White	[(150, 150, 150) , (255, 255, 255)]

- **Background (0,0,0):** represents the black background pixels of the image.
- **Black (1, 1, 1) - (50, 50, 50):** Encompasses very dark to medium-dark gray shades.
- **Red (150, 50, 50) - (255, 0, 0):** Includes shades ranging from medium to dark red and intense red hues.
- **Dark Brown (50, 0, 0) - (150, 100, 100):** Encompasses dark brown tones with potential hints of red or green, influenced by the RGB channel values.
- **Light Brown (150, 50, 6) - (200, 150, 100):** Includes shades spanning from light orange to light brown with orange or reddish hues.
- **Blue-Gray (0, 0, 51) - (100, 100, 200):** Incorporates various shades of blue, possibly including grayish or violet tones with bluish undertones.
- **White (150, 150, 150) - (255, 255, 255):** Encompasses light gray shades up to pure white, featuring subtle gray undertones.

Utilizing the defined color ranges, binary masks are generated for each specific color interval, highlighting pixels in the original image falling within the specified color range. For example, a 'red' mask would exhibit white pixels representing red pixels in the original image and black elsewhere. Counting the white pixels within each mask reveals the quantity of pixels in the original image corresponding to the specified color range.

A distinct 'background' range is defined, excluding the black background from the color analysis. This exclusion ensures the black background isn't considered as a color while setting a range specifically for black, excluding the background yet encompassing pixels with shades akin to black within the lesion.

Moreover, to ascertain the significant presence of a particular color, a criterion is applied: if the count of pixels within a color range, excluding the 'background,' amounts to less than 1% of the total colored pixels constituting the lesion area, it isn't regarded as a definitive presence of that color. This criterion serves to distinguish substantial occurrences of predefined colors within the image, acknowledging only those colors accounting for at least 1% of the colored pixels as effectively present.

Figures 4.25, 4.26, 4.27, and 4.28 provide four practical examples of applying the described methodology. Each image displays color masks associated with specific hues: Mask 1 represents the background, Mask 2 black color, Mask 3 red, Mask 4 dark brown, Mask 5 light brown, Mask 6 blue-gray, Mask 7 white, and the last image depicts the original picture. Specifically, Figure 4.24 illustrates a lesion with shades of dark brown, light brown, and blue-gray. Figure 4.25 highlights a lesion featuring dark brown and light brown tones. Figure 4.26 depicts a lesion displaying light brown and white hues, while Figure 4.27 presents a lesion with black, dark brown, light brown, and blue-gray tones.

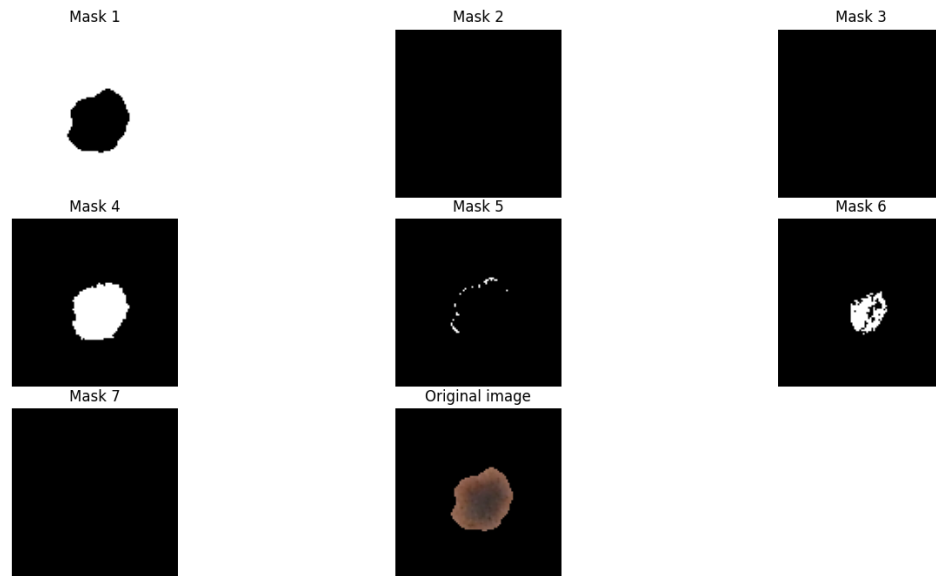


FIGURE 4.25: Lesion with dark brown, light brown and blue-gray colors: Mask 1= background; Mask 2= black; Mask 3= red; Mask 4= dark brown; Mask 5= light brown; Mask 6= blue-gray; Mask 7= white; Last image= original image.

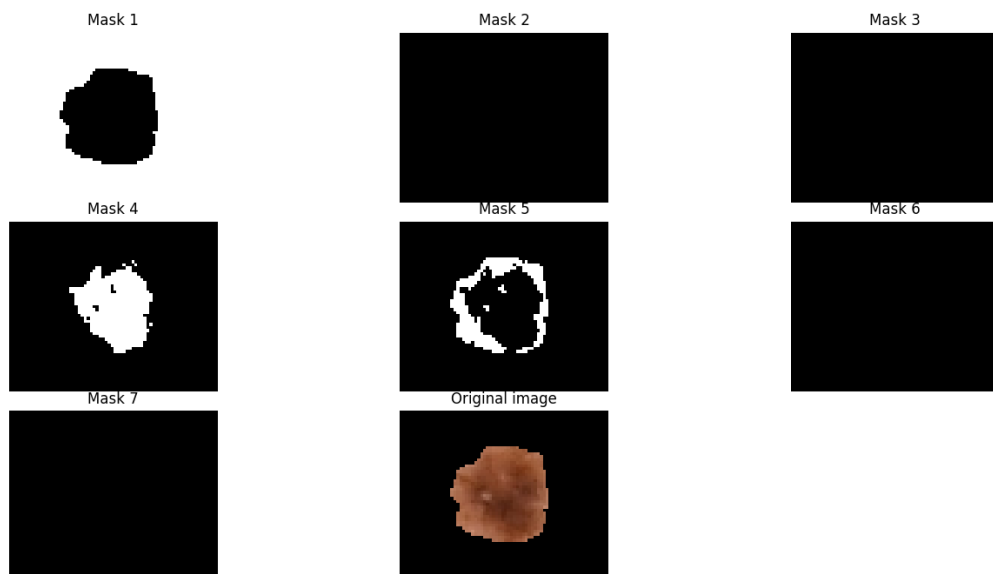


FIGURE 4.26: Lesion with dark brown and light brown colors: Mask 1= background; Mask 2= black; Mask 3= red; Mask 4= dark brown; Mask 5= light brown; Mask 6= blue-gray; Mask 7= white; Last image= original image.

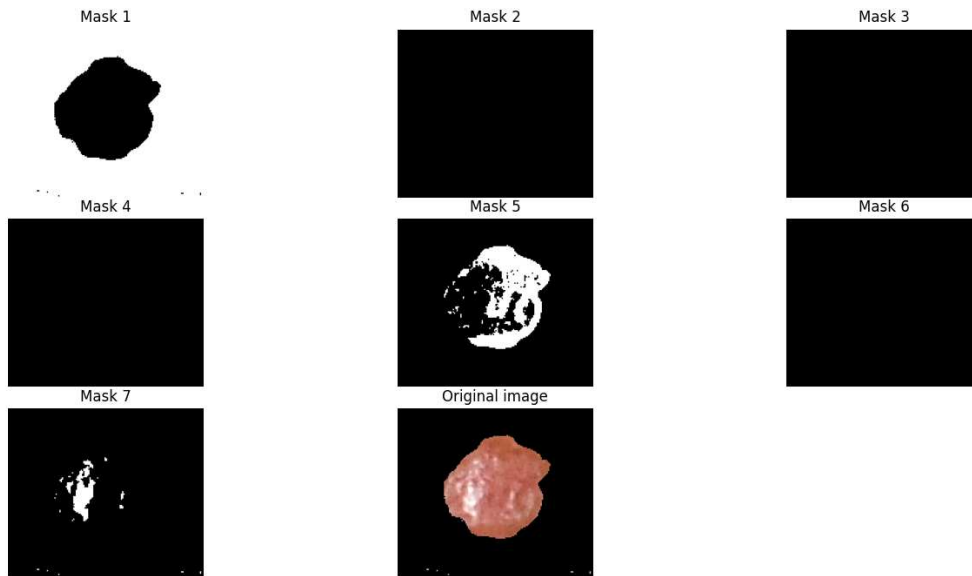


FIGURE 4.27: Lesion with light brown and white colors: Mask 1= background; Mask 2= black; Mask 3= red; Mask 4= dark brown; Mask 5= light brown; Mask 6= blue-gray; Mask 7= white; Last image= original image.

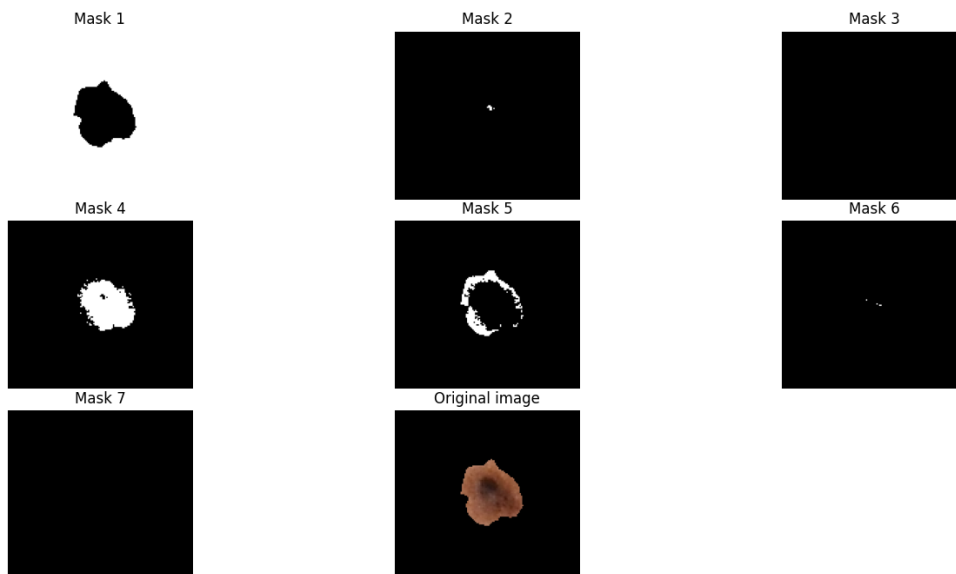


FIGURE 4.28: Lesion with black, dark brown, light brown and blue-gray colors: Mask 1= background; Mask 2= black; Mask 3= red; Mask 4= dark brown; Mask 5= light brown; Mask 6= blue-gray; Mask 7= white; Last image= original image.

Chapter 5

Results

In this chapter, the results obtained by applying the methodologies for quantitative assessment of asymmetry, border irregularity and color, outlined in Chapter 4, will be presented. The chapter begins with the validation of asymmetry and border irregularity methods using the dummy images; followed by the presentation of the results obtained by applying these techniques to images of both suspicious and benign skin lesion masks. Specifically, unsupervised learning through k-means clustering is employed to define three categories of asymmetry and three categories of border irregularity. Finally, the chapter concludes by presenting the results of applying the color methodology to images of both benign and suspicious lesions.

