

UNIVERSIDAD CARLOS III DE MADRID



**INGENIERÍA TÉCNICA INDUSTRIAL: ELECTRÓNICA
INDUSTRIAL Y AUTOMATIZACIÓN**

PROYECTO FIN DE CARRERA

DEPARTAMENTO DE ESTADÍSTICA

**RECONOCIMIENTO AUTOMÁTICO DE MELANOMAS
MEDIANTE TÉCNICAS DE VISIÓN POR ORDENADOR
Y RECONOCIMIENTO DE PATRONES**

AUTORA: GERTRUDIS KATERIN ASUNCIÓN BATUGO

TUTOR: DAVID DELGADO GÓMEZ

Leganés, Diciembre de 2012

Título: RECONOCIMIENTO AUTOMÁTICO DE MELANOMAS
MEDIANTE TÉCNICAS DE VISIÓN POR ORDENADOR Y
RECONOCIMIENTO DE PATRONES.

Autora: GERTRUDIS KATERIN ASUNCIÓN BATUGO.

Tutor: DAVID DELGADO GÓMEZ.

EL TRIBUNAL.

Presidente: _____

Vocal: _____

Secretario: _____

Realizado el acto de defensa y lectura del Proyecto Fin de Carrera el día ___ de _____ de 20__
en Leganés, en la Escuela Politécnica Superior de la Universidad Carlos III de Madrid, acuerda
otorgarle la CALIFICACIÓN de

VOCAL

SECRETARIO

PRESIDENTE

AGRADECIMIENTOS.

Quería dar la gracias a todas las personas que han estado presentes en mi vida y que han influido para que pudiera llegar hasta donde hoy me encuentro.

Gracias a mi tutor, David Delgado Gómez, por haberme permitido realizar este proyecto y por su dedicación a lo largo del mismo.

Gracias a mi familia, principalmente a mis padres, Abraham y Hernanie, que siempre me han brindado su apoyo incondicional, además de su paciencia y comprensión a lo largo de toda la carrera.

Gracias al grupo de amigos del colegio, con los cuales he compartido muy buenos momentos y quienes fundamentalmente hayan influido en mis ganas de seguir estudiando.

Gracias a los amigos y sobre todo a las amigas de la universidad, tanto de la técnica como de la superior, por todas esas horas dentro y fuera del campus.

Y en especial a David Tena, gracias por esa bendita paciencia que tienes conmigo, especialmente en estos últimos meses y por todo el cariño que me tienes.

RESUMEN.

El objetivo de este Proyecto Fin de Carrera consiste en clasificar cualquier lesión sospechosa de ser melanoma entre dos grupos. Un grupo correspondiente a las lesiones con cáncer de piel de tipo no melanoma, es decir, lesiones benignas y el otro grupo correspondiente a las lesiones con cáncer de piel de tipo melanoma, es decir, lesiones malignas. Para entender cómo se forman y detectan los melanomas y conocer el tratamiento necesario que se debe aplicar en caso de poseer melanomas, se presenta una breve introducción al melanoma.

Dado lo importante que es detectar el melanoma a tiempo, en este proyecto se propone realizar la clasificación mediante métodos de visión por ordenador, es decir, gracias a imágenes digitales. Para ello, se capturan y digitalizan las lesiones gracias a un dermatoscopio manual, que al ser digital, permite añadir una cámara digital. Además, el software empleado para acondicionar y analizar las imágenes digitales capturadas es el software de Matlab. Este software permite crear un programa automático y sencillo que maneja las imágenes.

Los principales pasos del programa desarrollado son:

- El acondicionamiento y la segmentación (con preprocesado y postprocesado), que consisten en manipular la imagen matemáticamente, de forma que la lesión resulte separada de la piel sana.
- La extracción de características, que consiste en determinar los parámetros a partir de la imagen de la lesión y de la segmentación de la misma, basándose en características que definen a los melanomas y que siguen la regla ABCD de la dermatoscopia (Asimetría, Borde, Color y Diferencias estructurales).
- La clasificación de las lesiones, que consiste en determinar si una lesión es un melanoma o no, a partir de las características extraídas previamente.

En este documento se mencionan numerosos métodos para realizar cada uno de los pasos anteriores. Pero se exponen y detallan: el método de captura, la digitalización, el acondicionamiento, la segmentación, la extracción de características y la clasificación escogidos para aplicarse en una serie de lesiones que forman nuestra base de datos y de los cuales obtenemos unos resultados que son mostrados y explicados al final de la memoria.

Posteriormente, se exponen unas breves conclusiones y trabajos futuros y finalmente, se muestra el presupuesto que costaría llevar a cabo este proyecto.

Palabras clave: melanoma, visión por ordenador, dermatoscopio digital, Matlab, segmentación, regla ABCD de la dermatoscopia, clasificación.

ABSTRACT.

The aim of this Thesis or PhD is to classify any suspicious lesion between two groups. One of them corresponds to cancer lesions of non – melanoma skin, called benign pigmented lesions and the other group corresponds to lesions of skin cancer melanoma, called malignant melanomas or melanomas. In order to understand how melanomas are formed and detected, and also to know the treatment that is necessary to be applied in case of having melanomas, it is presented a brief introduction to melanoma.

Detect an early melanoma is very important, so in this project, it is intended to make the classification by computer vision methods, i.e. by digital images. For this purpose, the lesions are captured and digitalized through a digital manual dermatoscope adding a digital camera. In addition, the software used to improve the conditions and to analyze the digital images is Matlab, because this software allows creating a simple automatic program that handles images.

The main steps of the developed program are:

- Conditioning and segmentation (with preprocessing and postprocessing), in these steps the program manipulate the image mathematically, so that the result is an image where lesion is separated from healthy skin.
- The feature extraction, in this step the program determines parameters from the image and the segmentation of the lesion, based on characteristics that define melanomas. These characteristics are in the ABCD rule of dermatoscopy (Asymmetry, Border, Color and texture).
- The classification of lesions, in this step it is determined whether a lesion is a melanoma or not, from the previously extracted features.

This document mentions several methods for each of the above steps. But are presented and detailed: the method of capture, digitalization, conditioning, segmentation, feature extraction and classification chosen to apply in a set of lesions that are our database and they give us a results which are shown and explained at the end of the memory.

Finally, it is presented some brief conclusions and future works and as a final point of this document, it is shown the budget that would cost to develop this project.

Keywords: melanoma, computer vision, digital dermatoscope, Matlab, segmentation, ABCD rule of dermatoscopy, classification.

ÍNDICE GENERAL.

AGRADECIMIENTOS.	III
RESUMEN.	IV
ABSTRACT.	V
ÍNDICE GENERAL.	1
ÍNDICE DE FIGURAS.	3
ÍNDICE DE ECUACIONES.	5
ÍNDICE DE TABLAS.	6
<u>CAPÍTULO 1. INTRODUCCIÓN Y OBJETIVOS.</u>	7
1.1. INTRODUCCIÓN AL MELANOMA.	8
1.2. DETECCIÓN Y TRATAMIENTO DEL MELANOMA.	10
1.3. OBJETIVOS Y FASES DE DESARROLLO.	17
1.4. ESTRUCTURA DE LA MEMORIA.	18
<u>CAPÍTULO 2. ESTADO DEL ARTE.</u>	19
2.1. CAPTURA Y ACONDICIONAMIENTO DE LA IMAGEN.	21
2.2. TÉCNICAS DE SEGMENTACIÓN DE LA IMAGEN.	23
2.3. EXTRACCIÓN DE CARACTERÍSTICAS.	25
2.3.1. REGLA ABCD APLICADA A SIMPLE VISTA [9] [14] [35].	26
2.3.2. REGLA ABCD DE LA DERMATOSCOPIA [14] [36].	26
2.3.3. MÉTODO “NEW 7 – POINT CHECKLIST” [38].	27
2.3.4. OTRAS CARACTERÍSTICAS PROPUESTAS.	28
2.4. MÉTODOS DE CLASIFICACIÓN [10].	30
2.4.1. CLASIFICADORES MÁS COMUNES.	30
2.4.2. EVALUACIÓN DE UN CLASIFICADOR.	32
<u>CAPÍTULO 3. CAPTURA Y ACONDICIONAMIENTO DE LA IMAGEN.</u>	33
3.1. CAPTURA DE LA IMAGEN DE LA LESIÓN.	34
3.2. ACONDICIONAMIENTO DE LA IMAGEN DE LA LESIÓN.	34
<u>CAPÍTULO 4. TÉCNICAS DE SEGMENTACIÓN DE LA IMAGEN.</u>	37
4.1. COMBINACIÓN DE IMÁGENES RGB.	38
4.1.1. MEZCLA DE GAUSSIANAS.	39
4.1.2. OBTENCIÓN DE LOS PARÁMETROS DE LA COMBINACIÓN ADECUADA DE LA IMAGEN.	41
4.2. IMAGEN SEGMENTADA (PREVIA Y DEFINITIVA).	43
4.2.1. DEFINICIÓN DE LA MÁSCARA BINARIA.	45
4.2.2. ELIMINACIÓN DE LOS ARTIFICIOS QUE NO PERTENECEN A LA LESIÓN.	46
4.2.3. RELLENADO DE LOS HUECOS QUE SE ENCUENTRAN EN EL INTERIOR DE LA LESIÓN.	48

CAPÍTULO 5. EXTRACCIÓN DE CARACTERÍSTICAS.	50
5.1. ANÁLISIS DE LA ASIMETRÍA DE LA LESIÓN.	51
5.2. ANÁLISIS DEL BORDE DE LA LESIÓN.	53
5.3. ANÁLISIS DEL COLOR Y DE LAS DIFERENTES ESTRUCTURAS DE LA LESIÓN.	59
5.4. ANÁLISIS DE LOS PÍXELES CONTENIDOS EN LA LESIÓN.	59
CAPÍTULO 6. TÉCNICAS DE CLASIFICACIÓN Y EVALUACIÓN.	62
6.1. CLASIFICACIÓN DE LAS CARACTERÍSTICAS MEDIANTE UN CONTRASTE DE HIPÓTESIS.	63
6.2. CLASIFICACIÓN DE LAS LESIONES MEDIANTE UN DISCRIMINANTE LINEAL [11] [30].	63
6.2.1. CLASIFICACIÓN DE LAS LESIONES CON TODAS LAS CARACTERÍSTICAS OBTENIDAS.	64
6.2.2. CLASIFICACIÓN DE LAS LESIONES CON CARACTERÍSTICAS SELECCIONADAS DEL CONJUNTO TOTAL DE CARACTERÍSTICAS.	64
6.3. EVALUACIÓN DE LAS LESIONES MEDIANTE VALIDACIÓN CRUZADA.	66
6.4. EVALUACIÓN DEL CLASIFICADOR.	66
CAPÍTULO 7. RESULTADOS DE LOS EXPERIMENTOS.	68
7.1. BASE DE DATOS.	69
7.2. RESULTADO DE LA CLASIFICACIÓN DE LAS CARACTERÍSTICAS EXTRAÍDAS.	69
7.3. RESULTADO DE LA CLASIFICACIÓN DE LESIONES.	73
7.4. RESULTADO DE LA EVALUACIÓN DEL CLASIFICADOR.	73
CAPÍTULO 8. CONCLUSIONES Y TRABAJOS FUTUROS	75
CAPÍTULO 9. PRESUPUESTO.	77
9.1. COSTE DEL MATERIAL.	78
9.1.1. COSTE DEL MATERIAL EN HARDWARE.	78
9.1.2. COSTE DEL MATERIAL EN SOFTWARE.	78
9.2. COSTE DEL PERSONAL.	78
9.2.1. COSTE DEL PERSONAL EN HORAS DE INGENIERÍA.	78
9.2.2. COSTE DEL PERSONAL EN HORAS DE REDACCIÓN.	79
9.3. COSTES TOTALES.	80
CAPÍTULO 10. REFERENCIAS.	81

ÍNDICE DE FIGURAS.