5.1 Validation of methodologies on dummy images

Validation is a crucial step in verifying the effectiveness of the methods employed. It involves comparing the results obtained through the application of the methods with a set of known or simulated data, such as dummy images. This serves to establish whether the evaluation parameters, accurately reflect the clinical feature present in the images. The validation of the methodologies is carried out using the dummy images to allow for a comparison between the known increase in asymmetry and the corresponding rise in values associated with asymmetry assessment. The same principle applies to border irregularity, where a known increase in irregularity corresponds to an increase in the evaluation parameter values.

An accurate validation confirms the robustness and reliability of the methodologies used in the context of skin lesion analysis.

5.1.1 Validation of methodologies for assessing asymmetry

As previously mentioned, the validation of asymmetry is conducted using the dummy images introduced in Chapter 4 (see Figure 4.3). For each image, both the percentage of non-overlapping area and the asymmetry index are calculated. The results obtained through the application of this methodology are summarized in Table 5.1.











Table 5.1 contains several relevant parameters: the parameters "NAOx" and "NAOy" represent the non-overlapping area of the object when the image is divided along the x and y axes, respectively. Both of these values are expressed in pixels. It is important to note that NAOx and NAOy are closely related to the dimensions of the object and should not be considered as definitive evaluation indicators.

On the other hand, the final evaluation parameters are "%NAO" (Percentage of Non-Overlapping Area) and "AI" (Asymmetry Index). The percentage of non-overlapping area represents the fraction of the total area of the object that does not overlap with other parts of the image. The asymmetry index quantifies the degree of asymmetry of the object, indicating on how many axes it can be considered asymmetric. A higher AI value indicates greater asymmetry.

The results presented in the table play a crucial role in validating the methodology used. The visual observation of an increase in asymmetry is reflected in an increase in both the percentage of the non-overlapping area and the asymmetry index. Furthermore, a higher percentage of non-overlapping area indicates that the asymmetry of the object is evident along a greater number of axes.

The correlation between quantitative parameters and visual perception of asymmetry confirms the effectiveness of the methodology employed in evaluating the object in question.

TABLE 5.1: Asymmetry validation results

Image	Image name	NOAx	NOAy	%NOA	AI
	Asymmetry 0	702	0	0,279	0
	Asymmetry 1	2543	1243,5	1,735	0
	Asymmetry 2	3126	2457	2,837	0
	Asymmetry 4	6013	815	2,916	1
	Asymmetry 5	3375	7638	6,023	1
	Asymmetry 6	14779,5	21287,5	22,561	2
	Asymmetry 7	13634,5	29722,5	30,447	2
	Asymmetry 8	11906,5	31806,5	34,167	2
	Asymmetry 9	21147	29936	45,305	2
	Asymmetry 10	19923,5	28062	45,793	2

5.1.2 Validation of methodologies for assessing border irregularity

Similarly to the assessment of asymmetry, the validation of methodologies developed for the quantitative evaluation of border irregularity is based on the use of dummy images.

For each image, various methodologies are applied, including the calculation of the standard deviation of distances from the border to the center of mass of the object, the

use of the compactness index, the implementation of the convex hull methodology, and the determination of the standard deviation of distances between the object's edge and the fitted ellipse.

In the following sections the results obtained through the analysis of these various methodologies are presented.

5.1.2.1 Validation of standard deviation of the center-of-mass border distances

The first methodology introduced involves calculating the standard deviation of the distances between the center of mass and the object's border.

Table 5.2 presents the standard deviation values for each section, along with their cumulative sum.











The *Sum* of standard deviation values can be a significant metric as it provides an aggregate measure of the variability in distances.

The data reported in Table 5.2, concerning the validation of the model based on the standard deviation of distances between the center of mass and the edge, demonstrate that the parameter defined as the cumulative sum of the standard deviations of the sections accurately reflects the increase in border irregularity. Specifically, as the visually observable irregularity increases, so does the value of the sum.

Table 5.3 provides key metrics for the histogram representation (refer to Figure 4.15), including the mean, minimum, maximum, 75th percentile, and 90th percentile of the standard deviation distribution.











The 75th and 90th percentiles serve as crucial statistical measures. The 75th percentile represents the value below which 75% of the data falls, providing insight into the central tendency of the standard deviations. This means that the majority of sections tend to cluster around this point. If the 75th percentile is high, it may indicate that a significant portion of sections have relatively high standard deviations, signifying greater irregularity. Similarly, the 90th percentile is particularly sensitive to outliers or extreme values compared to the 75th percentile. It represents the value below which 90% of the data falls. If the 90th percentile is significantly higher than the 75th percentile, it could suggest the presence of some sections with very high standard deviations, indicating a

TABLE 5.2: Standard deviation of center of mass-border distances validation results

Image	Std 1	Std 2	Std 3	Std 4	Std 5	Std 6	Std 7	Std 8	Sum
	0,303	0,303	0,257	0,355	0,259	0,259	0,355	0,257	2,348
	0,259	0,259	0,355	0,257	4,19	4,909	0,257	0,355	10,841
	3,89	11,86	0,56	0,981	6,11	6,244	0,744	0,783	31,172
	3,577	9,085	8,716	1,386	5,848	5,702	9,284	1,178	44,756
	16,467	10,981	6,173	0,65	2,098	10,665	12,623	13,605	73,262
	16,573	10,698	6,562	2,12	5,398	10,687	20,002	13,758	85,798
	10,933	7,927	11,058	9,251	11,959	13,271	42,919	11,255	118,573
	19,457	12,962	11,182	9,431	24,459	13,32	43,732	11,579	146,122
	18,37	12,067	10,448	11,381	24,111	21,03	43,14	26,569	167,116
	19,166	13,518	20,305	11,674	20,654	20,951	42,214	28,196	176,678

potential area of interest where irregularity is particularly pronounced. Analyzing these percentiles values allows for a deeper understanding of the distribution of standard deviations across different sections of the object, aiding in the identification of areas with pronounced irregularities. This enhances the overall assessment of the object's irregularity.

TABLE 5.3: Distribution Statistics for Std of Distances

Image	Mean	Min	Max	75th Per	90th Perc
	0,294	0,257	0,355	0,316	0,355
	1,355	0,257	4,909	1,314	4,406
	3,897	0,56	11,86	6,144	7,929
	5,595	1,178	9,284	8,808	9,145
	9,158	0,65	16,467	12,869	14,464
	10,725	2,12	20,002	14,462	17,602
	14,822	7,927	42,919	12,287	22,165
	18,265	9,431	43,732	20,706	30,241
	20,890	10,448	43,14	24,726	31,540
	22,085	11,674	42,214	22,762	32,401

The inclusion of mean, minimum, and maximum values in Table 5.3 further enriches the characterization of the standard deviation distribution. The mean serves as a central reference point, offering insight into the typical deviation observed across sections. Meanwhile, the minimum and maximum values pinpoint the range of variability present in the measurements. The minimum value indicates the lowest standard deviation recorded among the sections, highlighting areas with relatively low variability.

On the other hand, the maximum value showcases sections with the highest standard deviations, drawing attention to regions exhibiting notable variability.

By incorporating these additional statistical measures, it is possible to gain a more comprehensive understanding of how standard deviations are distributed across different segments of the object. This enhanced analysis provides valuable context for pinpointing areas of pronounced irregularity and strengthens the overall assessment of the object's border characteristics.

5.1.2.2 Validation of compactness index and convex hull methodologies

The second methodology focuses on the exploitation of the compactness index to evaluate border irregularity. This approach entails quantifying the compactness of the object, providing a measure of how closely its shape resembles a circle. Additionally, the third methodology employs the convex hull, which outlines the smallest convex polygon encompassing the object's contour. In particular, it is useful to analyse the variation of the perimeter length of the convex hull in relation to that of the original contour. This Percentage Discrepancy can offer important indications of local border features such as indentations or expansions. Both these techniques offer unique insights into the object's border characteristics.











Table 5.4 reports the values of the Compactness Index and Percentage Discrepancy parameters calculated on the dummy images.

As specified earlier, CI is defined within the range of 0-1, where 1 corresponds to a perfectly circular shape, and values lower than 1 indicate shapes progressively deviating from circularity. These characteristics are observed in the 10 dummy images as well. Specifically, the first two images exhibit values very close to 1, while the values gradually decrease, approaching zero for the subsequent eight images. This departure from circularity is also visually discernible in the images.

Turning to Percentage Discrepancy, the percentage values increase with increasing border irregularity. In this particular case, the values for images 4 and 5 do not follow a strict ascending pattern. Figure 5.1 offers an in-depth view of these two images alongside their respective convex hulls. This visual comparison underscores the accuracy of

PD's assessment, as figure "Border Irr 5" exhibits fewer indentations compared to image "Border Irr 4".

TABLE 5.4: Validation of CI and PD

Image	Image name	CI	PD
	Border Irr. 1	0,899	5,067
	Border Irr. 2	0,885	5,649
	Border Irr. 3	0,854	6,755
	Border Irr. 4	0,806	8,696
	Border Irr. 5	0,785	8,54
	Border Irr. 6	0,724	11,724
	Border Irr. 7	0,590	16,972
	Border Irr. 8	0,528	21,197
	Border Irr. 9	0,417	29,099
	Border Irr. 10	0,393	29,872

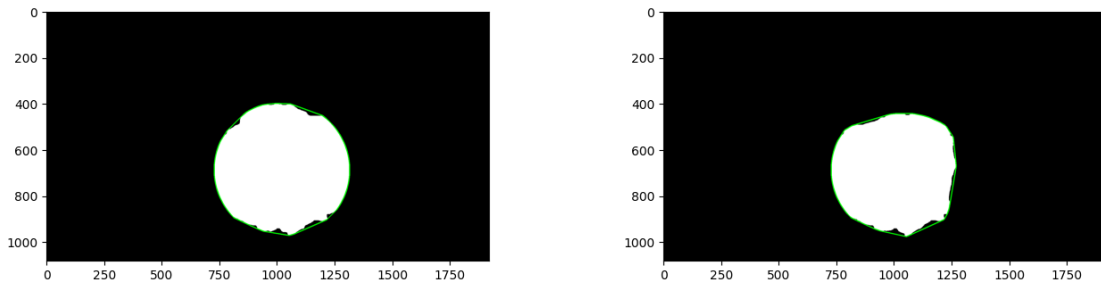


FIGURE 5.1: Comparison of PD: left image "Border Irr 4"; right image "Border Irr 5".

5.1.2.3 Validation of standard deviation of the ellipse-border distances

The fourth methodology involves utilizing the standard deviation of distances between the border and the fitted ellipse for assessing border irregularity. Like the previous methodologies, this approach also undergoes validation using dummy images.

Two tables are presented: Table 5.5 displays the standard deviation values for each section into which the object is divided, along with their cumulative sum. The second table (Table 5.6) showcases statistical values, found in the histogram distribution of the standard deviations (refer to Figure 4.20), including the mean, minimum, maximum, 75th percentile, and 90th percentile.

The application of these statistics serves the same purpose as outlined in the preceding section, where the methodology utilizing standard deviation of center-border distances was validated. This comprehensive analysis provides a multi-faceted understanding of border irregularities, contributing to a robust evaluation of the object's contours.