Figura 1 – Estructura interna de la piel cuando está sana [6].	8
Figura 2 – Estructura interna de la piel cuando existe un melanoma en etapa o estadio 0 [6].	12
Figura 3 – Estructura interna de la piel cuando existe un melanoma en etapa o estadio I [6].	12
Figura 4 – Estructura interna de la piel cuando existe un melanoma en etapa o estadio II [6].	13
Figura 5 – Estructura interna de la piel cuando existe un melanoma en etapa o estadio III [6].	14
Figura 6 – Estructura interna de la piel cuando existe un melanoma en etapa o estadio IV [6].	15
Figura 7 – Microscopio: Real [13].	19
Figura 8 – Microscopio: Esquema [13].	19
Figura 9 – Estereomicroscopio: Esquema [15].	20
Figura 10 – Estereomicroscopio: Real [16].	20
Figura 11 – Dermatoscopio manual: Diferentes tipos [18] [19].	20
Figura 12 – Videomicroscopio: Real [22].	21
Figura 13 – Videomicroscopio: Esquema [23].	21
Figura 14 – Videomicroscopio: Esquema completo [21].	22
Figura 15 – Equipo Mole Max II [25].	22
Figura 16 – Tabla de contingencia o Matriz de confusión.	32
Figura 17 – Lesión 1: No melanoma.	34
Figura 18 – Lesión 2: Melanoma.	34
Figura 19 – Lesión 3: Melanoma.	34
Figura 20 – Lesión 4: No melanoma.	34
Figura 21 – Lesión 5: No melanoma.	34
Figura 22 – Lesión 6: Melanoma.	34
Figura 23 – Lesión 7: Imagen original.	35
Figura 24 – Lesión 7: Imagen original recortada.	35
Figura 25 – Lesión 8: Imagen original.	35
Figura 26 – Lesión 8: Imagen original recortada.	35
Figura 27 – Lesión 9: Imagen original recortada.	36
Figura 28 – Lesión 9: Imagen original recortada y sampleada.	36
Figura 29 – Lesión 10: Imagen original recortada.	38
Figura 30 – Lesión 10: Banda roja.	38
Figura 31 – Lesión 10: Banda verde.	38
Figura 32 – Lesión 10: Banda azul.	38
Figura 33 – Lesión 11: Imagen original recortada.	38
Figura 34 – Lesión 11: Banda roja.	38
Figura 35 – Lesión 11: Banda verde.	38
Figura 36 – Lesión 11: Banda azul.	38
Figura 37 – Representación gráfica de la relación entre coordenadas esféricas y coordenadas cartesianas [10].	40
Figura 38 – Lesiones 12 y 13: Combinación cualquiera cuyo histograma diferencia poco la lesión de la piel.	41
Figura 39 – Lesiones 12 y 13: Combinación cualquiera cuyo histograma diferencia muy bien la lesión de la piel.	41
Figura 40 – Representación gráfica tridimensional de las posibles combinaciones.	42
Figura 41 – Lesión 14: Imagen original.	43
Figura 42 – Lesión 14: Imagen recortada.	43
Figura 43 – Lesión 14: Imagen con mejor contraste.	43
Figura 44 – Lesión 15: Imagen original.	43
Figura 45 – Lesión 15: Imagen recortada.	43
Figura 46 – Lesión 15: Imagen con mejor contraste.	43

ÍNDICE DE FIGURAS.

Figura 47 – Lesión 16: Imagen original recortada.	44
Figura 48 – Lesión 16: Imagen segmentada (algoritmo de Otsu).	44
Figura 49 – Lesión 16: Imagen segmentada (Matlab, función "cluster").	44
Figura 50 – Lesión 17: Imagen original recortada.	44
Figura 51 – Lesión 17: Imagen segmentada (algoritmo de Otsu).	44
Figura 52 – Lesión 17: Imagen segmentada (Matlab, función "cluster").	44
Figura 53 – Lesión 18: Segmentada previa.	45
Figura 54 – Lesión 19: Segmentada previa.	45
Figura 55 – Lesión 18: Segmentada previa dividida en nueve partes.	45
Figura 56 – Lesión 19: Segmentada previa divide en nueve partes.	45
Figura 57 – Lesión 18: Segmentada con máscara binaria.	46
Figura 58 – Lesión 19: Segmentada con máscara binaria.	46
Figura 59 – Lesión 18: Segmentada con máscara binaria.	47
Figura 60 – Lesión 18: Segmentada con máscara binaria y artificios.	47
Figura 61 – Lesión 19: Segmentada con máscara binaria.	47
Figura 62 – Lesión 19: Segmentada con máscara binaria y artificios.	47
Figura 63 – Lesión 18: Segmentada con máscara binaria, sin artificios y con huecos.	48
Figura 64 – Lesión 18: Segmentada definitiva.	48
Figura 65 – Lesión 19: Segmentada con máscara binaria, sin artificios y con huecos.	48
Figura 66 – Lesión 19: Segmentada definitiva.	48
Figura 67 – Lesión 18: Imagen original recortada.	49
Figura 68 – Lesión 18: Imagen segmentada definitiva.	49
Figura 69 – Lesión 19: Imagen original recortada.	49
Figura 70 – Lesión 19: Imagen segmentada definitiva.	49
Figura 71 – Lesión 20: Imagen original recortada.	52
Figura 72 – Lesión 20: Imagen segmentada definitiva.	52
Figura 73 – Lesión 20: Imagen segmentada definitiva con ejes perpendiculares.	52
Figura 74 – Lesión 20: Histograma del eje principal.	52
Figura 75 – Lesión 20: Histograma del eje secundario.	52
Figura 76 – Lesión 21: Imagen de la segmentada definitiva indicando los 36 ejes tomados.	53
Figura 77 – Lesión 22: Imagen de la segmentada definitiva indicando los 36 ejes tomados.	53
Figura 78 – Lesión 21: Banda roja de la lesión.	54
Figura 79 – Lesión 21: Banda verde de la lesión.	54
Figura 80 – Lesión 21: Banda azul de la lesión.	54
Figura 81 – Lesión 21: Mejor combinación o imagen proyectada.	54
Figura 86 – Representación gráfica de la sigmoidea genérica [10].	56
Figura 82 – Lesión 22: Banda roja de la lesión.	56
Figura 83 – Lesión 22: Banda verde de la lesión.	56
Figura 84 – Lesión 22: Banda azul de la lesión.	56
Figura 85 – Lesión 22: Mejor combinación o imagen proyectada.	56
Figura 87 – Lesión 22: Píxeles en banda roja.	57
Figura 88 – Lesión 22: Píxeles en banda verde.	57
Figura 89 – Lesión 22: Píxeles en banda azul.	57
Figura 90 – Lesión 22: Píxeles de la mejor combinación o imagen proyectada.	57

ÍNDICE DE ECUACIONES.

<i>Ecuación 1 – Luminancia estándar.</i>	23
<i>Ecuación 2 – Método de Otsu: cálculo de la varianza intragrupal.</i>	24
<i>Ecuación 3 – Método de Otsu: cálculo de la varianza entre grupos.</i>	24
<i>Ecuación 4 – Regla ABCD de la dermatoscopia.</i>	26
<i>Ecuación 5 – XOR.</i>	32
<i>Ecuación 6 – Valor predictivo positivo.</i>	32
<i>Ecuación 7 – Valor predictivo negativo.</i>	32
<i>Ecuación 8 – Análisis para encontrar la mejor combinación lineal para proyectar las bandas RGB.</i>	39
<i>Ecuación 9 – Relación entre coordenadas esféricas y coordenadas cartesianas.</i>	40
<i>Ecuación 10 – Valores que toman las variables esféricas.</i>	40
<i>Ecuación 11 – Valores simplificados que toman las variables esféricas en el programa.</i>	40
<i>Ecuación 12 – Coeficiente de asimetría de Fisher o índice de Fisher.</i>	42
<i>Ecuación 13 – Valores φ, θ, S de una lesión cualquiera.</i>	42
<i>Ecuación 14 – Valores x, y, z de una lesión cualquiera.</i>	43
<i>Ecuación 15 – Sigmoidea genérica.</i>	56
<i>Ecuación 16 – Entropía de información.</i>	60
<i>Ecuación 17 – Precisión.</i>	66
<i>Ecuación 18 – Sensibilidad.</i>	66
<i>Ecuación 19 – Especificidad.</i>	67

ÍNDICE DE TABLAS.

<i>Tabla 1 – Características obtenidas analizando la asimetría de la lesión.</i>	<i>69</i>
<i>Tabla 2 – Características obtenidas analizando el borde de la lesión mediante el ajuste por regresión lineal.</i>	<i>70</i>
<i>Tabla 3 – Características obtenidas analizando el borde de la lesión mediante el ajuste por sigmoidea.</i>	<i>71</i>
<i>Tabla 4 – Características obtenidas analizando los píxeles contenidos en la lesión.</i>	<i>72</i>
<i>Tabla 5 – Características obtenidas analizando todos los píxeles de la imagen de la lesión.</i>	<i>73</i>
<i>Tabla 6 – Verdadero/Falso Positivo/Negativo del clasificador discriminante lineal.</i>	<i>73</i>
<i>Tabla 7 – Sensibilidad, Especificidad y Precisión del clasificador discriminante lineal.</i>	<i>74</i>
<i>Tabla 8 – Características más empleadas en la selección de características.</i>	<i>74</i>
<i>Tabla 9 – Coste del material en hardware.</i>	<i>78</i>
<i>Tabla 10 – Coste del material en software.</i>	<i>78</i>
<i>Tabla 11 – Costes del personal en ingeniería.</i>	<i>79</i>
<i>Tabla 12 – Costes del personal en redacción.</i>	<i>79</i>
<i>Tabla 13 – Costes totales.</i>	<i>80</i>
<i>Tabla 14 – Presupuesto final del proyecto.</i>	<i>80</i>

Capítulo 1.

INTRODUCCIÓN Y

OBJETIVOS.

El melanoma es un tipo de cáncer de piel que conforma menos del 5% de los cánceres de piel [1] [2]. Aunque no es el cáncer de piel más común, es el que produce un mayor número de muertes y durante los últimos años, estas cifras han aumentado. Desde un punto de vista económico, se invierte una gran cantidad de recursos en detectar y tratar pacientes con este tumor. Por ejemplo, en Estados Unidos, únicamente en 1992, se gastó más de un billón de dólares en la realización de biopsias [3]. Los cálculos más recientes de la Sociedad Americana Contra El Cáncer en cuanto al melanoma indican que para el 2012, en Estados Unidos [1] [2]:

- Serán diagnosticados aproximadamente 76.250 nuevos casos de melanoma (alrededor de 44.250 en hombres y 32.000 en mujeres).
- Y a causa del melanoma morirán aproximadamente 9.180 personas (6.060 hombres y 3.120 mujeres).