Examining the results of the methodology application reveals that in the case of highly irregular objects, it's important to acknowledge that the cumulative sum of standard deviations may not provide a complete representation of the escalating border irregularity. This phenomenon is particularly evident when comparing highly irregular objects, as seen in the images labeled as "Border Irr 8", "Border Irr 9", and "Border Irr 10". Visually, there is a noticeable increase in irregularity, but this effect is not fully reflected in the standard deviation values.

This discrepancy may be attributed to the behavior of the fitting ellipse in the presence of highly irregular shapes. The approximation ellipse tends to better trace the contours of the object, resulting in reduced variability in the measured distances compared to less irregular objects. This behavior can influence the standard deviation results, making the cumulative sum of standard deviations less indicative of border irregularity in these specific cases.

Furthermore, it's possible that the cumulative sum of standard deviations is more sensitive to certain types of irregularities than others. For instance, it may be more influenced by small, regular variations compared to large and irregular changes in the object's shape.

TABLE 5.5: Standard deviation of ellipse-border distances validation results





















Image	Std 1	Std 2	Std 3	Std 4	Std 5	Std 6	Std 7	Std 8	Sum
	0,303	0,303	0,257	0,355	0,202	0,202	0,355	0,257	2,221
	0,191	0,208	0,247	0,403	3,525	3,633	0,320	0,277	8,804
	2,675	6,540	1,426	1,204	4,646	4,696	0,364	1,576	23,128
	3,827	4,633	3,493	3,286	3,826	4,221	5,785	2,966	32,038
	7,891	5,541	3,240	2,126	3,630	7,114	4,597	5,501	39,639
	8,081	5,579	3,095	2,125	5,094	5,630	9,462	6,249	45,316
	9,250	13,062	12,715	7,911	12,382	13,825	30,154	17,420	116,717
	12,823	16,256	13,153	7,9920	16,198	14,615	28,866	17,900	127,803
	13,256	15,194	12,574	8,172	14,979	12,122	26,897	17,520	120,714
	14,072	13,849	7,724	9,643	11,249	15,826	26,914	18,087	117,363

TABLE 5.6: Distribution Statistics for Std of Distances

Image	Image Name	Mean	Min	Max	75th Per	90th Perc
	Border Irr. 1	0,278	0,203	0,348	0,314	0,348
	Border Irr. 2	1,101	0,191	3,633	1,184	3,558
	Border Irr. 3	2,891	0,364	6,540	4,659	5,249
	Border Irr. 4	4,005	2,966	5,786	4,324	4,979
	Border Irr. 5	4,955	2,127	7,891	5,934	7,347
	Border Irr. 6	5,664	2,125	9,462	6,707	8,496
	Border Irr. 7	14,590	7,911	30,154	14,723	21,234
	Border Irr. 8	15,975	7,992	28,866	16,667	21,190
	Border Irr. 9	15,089	18,172	26,897	15,775	20,333
	Border Irr. 10	14,670	7,724	26,914	16,391	20,735

5.1.3 Parameters chosen to assess border irregularity

The final parameters selected for evaluating border irregularity are as follows:

- **Cumulative Sum of Standard Deviations of center-border distances (Sum):** This parameter has demonstrated its effectiveness in quantifying the gradual increase in overall border irregularity. It provides a comprehensive measure of the variability in distances between the center of mass and the object's edge.
- **Compactness Index:** CI allows for the consideration of the object's deviation from perfect circularity. It provides valuable insights into the shape irregularities, especially in cases where the object's contour differs significantly from a circle.
- **Percentage Discrepancy:** PD takes into account the presence of indentations along the object's border and assigns greater weight to the presence of deeper indentations. This parameter proves essential in capturing finer details of irregularities.

By combining these parameters, it is possible to create a comprehensive evaluation framework that addresses various aspects of irregularity, from overall trends to specific shape deviations and fine details in the border's contour. This multi-dimensional approach enhances the accuracy and depth of the assessment.

The cumulative sum of distances between the fitted ellipse and the object's border will not be considered in subsequent stages, as it did not prove effective in evaluating the increase in irregularity observed in the dummy images. It's important to note that while these parameters have shown promise in assessing border irregularity, further analysis may be needed to understand their behavior in more complex scenarios or with different types of objects. Additionally, considering potential interactions between these parameters in capturing specific types of irregularities could provide even deeper insights into the object's characteristics.

5.2 Evaluation of asymmetry in benign and suspicious skin lesions

In this section, the asymmetry of benign and suspicious skin lesions will be examined. For each image, the percentage of Non-Overlapping Area and the Asymmetry Index

is calculated. These parameters provide a detailed assessment of lesion asymmetry, a critical element in dermatological diagnosis.

Subsequently, an unsupervised learning approach will be applied using k-means clustering. Unlike supervised methods that rely on predefined class labels, this approach considers only the intrinsic characteristics of the lesions. This methodology aims to unveil potential patterns or similarities among lesions, with the ultimate goal of classifying them based on their asymmetry.

It's essential to recognize the observed degree of asymmetry is not the only feature used by professional dermatologists to classify a skin lesion as suspicious melanoma. For instance, certain lesions labeled as benign may exhibit a significant level of asymmetry, while some labeled as suspicious may appear perfectly symmetrical but later be confirmed as melanomas through biopsy. This emphasizes the significance of considering asymmetry alongside other parameters for a comprehensive and accurate diagnosis.

The incorporation of k-means clustering serves to extract meaningful patterns from the asymmetry-related data. These discernible patterns play a crucial role in categorizing lesions based on their varying degrees of symmetry or asymmetry.








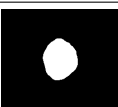


Table 5.7 presents significant results obtained from applying the methodology to the masks of suspicious and benign lesions. Lesions starting with "s" are from the suspicious lesions dataset, while those labeled with "b" come from the benign dataset. Both datasets showcase instances of both symmetric and asymmetric lesions.

The table includes labels for the images. Suspicious lesions underwent biopsy, and their labels correspond to the ground truth. In contrast, benign lesions lack a defined label as they were classified as benign by dermatologists and did not undergo biopsy.

One noteworthy case is Lesion "s3," which exhibits a high percentage of non-overlapping area. Upon closer inspection of the corresponding image (depicted in Figure 5.2 (a)), it becomes evident that the mask has a notably low resolution, with a limited number of pixels. This explains the elevated percentage of non-overlapping area, making it comparable to lesions with high asymmetry. Moreover, when rotating an image, especially if the object within the image has low resolution, interpolation becomes necessary to estimate pixel values in new positions. Interpolation is a mathematical process that

estimates pixel values at non-integer positions based on surrounding pixel values. Essentially, it 'fills in' the gaps between original pixels to align with the new object position after rotation. However, as observed in this case, if the object in the image initially has low resolution, interpolation may introduce additional information or distortions during rotation (as evident in Figure 5.2(b)). Therefore, when conducting asymmetry assessments on rotated lesions, it is crucial to carefully consider the rotational context and the impact of interpolation on measurement accuracy.

TABLE 5.7: Asymmetry evaluation on suspicious and benign lesion masks

Image	Image Name	%NOA	AI	Label
	s1	8,553	2	low grade dysplasia
	s2	6,042	1	basalioma
	s3	10,183	1	low grade dysplasia
	s4	10,259	2	low grade dysplasia
	s5	6,787	2	melanoma in situ
	s6	1,743	0	low grade dysplasia
	b1	2,138	0	benign
	b2	2,133	0	benign
	b3	5,681	1	benign
	b4	9,527	2	benign

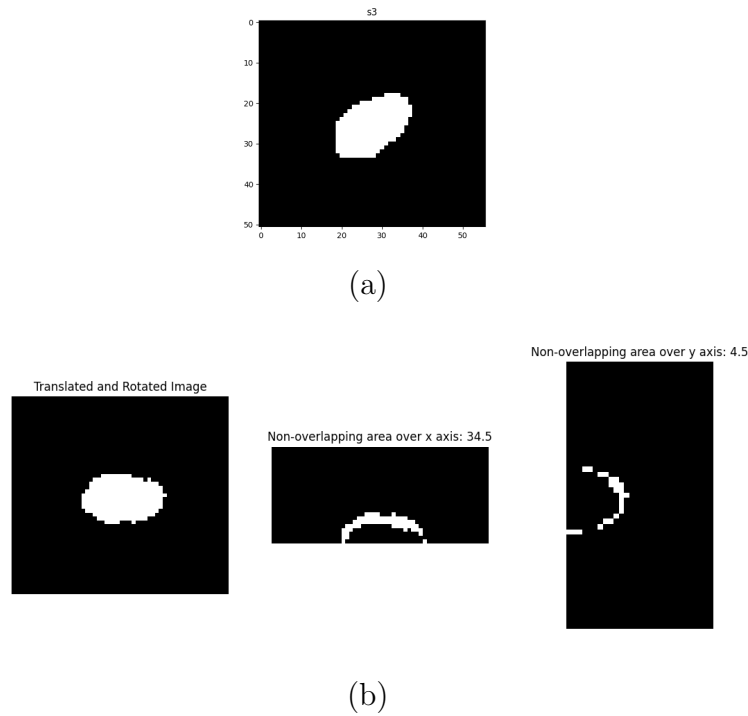


FIGURE 5.2: Asymmetry on low resolution image.

The following phase of the analysis consists of the use of the data obtained from the application of methodologies to mask of benign and suspicious lesions to do an unsupervised learning using k-means clustering.

The initial step of this process involves estimating the optimal number of clusters to use. To achieve this goal, the elbow method is used, an evaluation technique that helps identify the ideal number of clusters.

The elbow method works by varying the number of clusters and observing the clustering error trend as this number changes. The fundamental idea is to identify the point where adding more clusters does not provide a significant improvement in error reduction. This point, visually resembling an "elbow" in the error curve, represents the optimal number of clusters for the data.

In Figure 5.3, the results of the elbow method are presented, from which it is clear that the error trend suggests 3 clusters as the optimal outcome for the analysis. The error metric utilized in this instance is inertia, also known as within-cluster sum of squares. This metric is a measure of how compactly the clusters are gathered. It quantifies the sum of squared distances between each data point and the centroid of its assigned cluster.

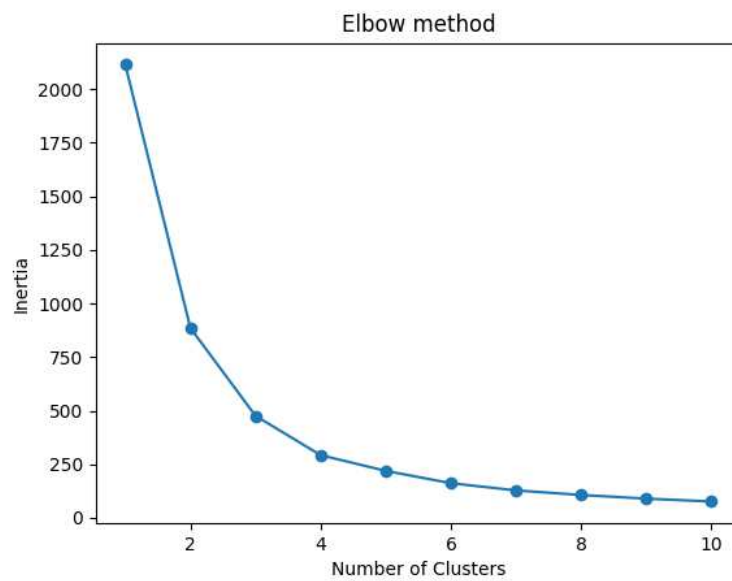


FIGURE 5.3: Number of clusters for asymmetry dataset.