La frecuencia del melanoma es más de 10 veces mayor en estadounidenses de piel blanca que en estadounidenses de raza negra. Aunque para las personas menores de 40 años, las tasas de incidencia son mayores en las mujeres que en los hombres, después de esta edad, las tasas son casi dos veces mayores en los hombres que en las mujeres. En general, el riesgo de padecer melanoma en el transcurso de la vida es de aproximadamente 2% (1 en 50) para las personas de piel blanca, 0.1% (1 en 1.000) para las personas de piel negra y 0.5% (1 en 200) para los hispanos.

Contrario a los otros cánceres comunes, el melanoma tiene una distribución por edades amplia, presentándose tanto en personas jóvenes como en personas de avanzada edad. El melanoma es uno de los cánceres más comunes en adultos jóvenes (menores de 30 años), además las tasas aumentan con la edad y son más altas en personas de entre 80 y 89 años.

En España, en los últimos años, la incidencia del melanoma maligno ha aumentado de forma importante. La tasa de melanoma se ha multiplicado por 3,3 en varones y por 2,5 en mujeres, durante los últimos veinte años. A pesar de esto, representa menos del 3% de todos los tumores y las campañas de diagnóstico precoz han permitido que la mortalidad se reduzca un 30% desde los años setenta [4].

En este capítulo se explica una breve introducción al melanoma, destacando:

- La función importante de la piel.
- Los tipos de tumores o lesiones que existen, malignos y benignos.
- Los factores de riesgo que pueden influir en la aparición de melanomas.
- Los síntomas presentes en un melanoma y maneras de detectarlo.
- Y los tratamientos existentes para eliminar un melanoma según su estadio.

Tras la introducción, se presentan los objetivos y fases del desarrollo del proyecto y la estructura de la memoria.

1.1. INTRODUCCIÓN AL MELANOMA.

Como ya se ha comentado anteriormente, el melanoma es un tipo de cáncer de piel. La piel es el órgano más grande del cuerpo y se encarga de protegerlo frente al calor, la luz, las infecciones y las lesiones. La piel está constituida por dos capas principales: epidermis (capa externa de la piel) y dermis (capa interna de la piel) [5]. En la Figura 1 se puede observar que en la epidermis se encuentran numerosas células, como son las células escamosas, las células basales o las células llamadas melanocitos. Estas últimas son las encargadas de producir una sustancia llamada melanina, que es el pigmento que da color a la piel.

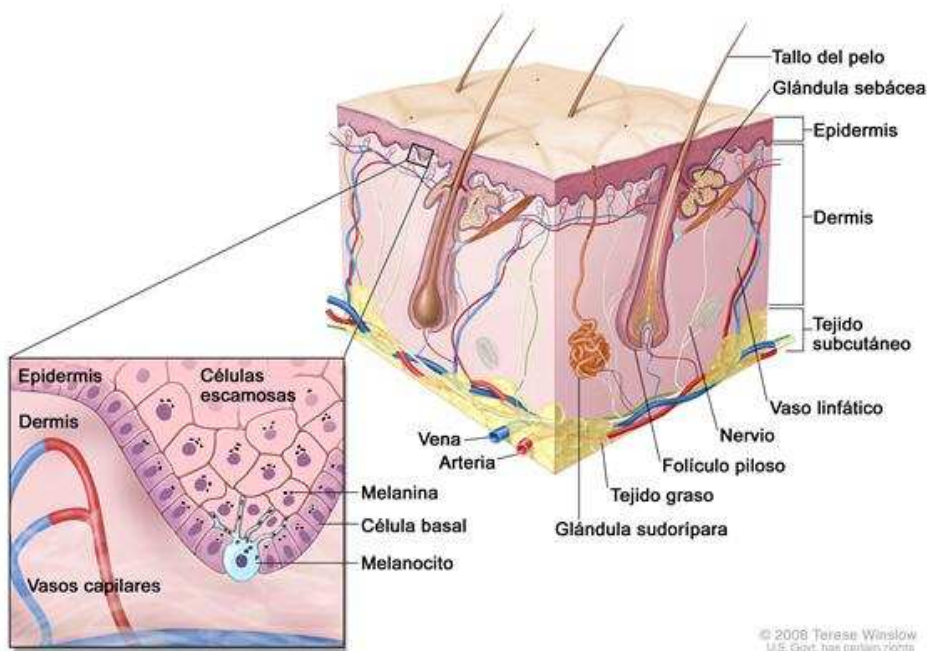


Figura 1 – Estructura interna de la piel cuando está sana [6].

Normalmente, estas células crecen, se dividen y mueren. Pero, a veces, las células mutan (cambian) y comienzan a crecer y dividirse con mayor rapidez que las células normales. En lugar de morir, estas células anormales se aglutinan dando lugar a tumores [7]. Dependiendo del tipo de células que evolucionan anormalmente, se tienen dos tipos de cáncer de piel. El primero es el cáncer de piel de tipo no melanoma o benigno, producido por el desarrollo irregular de las células de la piel que no acumulan pigmento (como las escamosas o basales), es el más frecuente y tiene la característica de no expandirse. Por otro lado, se encuentra el cáncer de piel de tipo melanoma o maligno, es menos frecuente, más peligroso y se forma a partir de células de la piel que acumulan pigmento (melanocitos). Este tipo de tumores son cancerosos y pueden propagarse y formar nuevos tumores en otras partes del cuerpo, presentando la llamada metástasis.

Este proyecto se centra en el estudio del cáncer de piel de tipo melanoma, también llamado melanoma maligno o melanoma cutáneo [2]. Cuando las células llamadas melanocitos se amontonan en una zona, producen agrupaciones en forma de bultitos sobre la piel que son comúnmente conocidos como lunares (o, en términos médicos, nevus). Por lo general, no causan ningún tipo de problemas y pueden aparecer en cualquier parte del cuerpo. Sin embargo, debido a un cambio genético o mutación, estos lunares recientes o existentes pueden comenzar a crecer de forma desmesurada dando lugar a los llamados melanomas que dañan las células sanas y el tejido que se encuentra a su alrededor [8].

El melanoma como la mayoría de cánceres, detectado a tiempo, es decir, cuando todavía se encuentra en la superficie de la piel, en la epidermis, se puede curar. Pero si no es así, puede propagarse hacia la dermis, diseminarse por metástasis rápidamente a otras partes del cuerpo hasta alcanzar los vasos sanguíneos y el sistema linfático. Estos dos sistemas, se encuentran repartidos por todo el cuerpo y pueden permitir que las células cancerosas accedan fácilmente a órganos alejados como los pulmones o el cerebro, llegando a provocar la muerte. Es por eso que la detección temprana es sumamente importante.

Aunque existen muchos factores que desencadenan la creación de melanomas, el principal factor es la alta exposición de la piel a los rayos ultravioleta del tipo A o del tipo B, conocidos como rayos UVA o UVB, procedentes de los rayos solares o de los centros de bronceado [7]. Es por esto que, los melanomas pueden afectar a cualquier persona, pero existen personas que tienen mayor probabilidad de presentar melanomas debido a su color de piel, edad, antecedentes familiares, etc. Algunos de los factores que favorecen la aparición de melanomas [2] [4] [5] [7] [8] [9], se mencionan a continuación, y son los llamados factores de riesgo. Estos factores de riesgo, indican que las personas que poseen estas características tienen más probabilidad de presentar melanomas:

- Tener la piel clara, los ojos claros (azules, verdes o grises) o el pelo claro (rubio o pelirrojo).
- Tener una piel propensa a sufrir quemaduras o en la que aparecen pecas con facilidad.
- Tener una piel sensible al sol y dificultad para broncearse.
- Tener lunares, pecas o manchas de nacimiento.
- Tener muchos lunares en todo el cuerpo (más de 50 lunares de cualquier tamaño).
- Tener cierto tipo de lunares (displásticos y atípicos) o múltiples marcas de nacimiento.
- Tener muchos lunares más grandes que el borrador de un lápiz (6 mm).
- Padecer de xeroderma pigmentosum.

- Haberse expuesto mucho al sol de niño.
- Haber sufrido quemaduras solares graves (con ampollas) durante la niñez.
- Haber tenido cáncer de piel o melanomas anteriormente.
- Tener un sistema inmunitario debilitado debido a enfermedades y medicamentos, es decir, padecer de inmunosupresión.
- Tener una edad avanzada.
- Tener antecedentes familiares que hayan tenido: un cáncer de piel, lunares irregulares, lunares cancerosos o melanomas.

- Vivir en climas con sol o a grandes alturas.
- Vivir en el Ecuador, donde los rayos solares inciden con mayor intensidad.
- Exponerse de forma prolongada o intensa a altos niveles de luz solar intensa, debido al trabajo o a otras actividades (como jugar al aire libre, estar de vacaciones, usar dispositivos de bronceado o lámparas solares...).
- Exponerse a rayos ultravioletas (rayos UV) solares o de cama solar.
- Exponerse a agentes químicos que puedan causar cáncer como el arsénico, el alquitrán de hulla o la creosota.

Los melanomas no discriminan, es decir, una persona que no cumpla algunas de las características mencionadas anteriormente, no quiere decir que esté exento de la aparición de melanomas, si no que tiene menos probabilidades de presentar melanomas, pero no de librarse de ellos. Actualmente:

- Este tipo de cáncer de piel ya no se considera una enfermedad que afecta exclusivamente a personas mayores. Debido a que el porcentaje de gente afectada por esta enfermedad está en ascenso, especialmente entre adolescentes y jóvenes de veinte a treinta años. Las investigaciones han demostrado que el aumento puede estar relacionado en parte con la exposición a los rayos ultravioletas (UV) del sol durante la niñez. Sólo en los Estados Unidos, el porcentaje de las personas que desarrollan un melanoma ha aumentado más del doble en las últimas tres décadas [8].

- Además de que la aparición de melanomas esté en aumento en jóvenes, las personas sin antecedentes familiares también pueden desarrollarlos. Y en la población negra u oriental, donde el melanoma es muy raro, cuando se da, aparece en zonas poco pigmentadas como en ojos, uñas, palmas de las manos o plantas de los pies, siendo su pronóstico aún peor [4].

1.2. DETECCIÓN Y TRATAMIENTO DEL MELANOMA.

Generalmente, la detección de melanoma se realiza a través de una inspección visual de la piel desarrollada por un especialista. El dermatólogo o médico examina la lesión con el objetivo de determinar si se encuentran presentes determinados patrones que suelen caracterizar al melanoma. Entre estos se encuentran determinados patrones de Asimetría, Borde irregular, cambios de Color y Diámetro (la llamada regla ABCD de la dermatología). Un lunar normalmente es simétrico, con un borde regular, un color homogéneo y pequeño. En cambio, los melanomas suelen ser asimétricos, presentan irregularidades en su borde, poseen diferentes colores y tienen un diámetro superior a 6 mm (aproximadamente el tamaño de un borrador de lápiz).

Cuando el especialista sospecha de la existencia de alguna anomalía, realizará una prueba más exacta, que consiste en la biopsia de la zona sospechosa. Con la biopsia se analiza el tejido bajo el microscopio para examinar el tipo de células que presenta. Dependiendo de donde se encuentre la alteración de la piel y del tipo que sea, se realizará un tipo de biopsia u otro [4]. Por tanto se tienen varios tipos de biopsia, de las que se explican las más comunes en caso de sospechar que el paciente presente un melanoma:

- Biopsia por raspado: se insensibiliza la zona de piel a biopsiar con anestesia local y se raspan las capas superiores de la piel con un bisturí.
- Biopsias incisionales y escisionales: se extrae una cuña de piel. Se realiza para tumores más profundos. Con la biopsia incisional, sólo se extrae una parte del tumor para analizarlo. Con la biopsia excisional, se extirpa todo el tumor. Si la extensión de la piel afectada es grande, se realizará una biopsia incisional, en un primer momento, para no deteriorar mucho la imagen de la persona.
- Biopsia por aspiración con aguja fina: se utiliza una jeringuilla con una aguja fina para extraer pequeñas partículas del tumor. No se utiliza para el diagnóstico de un lunar sospechoso pero sí para biopsiar los ganglios linfáticos cercanos a un melanoma.

Cuando se sospecha que el estadio del cáncer es alto, se realizarán otras pruebas para el diagnóstico de la extensión como escáneres, análisis y pruebas de medicina nuclear, entre otros.

Gracias a la biopsia, se puede conocer el tipo de cáncer de piel que tiene una persona. Una vez determinado el tipo de cáncer de piel y en caso de ser un melanoma maligno, es importante conocer en qué etapa se encuentra para poder tratarlo.

Tras confirmar la existencia de un melanoma, el tratamiento depende del tamaño y la profundidad de la lesión, en qué parte del cuerpo se encuentra y si se ha propagado a otras partes. Este proceso se denomina clasificación por etapas o estadios (estadificación) [2]. A pesar de que se trata de un proceso complicado que incluye tanto exámenes físicos como de laboratorio, la premisa básica es que cuanto más bajo es el estadio, mejor será el pronóstico [4]. Existen muchas maneras de clasificar los melanomas por etapas. En este documento se explica una de las más sencillas, la llamada clasificación de Clark [10], que clasifica al melanoma por nivel de invasión y que consta de las siguientes etapas o estadios:

- Etapa o estadio 0:

En esta etapa, como muestra la Figura 2, el melanoma se encuentra en la epidermis, es decir, que las células anormales se encuentran solamente en la capa exterior de las células de la piel y no invaden los tejidos más profundos [2].

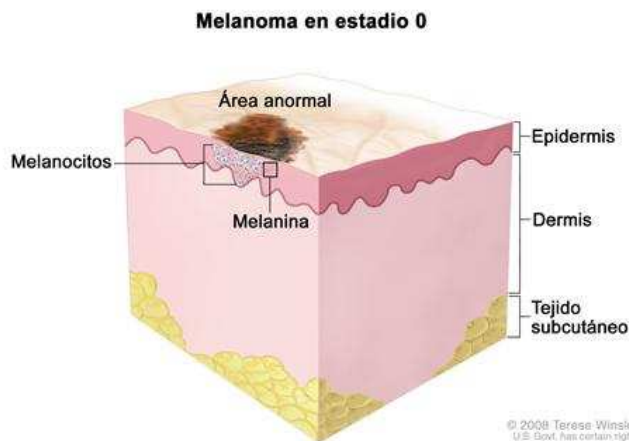


Figura 2 – Estructura interna de la piel cuando existe un melanoma en etapa o estadio 0 [6].

Por lo general, el tratamiento consiste en extirpar el melanoma con cirugía tomando un margen de aproximadamente de 5 mm de piel sana.

El índice de supervivencia cuando el melanoma es detectado en esta etapa es del 99%.

- Etapa o estadio I:

En esta etapa, como se indica en la Figura 3, el melanoma se encuentra en la epidermis y/o en la parte superior de la dermis, pero no se ha diseminado a los ganglios linfáticos vecinos. El tumor tiene un grosor de menos de 2 mm y es todavía un tumor de bajo riesgo [2].

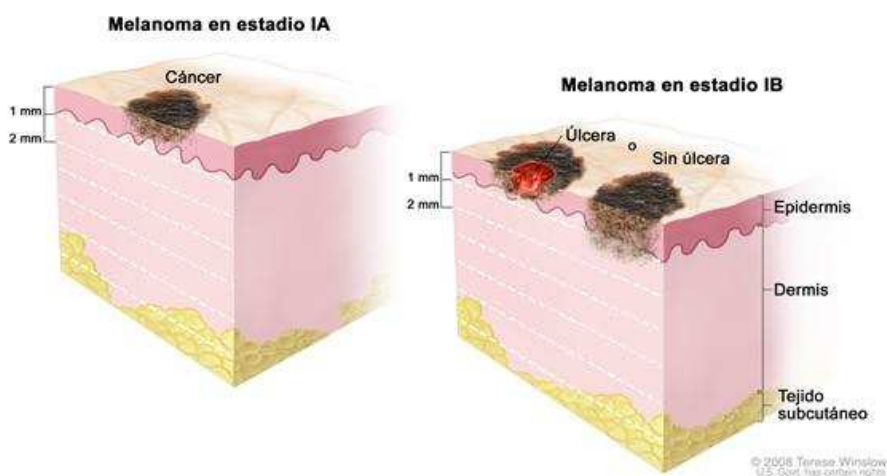


Figura 3 – Estructura interna de la piel cuando existe un melanoma en etapa o estadio I [6].

Al igual que en la etapa anterior, el tratamiento consiste en extirpar el melanoma con cirugía tomando un margen de piel sana, que depende del grosor del melanoma.

A veces se recomienda realizar una biopsia de ganglio centinela, especialmente si el melanoma es sospechoso de propagarse a ganglios linfáticos. Si el resultado de la biopsia de ganglio centinela es positivo, se recomienda una disección de ganglios linfáticos, extirpando todos los ganglios linfáticos cercanos al cáncer (aunque esto no mejora la supervivencia del paciente).

El índice de supervivencia cuando el melanoma es detectado en esta etapa es del 90% [4].

- Etapa o estadio II:

En esta etapa, como se observa en la Figura 4, el melanoma se ha diseminado a la parte inferior de la dermis, pero no al tejido situado debajo de la piel ni a los ganglios linfáticos vecinos [2].

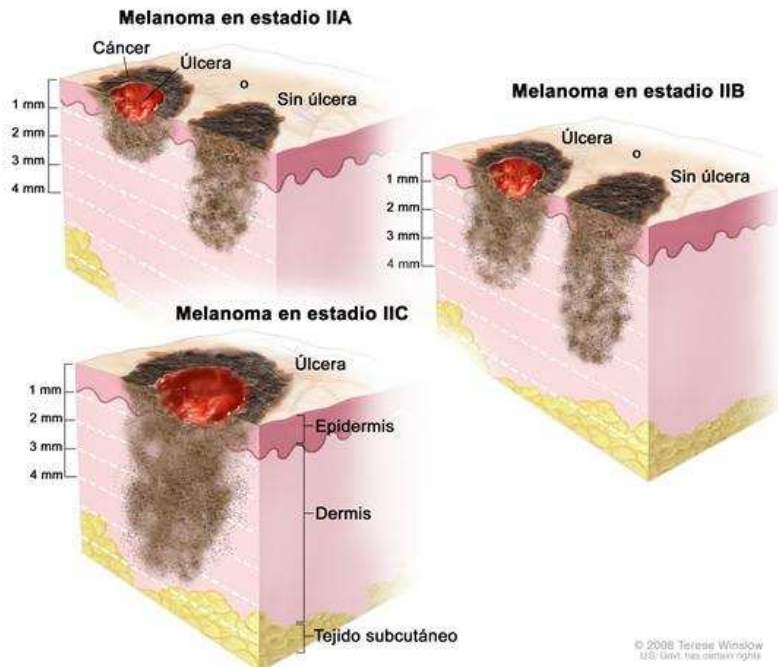


Figura 4 – Estructura interna de la piel cuando existe un melanoma en etapa o estadio II [6].

El tumor tiene un grosor de 2 mm a 4 mm y el tratamiento convencional consiste en emplear la cirugía para realizar una escisión amplia tomando un margen de piel sana, que depende del grosor del melanoma.

Debido a que el melanoma podría haberse propagado a los ganglios linfáticos cercanos al melanoma, se recomienda realizar también una biopsia del ganglio centinela. En caso de que el ganglio centinela contenga cáncer, se procede a realizar una disección de ganglios linfáticos, extirpando quirúrgicamente todos los ganglios linfáticos en esa área.

El índice de supervivencia cuando el melanoma es detectado en esta etapa es del 80% [4].

- Etapa o estadio III:

En esta etapa, el tumor es de más de 4 mm de espesor y presenta cualquiera de las siguientes situaciones, representadas en la Figura 5 [5] [6]:

- El cáncer se ha diseminado hasta uno o más ganglios linfáticos.
- Los ganglios linfáticos pueden estar unidos (enredados).
- El cáncer puede estar en un vaso linfático entre el tumor primario y los ganglios linfáticos cercanos.
- Se pueden encontrar tumores muy pequeños sobre la piel o debajo de ésta a menos de dos centímetros del lugar donde empezó el cáncer.

Es decir, que el tumor se ha diseminado al tejido corporal situado debajo de la piel, dispersándose a los ganglios linfáticos del área, o bien apareciendo tumores adicionales que han crecido entre el tumor original y los ganglios linfáticos del área, llamados tumores satélites.

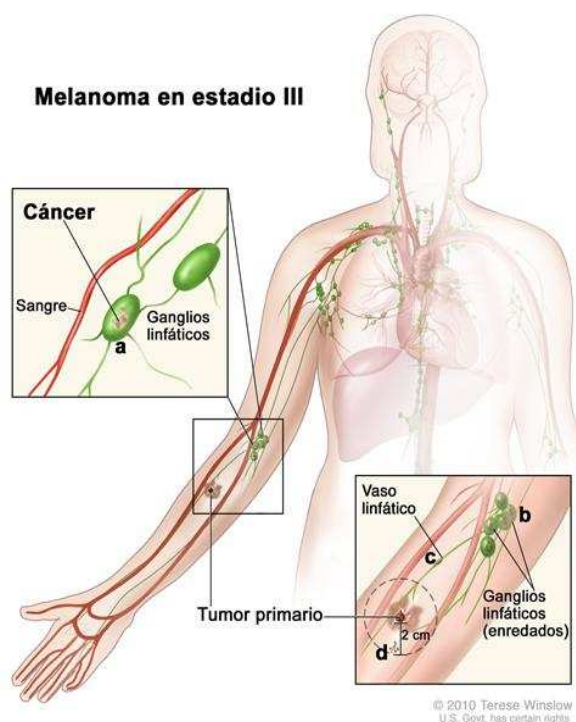


Figura 5 – Estructura interna de la piel cuando existe un melanoma en etapa o estadio III [6].

El tratamiento quirúrgico, requiere la disección de los ganglios linfáticos, junto con la escisión amplia del tumor primario como en la etapa II. Además, se recomienda una terapia adyuvante (tratamiento adicional) con Interferón después de la cirugía. O mediante un estudio clínico, por medicamentos y vacunas, reducir las probabilidades de que el melanoma regrese [2].