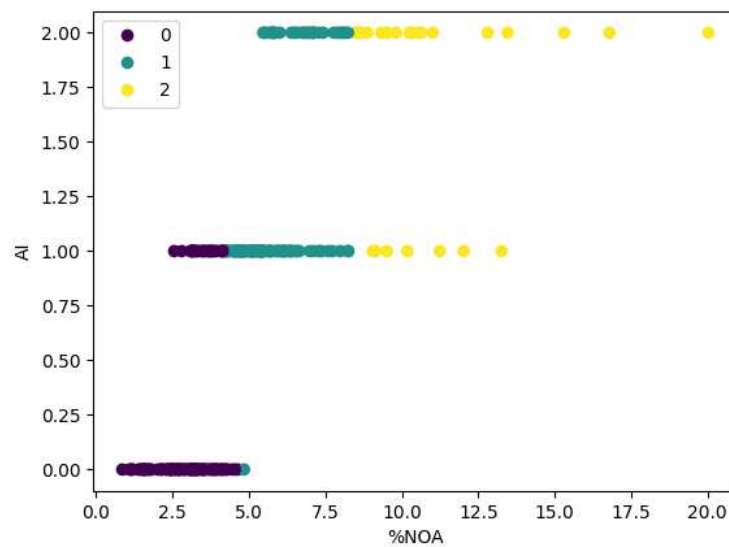


FIGURE 5.4: Clusters of asymmetry dataset.











Subsequently, k-means clustering is applied using the optimal number of clusters identified by the elbow method. The resulting clusters are illustrated in Figure 5.4, highlighting the subdivision of lesions into groups.

Table 5.8 reports cluster labels assigned to images from Table 5.7, defining three asymmetry levels:

- 0: symmetric lesion;
- 1: lesion with intermediate asymmetry;
- 2: highly asymmetric lesion.

This unsupervised clustering approach facilitates the identification of patterns or similarities among lesions, contributing to categorization based on the degree of asymmetry.

TABLE 5.8: Cluster labels obtained from asymmetry assessment

Image	Image Name	%NOA	AI	Cluster Label
	s1	8,553	2	2
	s2	6,042	1	1
	s3	10,183	1	2
	s4	10,259	2	2
	s5	6,787	2	1
	s6	1,743	0	0
	b1	2,138	0	0
	b2	2,133	0	0
	b3	5,681	1	1
	b4	9,527	2	2

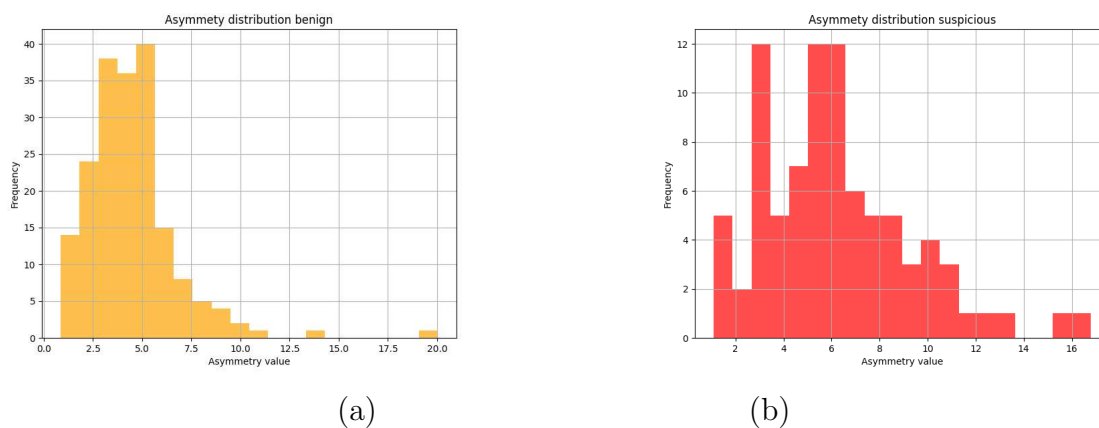


FIGURE 5.5: Asymmetry parameter distribution: (a) benign lesion; (b) suspicious lesions.

Following the clustering analysis, the examination proceeds to investigate potential statistical differences in the asymmetry values between the images within the benign and suspicious datasets. This phase of the study aims to discern whether there exist significant dissimilarities in the asymmetry patterns of lesions between the two datasets, thus contributing valuable insights into the distinct characteristics of benign and suspicious skin lesions. It's important to note that existing literature widely acknowledges the presence of such differences in asymmetry when derived from dermoscopic images. The primary objective here is to extend this understanding by exploring whether similar asymmetry disparities are observable in the indices derived from macroscopic images. This critical step will help determine the persistence and relevance of asymmetrical features previously observed in dermoscopic contexts within the broader-scale image data.

Figure 5.5 illustrates the distributions of the %NOA values within both groups, revealing characteristics that diverge from the typical bell-shaped normal distribution. This non-normality is further affirmed by the visual representation in Figure 5.6 through the boxplot of these distributions.

The Levene test was utilized to assess variance homogeneity among these distributions, resulting in a statistic of 11.5431 and a p-value of 0.0009. These findings indicate a significant difference in variances between the groups, specifically in their asymmetry values.

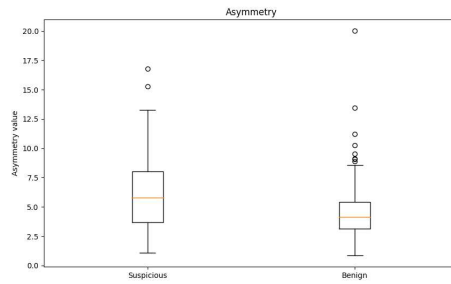


FIGURE 5.6: Boxplot of Asymmetry parameter distribution.

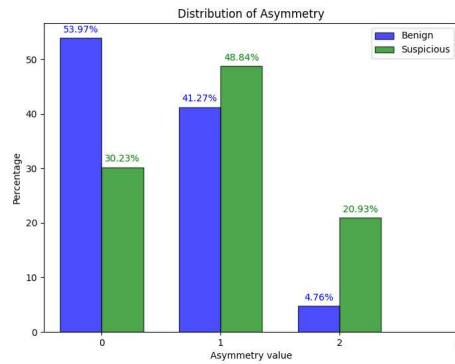


FIGURE 5.7: Distribution of clustering values for asymmetry.

Due to the non-Gaussian distribution and the observed heteroscedasticity, the Mann-Whitney U test, a non-parametric approach, was employed. This test evaluates differences between two independent samples to determine if they stem from the same population. In the case of asymmetry, the Mann-Whitney U test yielded a U Statistic of 9887.0, with a p-value of 6.274×10^{-7} . This considerably low p-value, below the conventional significance level of 0.05, strongly rejects the null hypothesis. These results confirm a substantial dissimilarity in asymmetry values between benign and suspicious lesions, highlighting notable disparities in asymmetry characteristics that potentially serve as discriminative factors between the two lesion types.

An important aspect of the analysis involved a statistical examination to determine if categorizing asymmetry into three classes (derived from clustering) would demonstrate a significant difference between the distributions of benign and suspicious lesions. The percentages representing these distributions is depicted in both Figure 5.7 and Table 5.9.

Cluster	Benign Lesions (%)	Suspicious Lesions (%)
0	53.97	30.23
1	41.27	48.84
2	4.76	20.93

TABLE 5.9: Distribution of clustering values for asymmetry

These values illustrate the proportions of asymmetry within each cluster for both benign and suspicious lesions. Statistical tests were then employed to validate the observed differences: Levene's test assesses variance between groups. The statistic of 0.0025, paired with a corresponding p-value of 0.9598, suggests a similarity in the variance of asymmetry values across clusters for both benign and suspicious lesions. Meanwhile, the Mann-Whitney U Test, a non-parametric comparison of asymmetry value distributions between benign and suspicious lesions, yielded a U-statistic of 10569.0, accompanied by an extremely low p-value of 9.74×10^{-6} . These results strongly indicate a statistically significant distinction in asymmetry distributions between the two groups.

These statistical analyses robustly support the conclusion that there is a notable distinction in asymmetry patterns between benign and suspicious lesions, as evidenced by both the clustering proportions and the outcomes of the conducted tests.

5.3 Evaluation of border irregularity in benign and suspicious skin lesions

In this section, the focus shifts to the evaluation of border irregularities in benign and suspicious skin lesions, specifically exploring the Cumulative Sum of standard deviations of distances from the lesion's centroid (Sum), the CI and the PD . These metrics play a key role in providing an understanding of irregularities within dermatological cases, serving as crucial indicators in the diagnostic process.












Similar to the previous examination of asymmetry, each image undergoes the calculation of the Sum , the CI and the PD parameters. The goal is to unravel insights into the irregularities present in the borders of skin lesions, shedding light on potential markers of malignancy or other dermatological conditions. Following the computation of the parameters, an unsupervised learning approach using k-means clustering is employed. This method, distinct from supervised techniques reliant on predefined labels, focuses solely on the inherent characteristics of the lesions. The objective is to uncover latent patterns or similarities among lesions, ultimately facilitating their classification based on irregularity in their borders.

It's crucial to note that labels assigned by dermatologists may not always align with the observed irregularities. Similar to the case of asymmetry, in this scenario as well, medical labels are not solely based on the assessment of border irregularity but encompass a set of clinical parameters. Indeed, some lesions labeled as suspicious may appear quite regular, while certain lesions labeled as benign may exhibit a certain degree of irregularity. Understanding this complexity emphasizes the importance of considering a variety of features in dermatological evaluation for a comprehensive and accurate diagnosis.

Table 5.10 presents noteworthy findings obtained from applying the methodology to the masks of suspicious and benign lesions. Lesions prefixed with "s" are from the suspicious lesions dataset, while those labeled with "b" are from the benign dataset. Both datasets showcase instances of lesions with diverse irregularities in their borders.