Si hay varios melanomas, todos éstos deben ser extirpados. Si no se puede, las opciones para el tratamiento son: administrar inyecciones de BCG o Interleucina – 2 directamente en el melanoma o aplicar la inmunoterapia tópica con Imiquimod. Para los melanomas en un brazo o en una pierna, otra posible opción es la perfusión aislada de la extremidad, es decir, infundir la extremidad con una solución calentada del medicamento de quimioterapia. En algunos casos, se puede administrar radioterapia como adyuvante a la cirugía en el área donde los ganglios linfáticos son extirpados, especialmente si se encuentran muchos ganglios que contienen cáncer. Otros posibles tratamientos incluyen quimioterapia, inmunoterapia con citocinas, o ambas combinadas (bioquimioterapia).

Los tratamientos más nuevos que se están probando en los estudios clínicos podrían beneficiar a algunos pacientes. Muchos pacientes con melanoma en etapa III puede que no sean curados con los tratamientos actuales, por lo que deben considerar la participación en estudios clínicos.

El índice de supervivencia cuando el melanoma es detectado en esta etapa es del 50% [4].

- Etapa o estadio IV:

En esta etapa, como se muestra en la Figura 6, el tumor se ha diseminado a otras áreas del cuerpo, ya sea a otros órganos o a ganglios linfáticos alejados del tumor original, por lo que los melanomas en esta etapa son muy difíciles de curar [2].

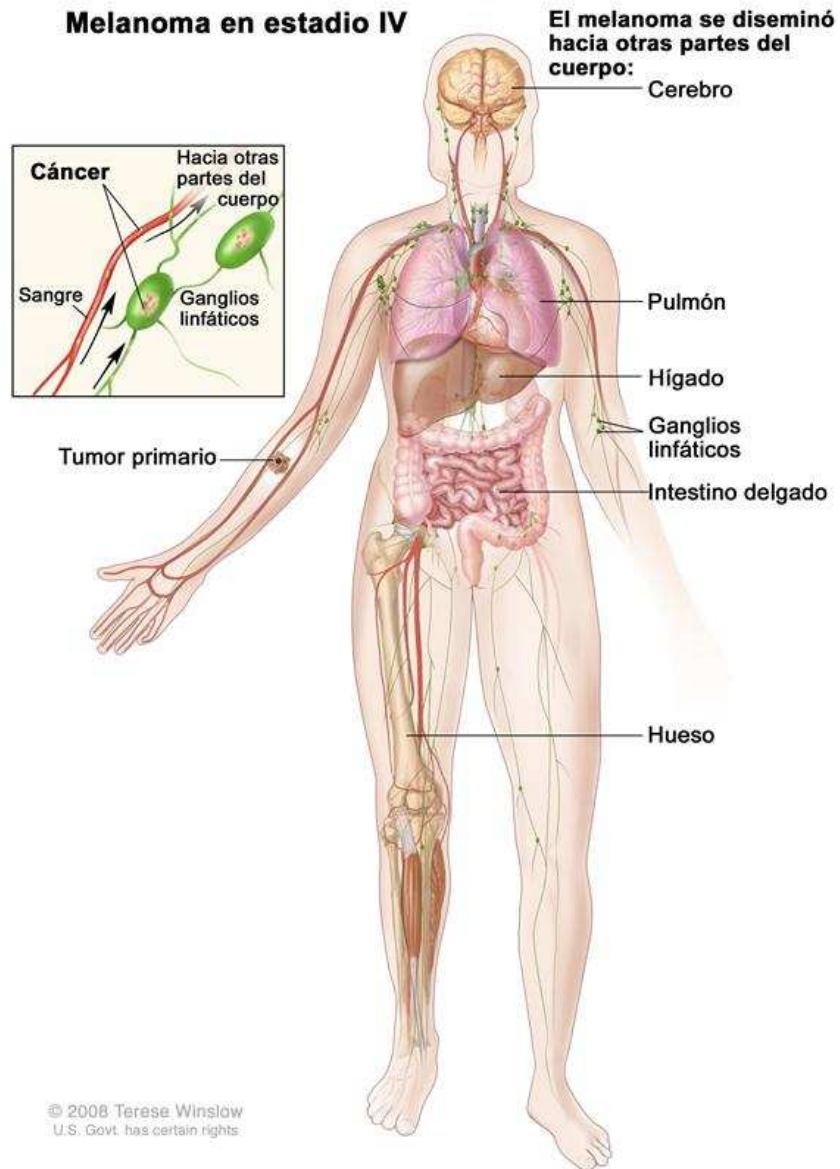


Figura 6 – Estructura interna de la piel cuando existe un melanoma en etapa o estadio IV [6].

Los tumores en la piel o las metástasis en los ganglios linfáticos que producen síntomas, a menudo se pueden extirpar mediante cirugía. Las metástasis en los órganos internos a veces se pueden extirpar, dependiendo de cuántas hay presentes, donde están localizadas y la probabilidad de que causen síntomas. Las metástasis que causan síntomas, pero que no se pueden extirpar quirúrgicamente, se pueden tratar con radiación, inmunoterapia, o quimioterapia. Aun cuando la quimioterapia puede reducir el tamaño de estos cánceres, el efecto a menudo solo es temporal, con un promedio de tiempo de 3 a 6 meses antes de que el cáncer comience a crecer nuevamente. No obstante, en pocas ocasiones, estos medicamentos son eficaces por períodos de tiempo más prolongados. La inmunoterapia puede ayudar a un pequeño número de pacientes con melanoma en esta etapa a vivir más tiempo. Las dosis más

altas de estos medicamentos parecen ser más eficaces, pero también tienen efectos secundarios más graves. O también se puede aplicar un tratamiento de bioquimioterapia, una combinación de quimioterapia e inmunoterapia.

Debido a que el melanoma en etapa IV es muy difícil de tratar con las terapias actuales, es posible que los pacientes quieran considerar la participación en un estudio clínico. Los estudios clínicos con nuevos medicamentos de quimioterapia, medicamentos dirigidos, nuevos métodos de inmunoterapia, tal como vacunas, así como las combinaciones de diferentes tipos de tratamientos, pueden beneficiar a algunos pacientes.

El índice de supervivencia cuando el melanoma es detectado en esta etapa es entre el 20% y 30% [4].

- Recurrente:

Enfermedad recurrente significa que el cáncer ha vuelto a aparecer (recurrido) después de haber sido tratado [5]. Puede reaparecer en el sitio original o en otra parte del cuerpo. El tratamiento del melanoma que reaparece después del tratamiento inicial depende de la etapa del melanoma original, del tipo de tratamiento inicial y del lugar de la recurrencia [2].

El melanoma puede regresar en la piel cercana al lugar donde se originó el tumor. En general, estas recurrencias locales en la piel se tratan con una cirugía similar a la recomendada contra el melanoma primario y puede incluir una biopsia del ganglio linfático centinela. Dependiendo del grosor y la localización del tumor, se pueden considerar otros tratamientos, como: perfusión de quimioterapia en extremidad aislada, quimioterapia sistémica, inmunoterapia, radioterapia, o inyección al tumor con la vacuna BGC o Interferón.

Si los ganglios linfáticos adyacentes no son extirpados durante el tratamiento inicial, el melanoma puede regresar a estos ganglios linfáticos cercanos y aparecer como una hinchazón o una masa tumoral. La recurrencia en los ganglios linfáticos se trata mediante una disección de ganglios linfáticos, y puede incluir terapia adyuvante, como Interferón o radioterapia.

El cáncer también puede regresar a áreas distantes. Casi cualquier órgano puede ser afectado. Con más frecuencia, el melanoma regresará en los pulmones, los huesos, el hígado o el cerebro. Por lo general, el tratamiento de estas recurrencias es el mismo que se usa en el melanoma que se encuentra en etapa IV. Además los melanomas que recurren en un brazo o una pierna pueden ser tratados con perfusión de quimioterapia en extremidad aislada.

El melanoma que regresa al cerebro puede ser difícil de tratar. Algunas veces localizaciones individuales de recurrencia se pueden extirpar mediante cirugía. La mayoría de los medicamentos de quimioterapia no pueden alcanzar el cerebro, aunque la temozolomida y la radioterapia pueden ser útiles.

Al igual que en otras etapas de melanoma, es posible que los pacientes con melanoma recurrente quieran considerar la participación en un estudio clínico.

1.3. OBJETIVOS Y FASES DE DESARROLLO.

El objetivo principal de este proyecto consiste en discriminar automáticamente mediante visión por ordenador imágenes de melanomas de no melanomas.

En este proyecto se parte de la imagen de una lesión previamente analizada, es decir, una lesión a la cual se le ha realizado una biopsia para saber realmente si era o no un melanoma y con la que se quiere obtener un programa automático que demuestre que sin realizar una biopsia es posible realizar una clasificación de la lesión simplemente con una imagen de la misma.

En un principio, este programa sería de uso clínico, para un hospital o clínica. Lo podrían emplear tanto expertos como principiantes. La principal ventaja es que evitaría tener que realizar una biopsia al paciente y además es una forma rápida y exacta de analizar lesiones sospechosas de ser melanomas.

Para alcanzar este objetivo, se deben seguir los siguientes pasos:

- **Obtener la imagen de la lesión:** la imagen de la lesión ha de estar digitalizada para poder analizarla gracias a un programa ejecutado por un ordenador y que está compuesto por las siguientes etapas.
- **Segmentación de la lesión:** es la primera etapa del análisis de la imagen y se debe realizar una vez capturada y digitalizada la imagen. La segmentación consiste en encontrar la imagen binaria que diferencia entre las regiones de piel y lesión.
- **Extracción de características:** es la siguiente etapa a realizar en el análisis de la lesión. La extracción de características consiste en obtener todas las características posibles que puedan ayudar en la clasificación de una lesión entre melanomas y no melanomas.
- **Clasificación de la lesión:** es la última etapa del proceso. La clasificación se desarrolla a partir de las características obtenidas en la etapa anterior y permite clasificar la lesión que se está examinando, como ya se ha mencionado anteriormente entre melanoma y no melanoma.
- **Evaluación de los resultados:** es el último paso del proyecto y consiste en demostrar que el sistema automático es rápido y preciso, porque clasifica correctamente la mayoría de las lesiones.

En este proyecto las lesiones son proporcionadas por el Departamento de Dermatología de la Universidad de Missouri – Rolla. Además, este proyecto se ha desarrollado con Matlab.

1.4. ESTRUCTURA DE LA MEMORIA.

La memoria del proyecto está compuesta por 10 capítulos. A continuación, se describen brevemente los siguientes capítulos:

- En el capítulo 2, se explican los métodos existentes para poder capturar y digitalizar las imágenes de las lesiones, los métodos para poder realizar la segmentación de las mismas y los métodos para poder clasificarlas y evaluarlas.
- En el capítulo 3, se propone el método de captura y acondicionamiento para las imágenes de las lesiones.
- En el capítulo 4, se expone el método de segmentación, explicando el preprocesado y postprocesado del mismo. Empleando las imágenes capturas y digitalizadas en el capítulo 3.
- En el capítulo 5, se proponen los parámetros que se deben extraer de las lesiones, para poder diferenciar entre melanomas y no melanomas, teniendo en cuenta la regla ABCD de la dermatoscopia y empleando las imágenes acondicionadas y segmentadas, obtenidas de los capítulos 3 y 4, respectivamente.
- En el capítulo 6, se presentan los métodos para clasificar y evaluar las lesiones según las características extraídas en el capítulo 5.
- En el capítulo 7, se exponen los resultados obtenidos tras la clasificación de las lesiones realizada en el capítulo 6, empleando las características extraídas anteriormente en el capítulo 5.
- En el capítulo 8, se presentan algunas de las posibles mejoras que podría tener el programa en el futuro.
- En el capítulo 9, se plantea el coste del proyecto en caso de llevarse a cabo.
- En el último capítulo, el capítulo 10, se presentan las referencias empleadas y mencionadas a lo largo de todo el proyecto.

Capítulo 2.

ESTADO DEL ARTE.

Desde la aparición del melanoma a principios del siglo XX (1930, en Estados Unidos) [11], la única manera de comprobar si una persona tenía un melanoma o no, era mediante la realización de una biopsia.

La biopsia, es una técnica invasiva, que consiste en realizar una extracción parcial de tejido de la zona sospechosa y analizarla bajo el microscopio para examinar el tipo de células que presenta [4]. El microscopio, mostrado en la Figura 7 y Figura 8, es un instrumento que permite observar objetos que son demasiado pequeños para ser vistos a simple vista. En él se observan laminillas de vidrio que contienen el espécimen, observándose en un solo plano [12].

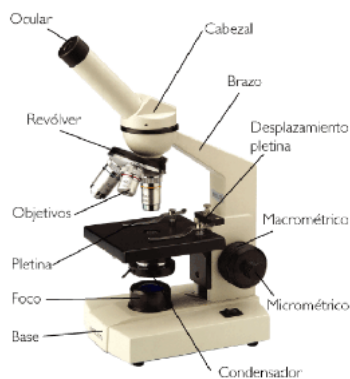


Figura 7 – Microscopio: Real [13].

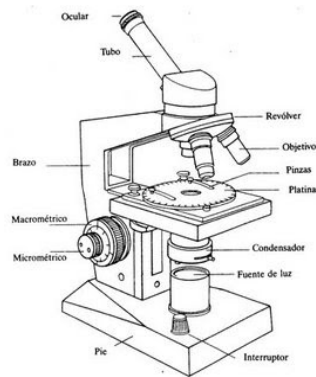


Figura 8 – Microscopio: Esquema [13].

Tras numerosos estudios en donde se desarrollaron diversos métodos para analizar lesiones y clasificarlas a simple vista, en 1971, se reconoce la importancia del estudio de la superficie de la piel microscópica en la evaluación preoperativa de una lesión melanocítica: nace la dermatoscopia. Numerosos grupos austríacos y alemanes continuaron estos estudios durante los años 80 [14].

Al principio se realizaban los análisis de la superficie de la piel microscópica mediante un estereomicroscopio [14]. Esta herramienta, consta de dos microscopios completos, cada uno con su objetivo y ocular en los que, al no coincidir sus ejes ópticos, las imágenes formadas en los oculares son distintas (como ocurre en la visión ocular), por lo que se pueden observar imágenes en tres dimensiones. Además, el estereomicroscopio, expuesto en la Figura 9 y Figura 10, funciona con pocos aumentos (20, 40, 60) y sirve para observar objetos de mayor tamaño que en el microscopio [12].

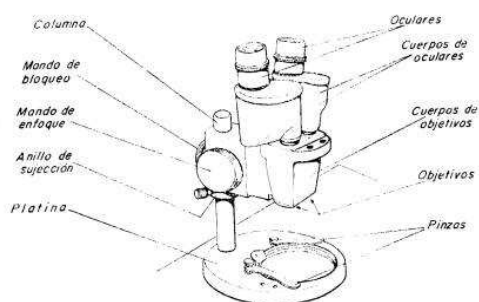


Figura 9 – Estereomicroscopio: Esquema [15].

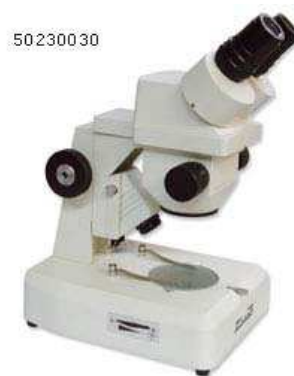


Figura 10 – Estereomicroscopio: Real [16].

Pero este instrumento, obstaculizaba la aplicación clínica por ser muy pesado, caro e incómodo. Por lo que la técnica tenía un procedimiento difícil y consumía mucho tiempo [14].

Actualmente, se emplea para analizar lesiones dermatológicas in situ, el llamado dermatoscopio manual, que es más sencillo, menos costoso y más cómodo. Este instrumento, representado en la Figura 11, contiene un sistema de lentes que aumentan de 10 a 400 veces la imagen, y un sistema de iluminación, que permite que las imágenes sean comparables y reproducibles, haciendo que las capas superficiales de la piel sean transparentes. Esto último permite visualizar estructuras profundas, que normalmente son invisibles al ojo humano [17].



Figura 11 – Dermatoscopio manual: Diferentes tipos [18] [19].

Gracias al avance de la tecnología, se opta por desarrollar un sistema basado en visión por ordenador automático, fiable y lo más exacto posible a la hora de clasificar las lesiones. Se pretende que el paciente no pase por una biopsia (sistema invasivo) que suele ser doloroso para el propio paciente.

Como se ha señalado previamente, este sistema automático consiste fundamentalmente de las siguientes etapas:

- Captura y acondicionamiento de la imagen de cada lesión.
- Técnicas de segmentación aplicadas a la imagen de la lesión.
- Extracción de las características de la imagen de la lesión.
- Clasificación de la lesión.

A continuación, se realiza una breve descripción de cada una de estas etapas.

2.1. CAPTURA Y ACONDICIONAMIENTO DE LA IMAGEN.

Para poder capturar la imagen de cada lesión, se pueden emplear varios métodos, de los que se destaca:

- El método ELM (epiluminescence microscopy) [11]:

El ELM es una técnica no invasiva que emplea la inmersión por aceite, haciendo que las estructuras de la subsuperficie de la piel, sean accesibles, por lo que proporciona un diagnóstico de criterios basándose en características que no se pueden apreciar a simple vista [20].

Este método, se puede aplicar empleando alguno de los instrumentos que se mencionan a continuación, como son los videomicroscopios o los dermatoscopios digitales.

- El videomicroscopio, mostrado en la Figura 12 y Figura 13, es la unión de dos tecnologías: microscopio y vídeo, por lo tanto permite grabar objetos muy pequeños y mostrarlos por pantalla [21]. Es un aparato fácil de manejar por ser pequeño y su precio es razonable.



Figura 12 – Videomicroscopio: Real [22].

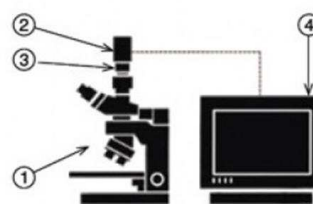


Figura 13 – Videomicroscopio: Esquema [23].

Por tanto, para obtener un videomicroscopio se debe colocar en el ocular de un microscopio una cámara de video por medio de un adaptador y esta cámara a un monitor o PC, como se indica en la Figura 14. Así es posible observar lo que sucede en el microscopio, además de poder grabarlo en video o fotograma [12].

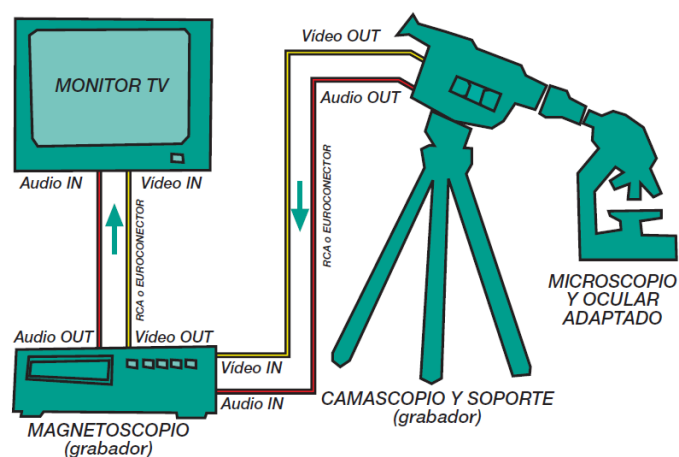


Figura 14 – Videomicroscopio: Esquema completo [21].

El microscopio óptico es un instrumento que permite formar, y por tanto observar imágenes virtuales, invertidas y aumentadas de objetos pequeños y además está formado básicamente por dos sistemas de lentes: el ocular y el objetivo. Mientras que el camascopio (cámara de vídeo) es un aparato óptico – electrónico que permite registrar electrónicamente o grabar sobre un soporte magnético o cinta imágenes en movimiento [21].

Existen dos modos de observación con el videomicroscopio: un modo es la observación polarizada y el otro modo es la observación aplicando un poco de aceite y añadiendo una lámina de cristal. Ambos modos, proporcionan una imagen digital, cuyos valores, en la mayoría de los casos, no difiere estadísticamente. Por lo que aplicando el modo polarizado (es decir, sin aceite) se puede convertir en un procedimiento de examinación rápido y con el mismo resultado que aplicando el método por inmersión en aceite [20].

- Otro instrumento frecuentemente usado en la adquisición de imágenes dermatoscópicas es el equipo Mole Max, mostrado en la Figura 15, que es un aparato digital que permite diagnosticar lesiones. Incluye un dermatoscopio digital y permite realizar un seguimiento de la lesión, puesto que puede almacenar la imagen de la lesión gracias a la tecnología computarizada que lleva asociada. Es un equipo amplio, costoso y difícil de utilizar [24].



Figura 15 – Equipo Mole Max II [25].

2.2. TÉCNICAS DE SEGMENTACIÓN DE LA IMAGEN.

Una vez capturada la imagen de la lesión y digitalizarla, se debe segmentar dicha imagen para poder más tarde extraer de ella las características. Tras la segmentación de la imagen de la lesión, es decir, diferenciadas la lesión y la piel sana, se obtiene la imagen o máscara binaria [26] [27].

Para optimizar el resultado de la segmentación, es conveniente realizar un preprocesado y un postprocesado de la imagen que haga la segmentación más precisa.

- En el preprocesado, se realiza una transformación del espacio de color, para mejorar el contraste y eliminar artefactos. Inicialmente se desarrollaron unas transformaciones que pretendían mejorar el contraste para cualquier tipo de lesión. En donde se emplea un escalar (en este caso un solo canal), es decir, que las imágenes en color RGB (red-green-blue) se convierten a un escalar empleando unos de los siguientes métodos:
 - Conservando sólo el canal azul, debido a que a menudo las lesiones son más prominentes en este canal [28].
 - Aplicando la transformación de la luminancia [3] [26] [28], expresada por la Ecuación 1:

$$Luminance = 0.299 \times R + 0.587 \times G + 0.114 \times B$$

Ecuación 1 – Luminancia estándar.

En donde:

La fórmula de la luminancia estándar, llamada *Luminance*, transforma las imágenes a color a imágenes en blanco y negro y los parámetros *R*, *G* y *B* representan los valores de intensidad de píxeles del rojo, verde y azul, respectivamente.