The subsequent phase of the analysis involves the use of data acquired through the

TABLE 5.10: Border irregularity evaluation on suspicious and benign lesion masks

Image	Image Name	Sum	CI	PD
	s1	44,972	0,703	6,913
	s2	18,327	0,754	7,052
	s3	6,632	0,789	2,972
	s4	27,169	0,635	11,153
	s5	12,581	0,701	6,994
	s6	3,539	0,871	4,588
	s7	36,883	0,460	20,286
	b1	7,169	0,870	4,903
	b2	10,986	0,871	4,499
	b3	14,598	0,717	9,095
	b4	25,314	0,668	11,274

application of methodologies to the masks of benign and suspicious lesions for unsupervised learning using k-means clustering. The first step in this process is the scaling of the variables. In k-means clustering with three variables, scaling is crucial to standardize units of measurement and balance the influence of variables, preventing those

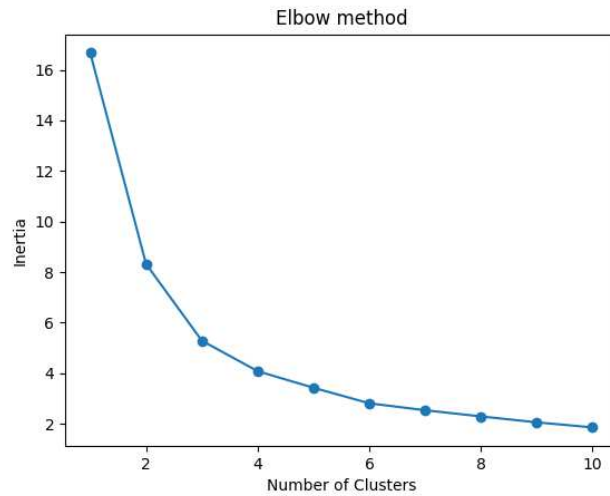


FIGURE 5.8: Elbow method used in border irregularity assessment.

with a wide range from dominating the clustering process. This ensures fair weighting of variables and enhances clustering effectiveness. The Min-Max Scaling transformation is applied to a variable X using the following formula:

$$X_{\text{scaled}} = \frac{X - \min(X)}{\max(X) - \min(X)} \quad (5.1)$$

where:

- X_{scaled} is the scaled value of the variable,
- X is the original value of the variable,
- $\min(X)$ is the minimum value of the variable,
- $\max(X)$ is the maximum value of the variable.

This transformation scales the variable's values to a range between 0 and 1, maintaining the proportionality between the original values.

After the scaling, the optimal number of clusters has to be determined. Similar to the approach taken for asymmetry, the optimal number of clusters is determined using the elbow method, utilizing inertia as the error parameter. In Figure 5.8, the results of the elbow method are presented, from which it is clear that the error trend suggests 3 clusters as the optimal outcome for the analysis.

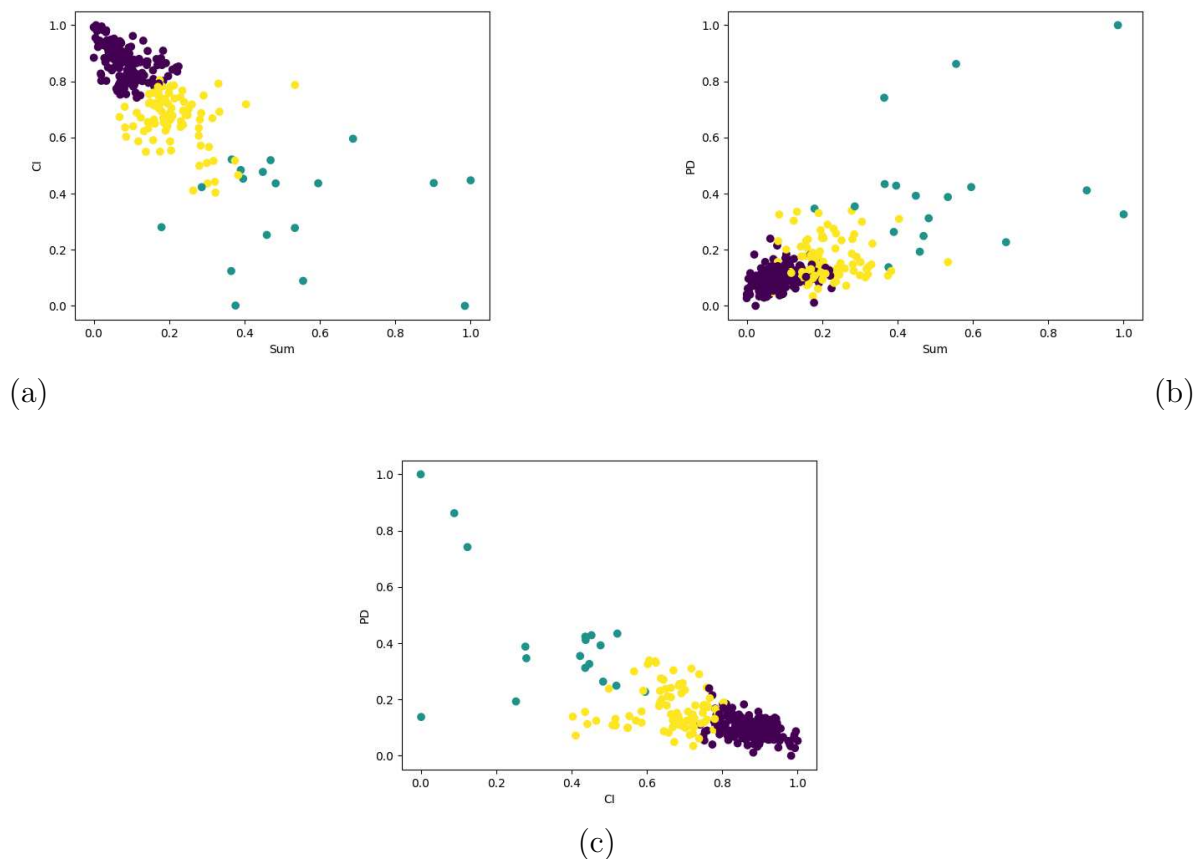


FIGURE 5.9: Clusters of border irregularity.

Following the determination of the optimal number of clusters as 3 using the elbow method, k-means clustering is subsequently applied. The resulting clusters are visually represented in Figure 5.9, illustrating the segmentation of lesions into distinct groups. Figure 5.9 (a) depicts the representation of Sum and CI, (b) showcases the representation of Sum and PD, and (c) displays CI and PD. In particular, when viewing these three images, one can see an overlap of points within the clusters. The overlap of points within the clusters, as observed in the visualizations, suggests some degree of similarity or proximity between lesions in different groups. This phenomenon could be attributed to the inherent complexity of the data, where certain lesions may share common characteristics that lead to their inclusion in multiple clusters.

To enhance visual clarity, Principal Component Analysis (PCA) was performed. PCA is a technique used to transform the original variables into a new set of uncorrelated variables, known as principal components, which capture the maximum variance in the data. In this case, the first two principal components were utilized to visualize the

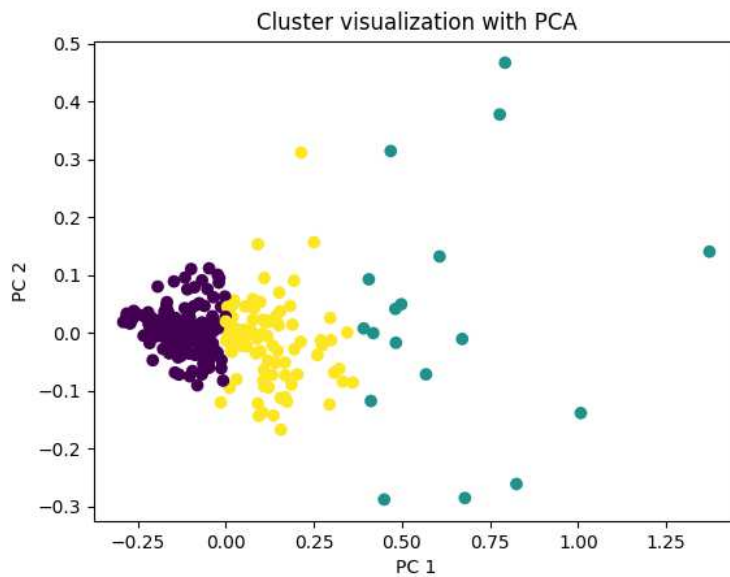









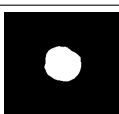
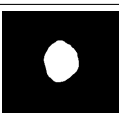


FIGURE 5.10: Clusters visualization with PCA.

clusters, as reported in Figure 5.10. The implementation of PCA aims to condense information into a more visually digestible format, thereby enhancing the clarity of cluster distribution in the reduced-dimensional space defined by the first two principal components. This simplification facilitates the identification of potential clusters, making it significantly easier for visual interpretation.

The cluster labels assigned to the images in Table 5.10 are detailed in Table 5.11. The cluster labels in Table 5.11 categorize the images into three groups:

- 0: regular borders,
- 1: slightly irregular borders,
- 2: highly irregular borders.

TABLE 5.11: Cluster labels obtained from border irregularity assessment

Image	Image Name	Sum	CI	PD	Cluster label
	s1	44,972	0,703	6,913	2
	s2	18,327	0,754	7,052	1
	s3	6,632	0,789	2,972	0
	s4	27,169	0,635	11,153	2
	s5	12,581	0,701	6,994	1
	s6	3,539	0,871	4,588	0
	s7	36,883	0,460	20,286	2
	b1	7,169	0,870	4,903	0
	b2	10,986	0,871	4,499	0
	b3	14,598	0,717	9,095	1
	b4	25,314	0,668	11,274	2

Following the clustering analysis, the examination proceeds to investigate potential statistical differences in the border irregularity values between the images within the benign and suspicious datasets. This phase of the study aims to discern whether there exist significant dissimilarities in the irregularity patterns of lesion borders between the two datasets, thus contributing valuable insights into the distinct characteristics of benign and suspicious skin lesions. It's pertinent to acknowledge the established findings in literature that highlight discernible differences in border irregularities when assessed through dermoscopic imagery. The goal is to extend this understanding by exploring whether similar disparities in border irregularity are discernible within the indices derived from macroscopic images.

The exploration of statistical differences in border irregularity involves a comprehensive comparison of metrics derived from image analysis techniques applied to both benign and suspicious lesions. The goal is to ascertain whether certain irregularity parameters exhibit noteworthy variations between the two groups, signifying potential discriminatory power in distinguishing between benign and suspicious lesions.

In Figure 5.11, histograms display the distribution of values for the Sum of variance of distances within benign and suspicious cases. Both distributions do not exhibit a clear Gaussian pattern, as observed from the histograms. The lack of symmetry around the center, along with the non-bell-shaped form, suggests a non-normal distribution, highlighting the presence of outliers. These values, outside the typical data distribution, might indicate lesions with unique or uncommon characteristics within the datasets. Furthermore, the discrepancy between mean and median values suggests a non-symmetric data distribution. This non-Gaussian distribution characteristic is also visible in Figure 5.12, which presents the box plot of these distributions.

To evaluate whether the distributions have equal variances, the Levene's test was conducted. This test assesses homoscedasticity, determining if sample variances significantly differ. The test results indicate a statistic of 13.1527 and a p-value of 0.00034. A low p-value suggests a significant difference in variances among lesion groups, indicating that the distributions might not have equivalent variances.

Subsequently, to assess if there is a significant difference in border irregularity values

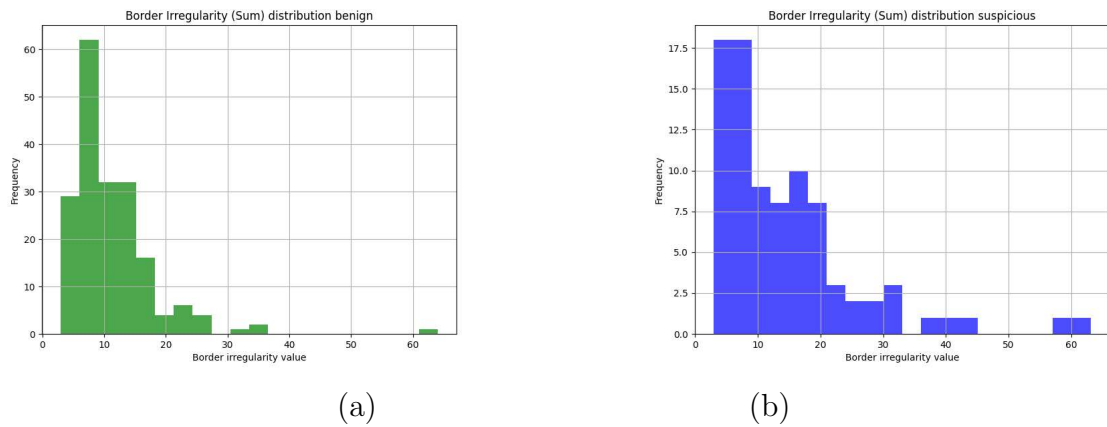


FIGURE 5.11: Sum parameter distribution: (a) benign lesion; (b) suspicious lesions.

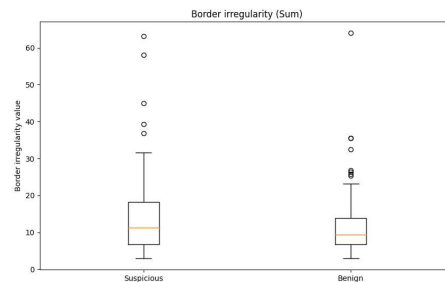


FIGURE 5.12: Boxplot of Sum parameter distribution.

between the two lesion groups, the Mann-Whitney U test was performed. This non-parametric test compares distributions of the two groups and evaluates if their central tendencies significantly differ. The Mann-Whitney U test results show a U statistic of 7881.5 and a p-value of 0.1738. In this case, the p-value suggests a non-significant difference between the two groups.

Following the assessment of Sum of variance of distances, a similar procedure was conducted for the Perimeter Discrepancy and Compactness Index parameters.

In Figure 5.13, histograms depict the distributions of PD values, while the corresponding boxplot is presented in Figure 5.14. Again, the distributions are non-Gaussian. The Levene's test was applied to evaluate variance equality, resulting in a statistic of 21.3810 and a p-value of 5.816810^{-6} . The obtained p-value suggests a substantial difference in variances between the benign and suspicious lesion datasets for the PD parameter.

Subsequently, the Mann-Whitney U test was employed to investigate potential differences in PD values between the two groups of lesions. The test yielded a U statistic of

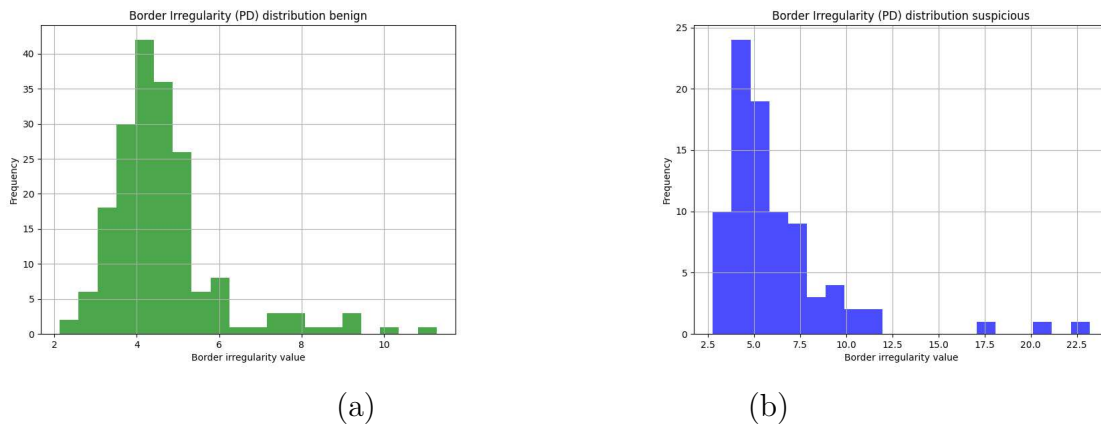


FIGURE 5.13: PD parameter distribution: (a) benign lesion; (b) suspicious lesions.

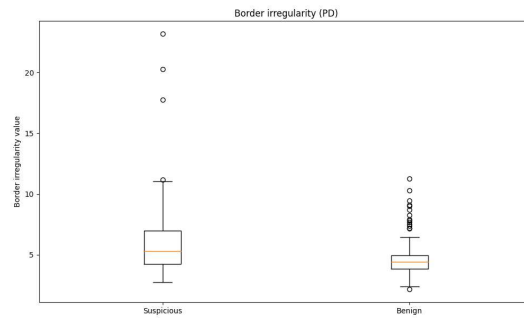


FIGURE 5.14: Boxplot of PD parameter distribution.

11036.0 and a p-value of 1.9657×10^{-6} . This outcome indicates a statistically significant difference in PD values between the benign and suspicious lesion datasets.

After examining PD, a parallel analysis was conducted for the CI parameter. Figure 5.15 displays histograms representing the distributions of CI values, while the corresponding boxplot is depicted in Figure 5.16. Upon applying Levene’s test to assess variance equality, the calculated statistic was 12.7976 with a p-value of 0.0004. The derived p-value suggests a notable difference in variances between the benign and suspicious lesion datasets concerning the CI parameter.

Subsequently, the Mann-Whitney U test was utilized to examine potential differences in CI values between the two lesion groups. The test resulted in a U statistic of 5201.0 and a p-value of 1.7123×10^{-6} , indicating a statistically significant dissimilarity in CI values between the benign and suspicious lesion datasets.

These statistical findings align with the earlier analyses of Perimeter Discrepancy, reinforcing substantial disparities in variances and CI values between the benign and

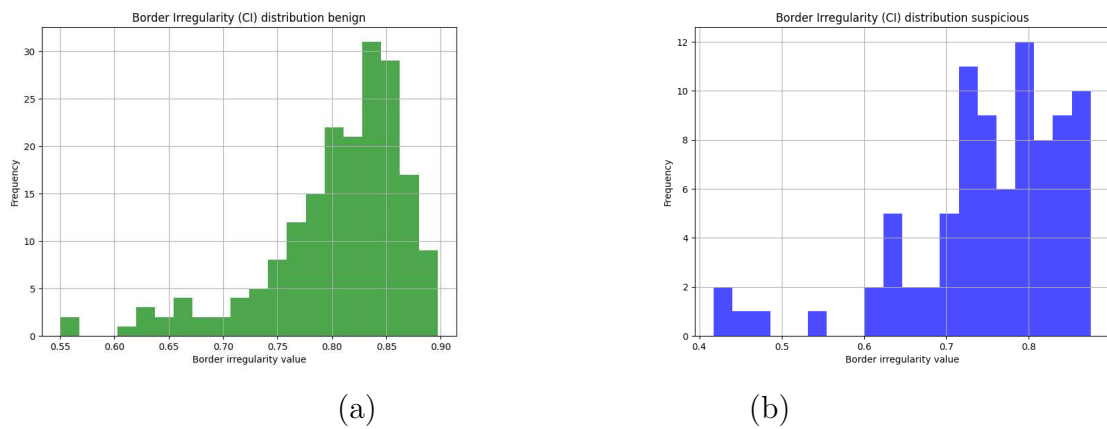


FIGURE 5.15: CI parameter distribution: (a) benign lesion; (b) suspicious lesions.

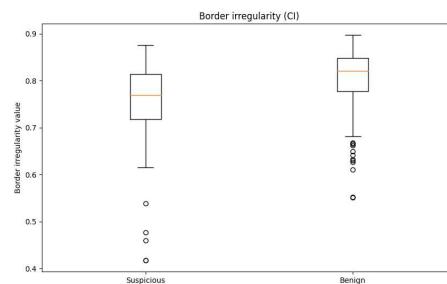


FIGURE 5.16: Boxplot of CI parameter distribution.

suspicious lesion groups. Such distinct irregularity patterns within these datasets reinforce the potential utility of these parameters in discriminating between the two lesion categories.

Additionally, scatter plots were generated to explore the relationship between asymmetry and border irregularity, including Sum, PD, and CI. These graphs visualize the distribution of lesions based on their asymmetry and the variation in border irregularity parameters, highlighting potential patterns or distinctions between benign and suspicious lesions.

The analysis of the scatter plots reveals a significant aspect: in all three cases - asymmetry-Sum (Figure 5.17), asymmetry-PD (Figure 5.18), and asymmetry-CI (Figure 5.19)- a substantial difference in the centroids of distributions between benign and suspicious lesions emerges. In particular, it can be seen that the centroid of benign lesions consistently tends towards lower values, indicating less asymmetry and border irregularity than in suspicious lesions. This phenomenon suggests that there are specific configurations of asymmetry and border irregularity parameters that are more frequent

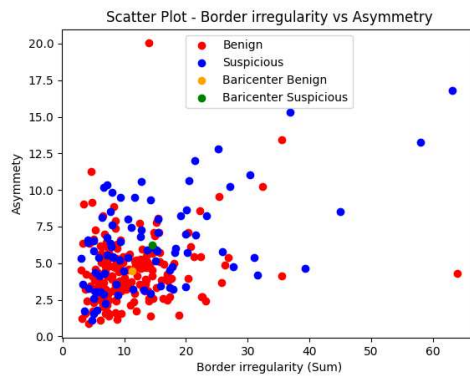


FIGURE 5.17: Scatter plot of Asymmetry-Sum.

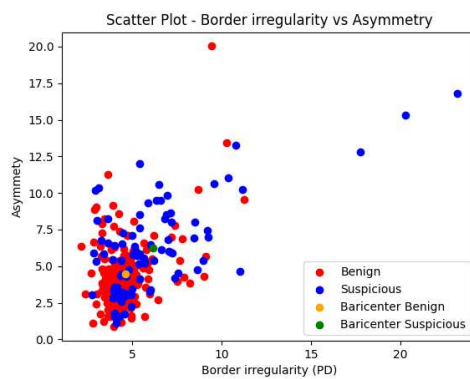


FIGURE 5.18: Scatter plot of Asymmetry-PD.

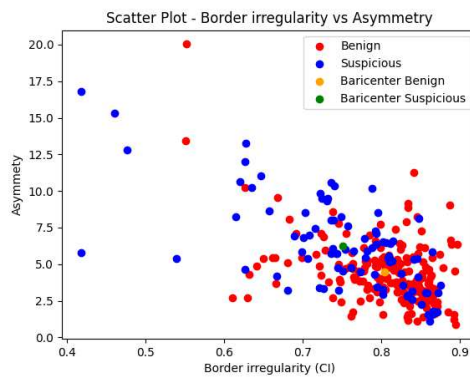


FIGURE 5.19: Scatter plot of Asymmetry-CI.

in one of the two groups of lesions, outlining a potential distinction between benign and suspicious lesions.