- Aplicando la transformación de Karhunen-Loève (KL) y conservando el canal con mayor varianza [28].

Una posibilidad mejor es encontrar la proyección de forma adaptativa. Una opción es encontrar la proyección del espacio original RGB que maximice la bimodalidad de los datos proyectados [27] [29]. La mejora de contraste facilita la segmentación de la imagen de forma que puede realizarse con un simple algoritmo de agrupación o clustering [28].

Para eliminar los artefactos se pueden emplear los siguientes filtros: gaussiana (Gaussian Filter GF), mediana (Median Filter MF) o difusión anisotrópico (Anisotropic Diffusion Filter ADF). O emplear métodos especializados para cada tipo de artefacto.

- El objetivo de la segmentación es obtener la imagen o máscara binaria. Para realizar la segmentación, existen varios métodos entre los que se encuentran: thresholding, clustering (por agrupación), edge-based, region-based, model-based y active contours [28].

A continuación se explican algunos ejemplos de cada método de segmentación:

➤ Thresholding [11] [26] [30]:

Con frecuencia, el algoritmo más empleado es el algoritmo de Otsu. Este algoritmo determina un valor umbral que permite diferenciar entre piel y lesión, basándose en la varianza. La varianza es una medida de dispersión calculada en los niveles de gris. El método de Otsu, establece un umbral cualquiera dividiendo los datos en dos grupos y calcula la varianza intragrupal aplicando la Ecuación 2:

$$\sigma_w^2(t) = \omega_1(t) \cdot \sigma_1^2(t) + \omega_2(t) \cdot \sigma_2^2(t)$$

Ecuación 2 – Método de Otsu: cálculo de la varianza intragrupal.

En donde:

La varianza intragrupal está representada por $\sigma_w^2(t)$, las proporciones de cada grupo son $\omega_1(t)$ y $\omega_2(t)$, respectivamente y las varianzas de cada grupo son $\sigma_1^2(t)$ y $\sigma_2^2(t)$, respectivamente.

Una vez calculada la varianza intragrupal $\sigma_w^2(t)$, se debe calcular la varianza entre grupos representada por $\sigma_b^2(t)$. Para ello, se debe aplicar la Ecuación 3 que relaciona ambas varianzas con la varianza total de todo el histograma σ^2 :

$$\sigma_b^2(t) = \sigma^2 - \sigma_w^2(t)$$

Ecuación 3 – Método de Otsu: cálculo de la varianza entre grupos.

El método de Otsu, realiza el mismo cálculo para todos los umbrales posibles y determina como valor umbral que permite diferenciar entre piel y lesión, aquel umbral que maximice la varianza entre grupos [10] [26] [31].

➤ Agrupación mediante k-medias.

El método de agrupación k – means, consiste en que dado un conjunto de puntos, se debe colocar aleatoriamente el número de medias deseadas, que corresponde al número de grupos en los que se quiere dividir el conjunto. En nuestro caso, dos medias, porque se quieren tener dos conjuntos finales (lesión y piel). Una vez colocadas las medias, se calculan qué puntos pertenecen por aproximación a cada media, es decir, que cada punto corresponde a una media por estar más cercana a ella. Tras la agrupación de los puntos, por la media inicial, se calcula la nueva media y se agrupan los puntos en función de la nueva media. Repitiendo el proceso hasta que las medias sean constantes, el resultado será el conjunto de puntos agrupado en dos grupos, en nuestro caso, en un grupo de piel y en otro de lesión [10] [32].

➤ Edge-based:

El método mediante snakes o serpientes, consiste en emplear una silueta que contenga a la lesión. Para encontrar el borde de la lesión, se contrae dicha silueta hacia la lesión, es decir, hacia dentro, y colocándose en el borde de la misma [24] [27] [33].

➤ Region-based (basado en la región):

El método mediante watersheds o cascadas, consiste en emplear los valores de los píxeles de la imagen y representarlos. Dado que el valor de los píxeles más oscuros tendrán un valor muy bajo y los píxeles más claros tendrán un valor muy alto (por ser el blanco en RGB (255,255,255) y el negro en RGB (0,0,0)), se obtendrá una representación del valor de los píxeles muy variable. Este método consiste en aumentar los valores más bajos hasta un cierto valor, este valor será el que represente gráficamente un llenado del hueco con menor valor en píxeles. Y así sucesivamente hasta completar la imagen [27] [34].

➤ Model-based (basado en un modelo):

El método que realiza la segmentación de la lesión aplicando un modelo de segmentación a tres niveles: piel saludable, parte clara del melanoma y parte oscura del melanoma [35].

- En el postprocesado, se fusionan regiones y se eliminan artificios [28].

Como resultado del proceso de segmentación se tiene una imagen etiquetada o un mapa de borde binario. Para obtener una segmentación más precisa, suele ser necesario postprocesar la imagen segmentada realizando los siguientes pasos:

- Fusionar o combinar regiones (region merging).
- Eliminar artificios (island removal).

2.3. EXTRACCIÓN DE CARACTERÍSTICAS.

Para poder clasificar las lesiones, se extraen de las imágenes de dichas lesiones diversas características o indicadores que revelan si dicha lesión es maligna.

Entre otras, numerosos estudios han aplicado las siguientes reglas, métodos o algoritmos:

- La regla ABCD básica, aplicada a simple vista [9] [14] [35].
- La regla ABCD de la dermatoscopia [14] [36], que es un algoritmo semicuantitativo basado en la evaluación de la forma y la distribución del pigmento, es decir, que evalúa la asimetría, la regularidad del borde, los colores y la estructuras diferenciales dentro de la lesión.
- El método de la puntuación de Menzies [35].
- Análisis de patrones [35] [37].

- El New 7 Point Checklist [35], que es un algoritmo semicuantitativo que explica las pautas de análisis [38].
- O simplemente características que intentan clasificar la lesiones:
 - La epiluminiscencia microscópica descrita, permite la visualización in vivo de las estructuras pigmentadas anatómicas de la epidermis, dermoepideral junction y papillary dermis superficial [38] [39].
 - Otros conjuntos de características indicativas de la malignidad en imágenes ELM se mencionan, pero la mayoría no están suficientemente descritas [39].

2.3.1. Regla ABCD aplicada a simple vista [9] [14] [35].

En 1985, Friedman *et al* [9], propuso aplicar la regla ABCD, una regla muy útil para diagnósticos clínicos de lesiones melanocíticas, es decir, de lesiones que derivan del melanocito. Esta regla se basa en que una lesión pigmentada tiene más probabilidades de convertirse en melanoma maligno si a tiende a los siguientes criterios: Asimetría, Borde irregular, Colores diferentes y Diámetro superior a 6 mm. El inconveniente de este método pionero es que es aplicada de manera subjetiva y a simple vista por el dermatólogo [14].

Existe una extensión menos común donde también se considera el criterio de la Elevación, la llamada regla ABCDE [35].

2.3.2. Regla ABCD de la Dermatoscopia [14] [36].

En 1994, Stolz *et al* [14] y Nachbar *et al* [36] desarrollaron la nueva regla ABCD de la dermatoscopia, basándose en el análisis semicuantitativo de cuatro características extraídas de la superficie microscópica de la piel. Esta regla es parecida a la anterior, con la diferencia de que el criterio del Diámetro es reemplazado por las Diferencias estructurales y que para extraer las características se emplea un estereomicroscopio (como se hacía antes) o un dermatoscopio (como se hace en la actualidad).

La regla ABCD de la dermatoscopia, se emplea para diferenciar las lesiones malignas de las benignas. Es una técnica fácil de aprender y rápida de calcular [9], por lo que expertos y principiantes pueden aplicarla, empleando la Ecuación 4:

$$FDS = A \times 1.3 + B \times 0.1 + C \times 0.5 + D \times 0.5$$

Ecuación 4 – Regla ABCD de la dermatoscopia.

En donde:

- El valor de la Asimetría, *A*, depende del número de ejes en los que la lesión es asimétrica. Es decir, trazando dos ejes perpendiculares en la lesión, si la lesión es simétrica, el valor de *A* será 0, si la lesión es asimétrica respecto a un eje, el valor de *A* será 1 y si la lesión es asimétrica con respecto a los dos ejes, el valor de *A* será 2.
- El valor del Borde, *B*, depende del número de segmentos en los que se observa un corte abrupto en el borde de la lesión. Un corte abrupto significa que existe una gran diferencia entre la lesión y la piel en la zona del borde, o dicho de otra manera, que el

color de la lesión no se va degradando a medida que se acerca a la piel. Para asignar el valor de B , se debe dividir la lesión en 8 segmentos, siendo cada uno de esos segmentos los que asignen el valor de B y cuando exista un borde abrupto el valor de B aumenta. Por tanto el valor de B varía entre 0 (no hay segmentos con borde abrupto, es decir, que la lesión tiene un borde que se va degradando) y 8 (la lesión tiene un borde abrupto y se diferencia notablemente con respecto a la piel).

- El valor del Color, C , depende del número de colores que intervienen en la lesión. Cuantos más colores, mayor valor de C . Los posibles colores son: blanco, rojo, marrón claro, marrón oscuro, azul grisáceo y negro, por lo que el valor de C varía entre 1 y 6.
- El valor de las Diferencias estructurales, D , depende del número de estructuras que posee la lesión. Desde áreas homogéneas, cadenas, venas, puntos y/o glóbulos, por lo que el valor de D varía entre 1 y 5.
- Final Dermatoscopy Score, FDS , corresponde a la puntuación final de la combinación de las cuatro características que abarca la regla ABCD de la dermatoscopia. Dependiendo del valor de FDS se puede tener: un melanoma maligno, una lesión sospechosa de ser maligna o una lesión benigna.
 - Si FDS excede de 5.45, la lesión tiene una gran probabilidad de ser un cáncer de piel de tipo melanoma maligno. Por tanto, requiere una extirpación lo antes posible y su tratamiento depende del estadio en el que se encuentre el melanoma.
 - Si FDS se encuentra entre 4.75 y 5.45, la lesión tiene probabilidad moderada de ser melanoma maligno. En este caso, es necesaria una vigilancia constante de la evolución de la lesión para verificar si puede convertirse o no en un melanoma maligno.
 - Si FDS es menor de 4.75, es probable que la lesión sea un cáncer de piel de tipo benigno. Por lo que no es necesaria una extirpación, pero quizás requiera un tratamiento diferente, dependiendo del cáncer que sea.

2.3.3. Método “New 7 – Point Checklist” [38].

Más tarde, en 1998, Argenziano *et al* [38] presentó el método New 7 Point Checklist, basado en una simplificación de patrones de análisis ELM presentados en 1989. Este método es un listado de siete características extraídas de diversas lesiones y que ayudan en la clasificación de las mismas. Las características que forman dicho listado se citan a continuación:

- Retículo pigmentado atípico (Atypical pigment network).
- Zonas gris - azul (Gray - blue areas).
- Patrón vascular atípico (Atypical vascular pattern).
- Transmisión radial (rayas) (Radial streaming (streaks)).
- Pigmentación difusa irregular (manchas) (Irregular diffuse pigmentation (blotches)).
- Puntos y glóbulos irregulares (Irregular dots and globules).
- Patrón de regresión (Regression pattern).