An essential component of the analysis entailed a statistical investigation to assess whether categorizing border irregularity into three classes (derived from clustering) would reveal a notable difference between the distributions observed in benign and

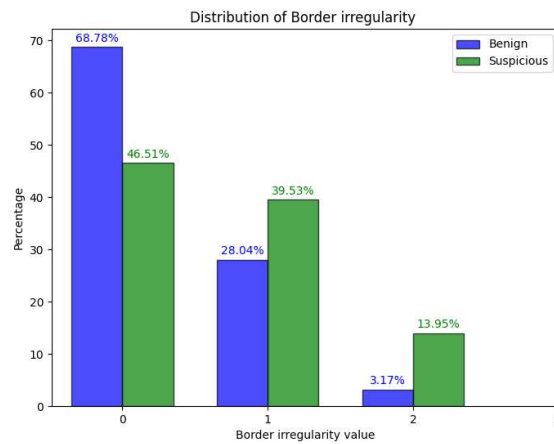


FIGURE 5.20: Distribution of clustering values for border irregularity.

suspicious lesions. The percentages delineating these distributions are visually depicted in both Figure 5.20 and Table 5.12.

This examination aimed to discern distinct patterns in border irregularity between benign and suspicious lesions based on the defined clusters. The visual representations and tabulated data provide insights into the prevalence and distribution of irregularity across these lesion categories, aiding in the identification of potential distinguishing characteristics between benign and suspicious lesions.

Cluster	Benign Lesions (%)	Suspicious Lesions (%)
0	68.78	46.51
1	28.04	39.53
2	3.17	13.95

TABLE 5.12: Distribution of clustering values for border irregularity

The statistical analyses yielded compelling insights. Levene's test for border irregularity demonstrated a substantial variance difference among clusters for both benign and suspicious lesions, producing a statistic of 14.59 and a significant p-value of 0.0001657.

Moreover, the Mann-Whitney U test for border irregularity showcased strong statistical significance, revealing a U-statistic of 10153.0 and a notably low p-value of 0.0001078. These results underscore distinct distributions of border irregularity between benign and suspicious lesions.

Overall, these findings solidify the presence of marked differences in irregularity patterns observed across diverse clusters within both benign and suspicious lesions.

5.4 Evaluation of color in benign and suspicious skin lesions

Understanding color variations within skin lesions is pivotal in diagnosis. While benign lesions commonly display a single shade of brown, melanomas often exhibit multiple shades of brown, light brown, or black. As melanomas progress, they might introduce additional colors like red, white, or blue. The presence of multiple colors within a lesion serves as a crucial indicator, prompting a more detailed examination in the diagnostic process.

Following this understanding, the results obtained from the application of the color detection methodology to images of suspicious and benign lesions are respectively detailed in Table 5.13 and Table 5.14. These tables include a cropped image of the lesion against a black background and 7 columns representing various colors: background, black, red, dark brown, light brown, blue-gray, and white. The values within each color cell denote the pixel count belonging to that color within the examined lesion. The last column presents the total count of colors, excluding the background. As explained in Chapter 4, a color is considered present only if the percentage of pixels attributed to that color exceeds 1% of the total coloured pixels of the lesion.

Among the images presented in the tables, some depict lesions with unique characteristics. For instance, among suspicious lesions, there are cases where lesions exhibit black, blue, or white coloration, sometimes featuring more than two colors. Similarly, among benign lesions, distinctive characteristics are observed, including lesions with a single color, two colors (specifically in shades of light and dark brown), and one lesion where blue-gray is also present.

It's important to note that in some analyzed images, a small portion of the outer edge might appear, possibly representing skin rather than part of the lesion. This effect could be due to imperfect segmentation. During the study, the consideration of using an erosion technique arose. However, upon implementation, it was observed that this process excessively smoothed the colors, significantly impacting the accurate identification of color variations within the images. As a result, the erosion procedure was excluded from the analysis, and the decision was made to assess the images in their original state,

ensuring a more precise evaluation of the diverse color shades present within the skin lesions.

TABLE 5.13: Results of color assessment in suspicious lesions








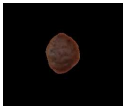

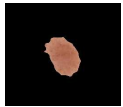

Image	Backgr.	Black	Red	Dark brown	Light brown	Blue-Gray	White	Tot
	8351	0	0	832	33	285	0	3
	3137	0	0	526	355	0	0	2
	16417	11	0	1124	520	4	0	2
	15970	0	0	0	2622	0	500	2
	2717	139	0	622	0	283	0	3
	5073	0	0	1433	55	496	0	3
	3675	246	0	359	0	0	0	2

TABLE 5.14: Results of color assessment in benign lesions

Image	Backgr.	Black	Red	Dark brown	Light brown	Blue-Gray	White	Tot
	39534	0	0	1775	1750	0	0	2
	49941	0	0	4526	0	606	0	2
	10589	11	0	1752	78	0	0	2
	53692	0	0	0	2774	0	0	1
	8954	0	0	1037	0	0	0	1
	8483	0	0	0	1718	0	0	1

Following this, for both quantitative and visual representation of the dataset's outcomes, two histograms were constructed. The first histogram (Figure 5.21) illustrates the distribution of the total number of colors within benign and suspicious lesions. The second histogram (Figure 5.22) portrays the distribution of "malignant colors," which specifically encompass black, red, blue-gray, and white. These colors are theoretically associated only with melanomas, serving as significant indicators during analysis. These histograms aim to offer a consolidated and accessible view of the color distribution patterns within the dataset, distinguishing between benign and suspicious lesions and emphasizing colors predominantly associated with melanoma lesions.

Figure 5.21 displays a histogram where each bar represents the percentage of samples within their respective groups exhibiting specific counts of distinct colors. Blue bars indicate the percentage of benign cases, while green bars represent suspicious cases. The graphic showcases notable differences in value distributions between benign and suspicious lesions.

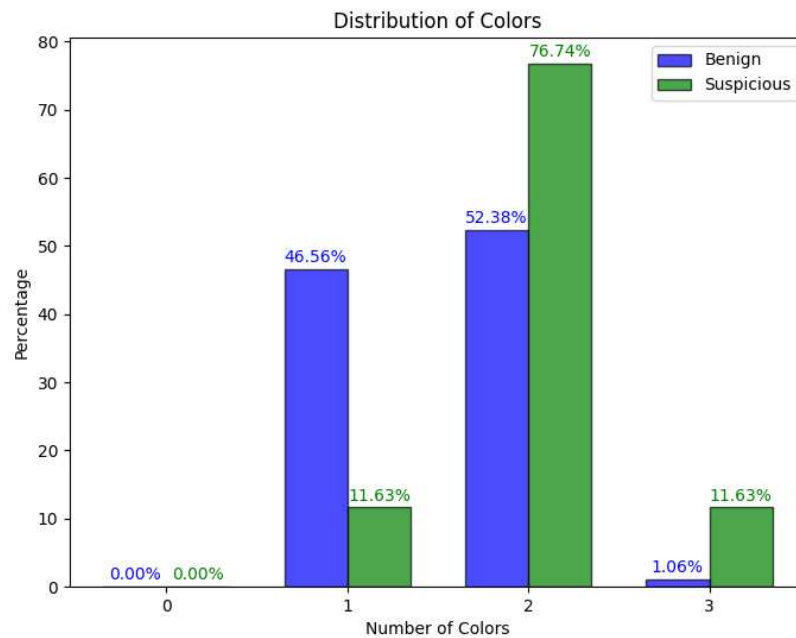


FIGURE 5.21: Color distribution in benign and suspicious lesions: Blue bars indicate the percentage of benign cases while green bars represent suspicious cases.

Tables 5.15 and 5.16 report numeric and percentage values of colors present in benign and suspicious lesions. These data offer a direct comparison of color distribution between the two lesion categories, presenting a numeric and percentage breakdown of color frequencies detected within the analyzed lesions.

TABLE 5.15: Numeric Values Table

Number of Colors	Benign	Suspicious
1	88	10
2	99	66
3	2	10

TABLE 5.16: Percentage values Table

Number of colors	Benign	Suspicious
1	46.56%	11.63%
2	52.38%	76.74%
3	1.06%	11.63%

A statistical test was conducted to determine whether a statistically significant difference exists between the distributions of color frequencies in benign and suspicious lesions. Given the non-normality observed in the distributions, the Mann-Whitney U test was employed. This test is a nonparametric method suitable for comparing two independent groups when the assumptions of normality are not met.

The Mann-Whitney U test results yielded a statistic $U = 11395.0$ and a p-value = 5.0374×10^{-10} .

A very low p-value (in this case, significantly smaller than the conventional significance level of 0.05) suggests strong evidence against the null hypothesis. Hence, the test results indicate a statistically significant difference between the distributions of color frequencies in benign and suspicious lesions. This highlights that the observed dissimilarity in color distribution between the two lesion categories is unlikely to be due to random chance but rather indicates a genuine distinction in color occurrences.

Figure 5.22 illustrates the distribution of 'malignant colors'. A substantial difference in the distribution of these malignant colors is noticeable. Particularly, for benign lesions, the presence of malignant colors is nearly negligible. Specifically Tables 5.17 and 5.18 report numeric and percentage values of malignant colors present in benign and suspicious lesions.

TABLE 5.17: Malignant colors numeric values

Number of malignant colors	Benign	Suspicious
0	182	65
1	7	18
2	0	3

TABLE 5.18: Malignant color percentage values Table

Number of malignant colors	Benign	Suspicious
0	96.30%	75.58%
1	3.70%	20.93%
2	0%	3.49%

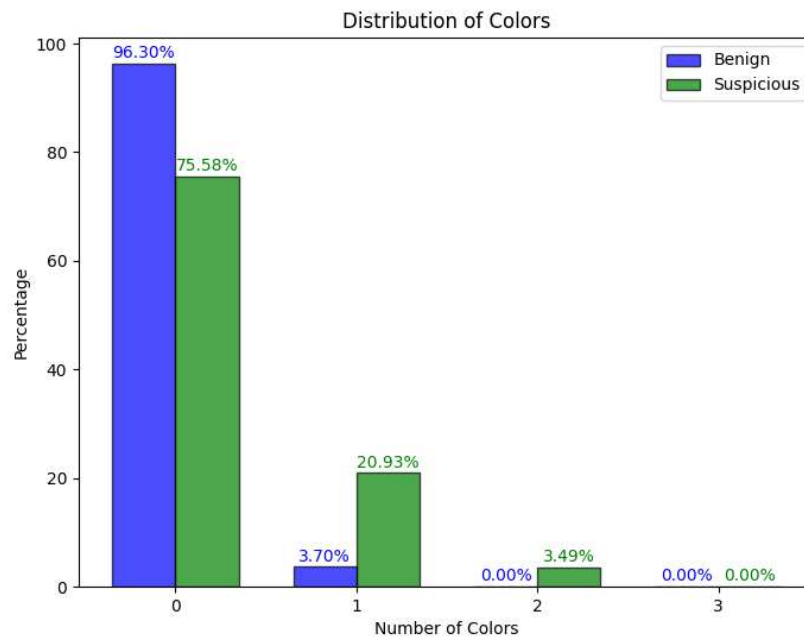


FIGURE 5.22: Malignant color distribution in benign and suspicious lesions: Blue bars indicate the percentage of benign cases while green bars represent suspicious cases.

In summary, the analysis highlights a notable distinction in the distribution of 'malignant colors,' primarily associated with melanomas, between benign and suspicious lesions. Specifically, benign lesions exhibit an insignificantly low presence of these malignant colors, implying a potential discriminatory power in differentiating between lesion types based on color profiles. This emphasizes the diagnostic utility of color analysis in discerning melanocytic lesions, wherein the presence or absence of certain colors, notably black, red, blue-gray, and white, could serve as crucial indicators for preliminary evaluation. Overall, the robust examination of color distribution offers a valuable adjunctive tool for clinical practitioners in the initial assessment and characterization of skin lesions.

Chapter 6

Conclusions

In this thesis, asymmetry, border irregularity and color of skin lesions were quantitatively assessed using macroscopic images, with the aim of addressing critical issues in current clinical practice. Asymmetry, border irregularity and color, an integral part of dermatological diagnosis, were examined on images acquired using an innovative image acquisition methodology employing a 12-camera system. The research then focused on the development of algorithms to objectively and reliably calculate these parameters, with the ultimate goal of providing an automatic tool for melanoma screening.

The obtained results demonstrate that the created algorithms yield parameters capable of quantitatively evaluating asymmetry, border irregularity and color. Concerning asymmetry, the sole use of the Non-Overlapping Area percentage allows for a progressive assessment, where an increase in %NOA corresponds to a heightened level of asymmetry. Employing k-means clustering facilitated the categorization of skin lesion masks into three categories: symmetric, moderately asymmetric, and highly asymmetric lesions.

Turning to border irregularity, the use of the Sum parameter enables a comparative evaluation between lesions, determining which exhibits greater irregularity. Consideration of the Compactness Index and Perimeter Discrepancy parameters provides a comprehensive view of border irregularity, incorporating factors such as deviation from perfect circularity and the presence of indentations of varying degrees in the border. The k-means clustering developed on these three parameters facilitated the classification of lesions into three categories of irregularity: lesions whit regular, moderately irregular,

and highly irregular borders.

In addition to analyzing asymmetry and border irregularity, this research extensively investigated color attributes within skin lesions. Employing advanced techniques, it quantified a spectrum of colors, particularly focusing on shades commonly linked to melanocytic lesions: blacks, browns, reds, blues, and whites. These colors play a vital role in identifying suspicious lesions, especially those associated with melanoma. The presence of multiple colors often signifies an advanced or malignant stage, prompting a more detailed diagnostic examination.

Furthermore, the statistical tests conducted to investigate differences between benign and suspicious lesions unveiled intriguing findings. The analyses, encompassing parameters such as %NOA, Sum, PD, CI, and color attributes, showcased distinct distributions between these two populations. Specifically, the %NOA values revealed statistically significant disparities between benign and suspicious lesions. Additionally, the PD and CI parameters exhibited marked differences, emphasizing variations in border irregularity characteristics. Notably, the color attributes analysis highlighted discernible differences in the spectrum of colors present within these lesion types. These findings underscore the potential diagnostic value of these quantitative parameters in distinguishing between benign and suspicious skin lesions.

While these tools aid dermatologists in initial lesion discrimination, they can't stand alone in melanoma definition, requiring consideration of other clinical indexes such as diameter, evolution parameters and other clinical features like personal melanoma risk and ugly duckling. The assessment of diameter faced challenges due to a lack of data, particularly regarding the precision of the theoretical distance between the patient and the camera. Low image resolution compounds these issues, emphasizing the need for precise data and improved resolution techniques.

The absence of follow-up images limits the understanding of lesion evolution, but acquiring such data could establish a longitudinal study, offering insights into parameter variations. Despite current limitations, the research lays the groundwork for future developments, aiming to refine algorithms and incorporate follow-up images to enhance early melanoma identification.

In conclusion, the quantitative analysis proves significant in dermatology, offering

effective lesion categorization. While the tools provide valuable insights, they complement, not replace, the comprehensive clinical approach of dermatologists. Addressing challenges related to resolution and follow-up datasets is crucial for refining algorithms and establishing the role of these parameters in early melanoma detection. The integration of these insights into comprehensive clinical evaluations remains essential for accurate melanoma diagnosis.

Bibliography

- [1] American Cancer Society. *Melanoma Skin Cancer*. Accessed: 2023. URL: <https://www.cancer.org/cancer/types/melanoma-skin-cancer.html>.
- [2] Istituto Superiore di Sanità. *Istituto Superiore di Sanità*. Accessed: 2023. URL: <https://www.epicentro.iss.it/melanoma/>.
- [3] Intergruppo Melanoma Italiano. *Intergruppo Melanoma Italiano*. Accessed: 2023. URL: <https://www.imi-melamed.it/home/cap-1-epidemiologia/>.
- [4] “European consensus-based interdisciplinary guideline for melanoma. Part 1: Diagnostics: Update 2022”. In: *European Journal of Cancer* 170 (2022), pp. 236–255. DOI: 10.1016/j.ejca.2022.03.008.
- [5] Robert J Friedman, Darrell S Rigel, and Alfred W Kopf. “Early detection of malignant melanoma: the role of physician examination and self-examination of the skin.” In: *CA: a cancer journal for clinicians* 35.3 (1985), pp. 130–151.
- [6] Naheed R Abbasi et al. “Early diagnosis of cutaneous melanoma: revisiting the ABCD criteria”. In: *Jama* 292.22 (2004), pp. 2771–2776.
- [7] Himanshu K Gajera, Mukesh A Zaveri, and Deepak Ranjan Nayak. “Patch-based local deep feature extraction for automated skin cancer classification”. In: *International Journal of Imaging Systems and Technology* 32.5 (2022), pp. 1774–1788.
- [8] Amanda Oakley and Tyler Stevens. *Dermoscopy*. Accessed: 2023. 2004. URL: <https://dermnetnz.org/topics/dermoscopy>.
- [9] S Chimenti et al. “Linee Guida in Dermoscopia”. In: *GIORNALE ITALIANO DI DERMATOLOGIA E VENEREOLOGIA* 140 (2005), pp. 329–348.
- [10] Vitaly Terushkin and Allan C Halpern. “Melanoma early detection”. In: *Hematology/Oncology Clinics* 23.3 (2009), pp. 481–500.

-
- [11] Amanda Levine and Orit Markowitz. “Introduction to reflectance confocal microscopy and its use in clinical practice”. In: *JAAD case reports* 4.10 (2018), pp. 1014–1023.
- [12] Lavinia Ferrante di Ruffano et al. “Computer-assisted diagnosis techniques (dermoscopy and spectroscopy-based) for diagnosing skin cancer in adults”. In: *Cochrane Database of Systematic Reviews* 2018.12 (2018).
- [13] Mahammed Messadi, Hocine Cherifi, and Abdelhafid Bessaid. “Segmentation and ABCD rule extraction for skin tumors classification”. In: *arXiv* (2021).
- [14] Abder-Rahman H Ali, Jingpeng Li, and Guang Yang. “Automating the ABCD rule for melanoma detection: a survey”. In: *IEEE Access* 8 (2020), pp. 83333–83346.
- [15] Nadia S Zghal and Nabil Derbel. “Melanoma skin cancer detection based on image processing”. In: *Current Medical Imaging* 16.1 (2020), pp. 50–58.
- [16] Adekanmi Adegun and Serestina Viriri. “Deep learning techniques for skin lesion analysis and melanoma cancer detection: a survey of state-of-the-art”. In: *Artificial Intelligence Review* 54 (2021), pp. 811–841.
- [17] Ying Wang et al. “A Collaborative Learning Model for Skin Lesion Segmentation and Classification”. In: *Diagnostics* 13.5 (2023), p. 912.
- [18] S Oukil et al. “Automatic segmentation and melanoma detection based on color and texture features in dermoscopic images”. In: *Skin Research and Technology* 28.2 (2022), pp. 203–211.
- [19] Md Kamrul Hasan et al. “A survey, review, and future trends of skin lesion segmentation and classification”. In: *Computers in Biology and Medicine* (2023), p. 106624.
- [20] Vincent Ng and David Cheung. “Measuring asymmetries of skin lesions”. In: *1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation*. Vol. 5. IEEE. 1997, pp. 4211–4216.
- [21] William V Stoecker, William Weiling Li, and Randy H Moss. “Automatic detection of asymmetry in skin tumors”. In: *Computerized Medical Imaging and Graphics* 16.3 (1992), pp. 191–197.

- [22] Arthur Tenenhaus et al. “Detection of melanoma from dermoscopic images of naevi acquired under uncontrolled conditions”. In: *Skin research and technology* 16.1 (2010), pp. 85–97.
- [23] Ihab Zaqout. “Diagnosis of skin lesions based on dermoscopic images using image processing techniques”. In: *Pattern Recognition-Selected Methods and Applications* 1320 (2019).
- [24] Reda Kasmi and Karim Mokrani. “Classification of malignant melanoma and benign skin lesions: implementation of automatic ABCD rule”. In: *IET Image Processing* 10.6 (2016), pp. 448–455.
- [25] Tim K Lee, David I McLean, and M Stella Atkins. “Irregularity index: a new border irregularity measure for cutaneous melanocytic lesions”. In: *Medical image analysis* 7.1 (2003), pp. 47–64.
- [26] Joanna Jaworek-Korjakowska et al. “Novel method for border irregularity assessment in dermoscopic color images”. In: *Computational and mathematical methods in medicine* (2015).
- [27] SH Carbonetto and SE Lew. “Characterization of border structure using fractal dimension in melanomas”. In: *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*. IEEE. 2010, pp. 4088–4091.
- [28] Li Ma et al. “Multi-scale descriptors for contour irregularity of skin lesion using wavelet decomposition”. In: *2010 3rd International Conference on Biomedical Engineering and Informatics*. Vol. 1. IEEE. 2010, pp. 414–418.
- [29] G Grammatikopoulos et al. “Simple matlab tool for automated malignant melanoma diagnosis.” In: *WSEAS Transactions on Information Science and Applications* 4.3 (2007), pp. 460–465.
- [30] Upender Kalwa et al. “Skin cancer diagnostics with an all-inclusive smartphone application”. In: *Symmetry* 11.6 (2019), p. 790.
- [31] Abder-Rahman Ali, Jingpeng Li, and Sally Jane O’Shea. “Towards the automatic detection of skin lesion shape asymmetry, color variegation and diameter in dermoscopic images”. In: *Plos one* 15.6 (2020).

-
- [32] Zhishun She, Y Liu, and A Damatoa. “Combination of features from skin pattern and ABCD analysis for lesion classification”. In: *Skin Research and Technology* 13.1 (2007), pp. 25–33.
- [33] A. Galgaro and G. Teza. “Three Dimensional Mapping for Melanoma Diagnosis”. In: *Italian Patent 102016000132357* (2018).
- [34] Philipp Tschandl. *The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions*. Version V4. 2018. DOI: 10.7910/DVN/DBW86T. URL: <https://doi.org/10.7910/DVN/DBW86T>.
- [35] OpenCV Contributors. *OpenCV Documentation*. 2023. URL: https://docs.opencv.org/3.4/d3/dc0/group__imgproc__shape.html%20%5C%5C%20#gaf259efaad93098103d6c27b9e4900ffa.