En donde el estudio realizado, calcula mediante Odds Ratio (OR), correspondiente al cociente entre el número de veces que ocurre y el número de veces que no ocurre, las características anteriores y puntúa con un valor de 2 a las características que obtengan un resultado de Odds Ratio superior a 5 y puntúa con un valor de 1 a las características que obtengan un resultado de Odds Ratio inferior a 5.

2.3.4. Otras características propuestas.

Varios estudios, demuestran claramente que la determinación semicuantitativa de solo cuatro criterios bien definidos (asimetría, borde, color y diferencia estructural) con el dermatoscopio, es suficiente para discriminar con buena precisión entre lesiones melanocíticas benignas y malignas [14].

Por tanto, a continuación se mencionan algunos de los indicadores o parámetros que se obtienen para expresar de forma automática las principales características ABCD y que están basados en visión por ordenador:

- Safi *et al* [35] propuso un conjunto de características obtenidas a partir de la regla ABCDE (Asimetría, Borde, Color, Diámetro, Elevación) mediante técnicas de visión por ordenador. Este conjunto de características fue obtenido a través de las propiedades geométricas, de color, de textura y de figura de la imagen de la lesión segmentada, obteniéndose un total de 8 características. Entre ellas se incluyen la excentricidad, el área o el centroide.
- En otro artículo, Seidenari *et al* [20] comparó los dos modos de observación de un videomicroscopio, es decir, el modo polarizado y el modo por inmersión en aceite. Para realizar la comparación se extrajeron numerosas características, las cuales se compararon estadísticamente. En este estudio, se extrajeron hasta 34 características formadas por parámetros geométricos, parámetros de las componentes RGB y de los momentos de inercia polares en los ejes x e y. Entre ellas se incluyen la media o los cuartiles y percentiles del histograma en banda roja, verde o azul.
- Green *et al* [40] desarrolló un software encargado del análisis de lesiones, el cual extraía características en relación al tamaño, el color (diversidad), la forma y el límite de cada lesión (regularidad y diferencias del límite con los alrededores de la piel). De un total de 22 características extraídas de la lesión, se emplearon finalmente 17 para clasificar las lesiones, es decir, para discernir entre melanomas malignos y lesiones pigmentadas benignas. Entre ellas se incluyen el tamaño, el color o el borde.
- En otro artículo, Cascinelli *et al* [41] propuso un conjunto de ocho indicadores binarios (on/off) extraídos de las lesiones. Los cuales fueron: la forma, la medida de la asimetría de la lesión, la información clínica (como el tamaño, el color o la presencia de úlcera), la distribución del color en la lesión, el tamaño de la lesión medida en milímetros, la oscuridad (es decir, el porcentaje de negro mezclado en el color), la saturación (es decir, el porcentaje de blanco mezclado en la lesión), el borde (es decir, la transición entre la lesión y la piel sana) y la textura dada por la superficie de la lesión.

A continuación, de forma más específica se indican varias de las medidas propuestas para representar la característica de asimetría:

- Cheng *et al* [42] obtuvo 17 características, entre las cuales se menciona: el área, el grado de redondez (thinness/roundness), la media RGB, la desviación típica RGB, el coeficiente de simetría (skew/skewness and kurtosis) RGB, la energía RGB y la entropía RGB.
- En otro artículo, Stoecker *et al* [3] realizó el análisis de lesiones empleando un algoritmo para encontrar un eje mayor de asimetría y calcular el grado de asimetría en el límite del tumor. Propuso tres algoritmos para determinar la asimetría:
 - El método de los ocho puntos vecinos (eight-point neighborhood), empleando la aproximación a una curva cerrada o a un polígono.
 - El método de simetría rotacional, por aproximación a una curva cerrada y a un círculo. Basándose en que cualquier figura plano o curva cerrada al menos posee 2 ejes principales de 90 grados entre sí y que se cruzan en el centro del área.
 - El método de los momentos de inercia, sabiendo que será máximo en un eje principal y mínimo en el otro eje principal. Y justificando que lesiones con una aproximación oval, al realizarles ajustes en las longitudes radiales, se asemejan a un óvalo o a una elipse.

De forma similar, a continuación se describen varios parámetros que permiten representar la característica del borde de la regla ABCD de la dermatoscopia. Algunos descriptores del borde son tales como, la irregularidad y la presencia de un borde brusco o abrupto que son considerados como indicios o síntomas de malignidad de la lesión [26].

- Celebi *et al* [28] explicó que para evaluar el borde de una lesión antes se debe mejorar el contraste de la imagen, empleando un método basado en el algoritmo lineal que transforma la imagen original RGB en un espacio en donde la lesión y el fondo de la piel son separados al máximo. Para luego aplicar un simple algoritmo de agrupación (clustering) y en el postprocesado además de unir y eliminar islas, se debe suavizar y expandir el borde.
- En otro artículo, Grana *et al* [26] presentó cinco parámetros que permiten calcular la característica correspondiente al borde de la lesión. Propuso una nueva aproximación matemática para evaluar la frontera de la lesión, considerada por valores de luminancia a lo largo de la dirección normal de cada punto del contorno. Requería dos condiciones para asegurar una correcta evaluación, la primera que el segmento estuviera completamente incluido en los límites de la imagen y la segunda que dicho segmento sólo debía cruzar el borde una vez. Este método empleaba líneas de tendencia que pasaban por los puntos obtenidos de los píxeles dentro y fuera de la lesión. Permitiendo una evaluación precisa del contorno, de la profundidad de las muescas y la irregularidad de la figura.
- Day *et al* [39] presentó un algoritmo interesante que medía automáticamente el número de segmentos con corte brusco o afilado. Definiendo el corte brusco del pigmento en un segmento, como un cambio grande en la luminosidad desde la lesión hasta la piel en ese mismo segmento. Para realizar este análisis, dicho algoritmo

dividía primeramente la lesión en 8 segmentos, obtenía el borde de la lesión y medía el gradiente de luminosidad en el borde de la lesión. Los gradientes eran asignados como "borrosamente" o "corte brusco" en base a un umbral obtenido por comparación de las opiniones de dermatólogos con el algoritmo. Para medir el cambio de luminosidad en cada punto del límite, llamado p , se recogían 60 valores de luminosidad (30 píxeles dentro del borde y 30 píxeles fuera del borde, estos píxeles estaban localizados en cada línea radial con origen el centroide de la lesión y centrados en p). Los 60 valores recogidos, se describían en una gráfica por posición (empezando en el interior del borde, y terminando en el exterior), siendo la pendiente de los valores de la luminosidad, una medida de brusquedad. Por lo tanto, para cada conjunto de 60 píxeles, se calculaba la pendiente de los valores de luminosidad por el método de mínimos cuadrados. Repitiendo el proceso para cada uno de los 40 puntos del límite (5 puntos por cada uno de los 8 segmentos en los que se dividía la lesión), resultando 40 valores correspondientes a las pendientes de cada lesión. Para agregar los 5 números obtenidos por cada segmento y obtener una puntuación simple para cada uno de los segmentos, se empleaba el valor máximo de agregación, es decir, que se empleaba para cada segmento, el máximo de las cinco pendientes. Obteniéndose 8 números reales que representaban el corte en el borde de cada uno de los 8 segmentos definidos. Además se necesitaba algún método de decisión para decidir cuáles de esos números reales representaban segmentos con cortes bruscos y cuáles representan segmentos con un desvanecimiento gradual. Para ello, determinó un umbral T de forma que cada valor por encima de T era considerado afilado o brusco, y cada valor por debajo de T era considerado borroso.

- Otras medidas objetivas y cuantitativas del borde en lesiones, incluyen gradientes de lesión, gradientes de lesión – piel, máximo y mínimo del ancho de borde, valores de luminancia a lo largo de una dirección radial desde el centro de la lesión hasta el límite o los alrededores de la piel y la diferencia entre la media de los valores de color en las regiones de la periferia dentro y fuera de los contornos de la línea, entre otros [26].

2.4. MÉTODOS DE CLASIFICACIÓN [10].

Un modelo de clasificación o clasificador es un algoritmo que permite decidir cuáles de un conjunto de instancias están relacionadas o no, por pertenecer a un mismo tipo o clase. El resultado del análisis es un valor numérico que indica el umbral por el cual limitar cada clase o el valor que indica directamente una de las clases.

2.4.1. Clasificadores más comunes.

A continuación se mencionan algunos de los métodos de clasificación más empleados:

- El Análisis Discriminante Lineal [11] [30], es un método equivalente al análisis de regresión. En donde la variable dependiente es categórica y tiene como categorías la etiqueta de cada uno de los grupos, y las variables independientes son continuas y determinan a qué grupos pertenecen los objetos. Este modelo trata de encontrar relaciones lineales entre las variables continuas que mejor discriminen en los grupos

datos a los objetos. Además, se trata de definir una regla de decisión que asigne un objeto nuevo, que no sabemos clasificar previamente, a uno de los grupos prefijados.

- El Árbol de Decisión, es un modelo de predicción que indica las acciones a realizar en función del valor de una o varias variables. Es una representación en forma de árbol cuyas ramas se bifurcan en función de los valores tomados por las variables y que terminan en una acción concreta. Se suele utilizar cuando el número de condiciones no es muy grande (en tal caso, es mejor utilizar una tabla de decisión). Un árbol de decisión tiene unas entradas las cuales pueden ser un objeto o una situación descrita por medio de un conjunto de atributos y a partir de esto devuelve una respuesta, que es la decisión tomada a partir de las entradas.
- El Clasificador Bayesiano, es un clasificador de patrones basado en teorías estadísticas de aprendizaje. El aprendizaje bayesiano calcula la probabilidad de cada hipótesis de los datos y realiza predicciones sobre estas bases. Es un aprendizaje casi óptimo, pero requiere grandes cantidades de cálculo debido a que el espacio de hipótesis es normalmente muy grande, o incluso puede ser infinito.
- La Red Neural Artificial [30], es una técnica de aproximación paramétrica útil para construir modelos de densidad. Está inspirado en la forma en que funciona el sistema nervioso de los animales, es decir, que es un sistema de interconexión de neuronas en una red que colabora para producir un estímulo de salida. El modelo de red neuronal habitual que utiliza este algoritmo consiste en una red con una capa de entrada con tantos nodos como entradas existan, una capa oculta con un número de nodos variable que dependerá de las características del problema y una capa de salida con tantos nodos como posibles salidas tenga.
- Análisis de componentes principales o Principal Component Analysis (PCA) [30] [42], es una técnica que busca reducir el número de variables, transformando el conjunto original de variables correlacionadas que tienen información en común, en un conjunto de variables no relacionadas, que no tienen repetición o redundancia entre ellas, llamado conjunto de componentes principales. Siendo las nuevas variables, combinaciones lineales de las anteriores. Técnicamente, busca la proyección según la cual los datos queden mejor representados en términos de mínimos cuadrados.
- Máquinas de soporte vectorial o Support Vector Machine (SVM) [35], son unas técnicas sencillas y de gran aplicación cuando se pretende construir un clasificador utilizando ejemplos. Dado un conjunto de ejemplos de entrenamiento o de muestras se pueden etiquetar las clases y entrenar una máquina de soporte vectorial para construir un modelo que prediga la clase de una nueva muestra. Intuitivamente, este modelo representa los puntos de muestra en el espacio, separando las clases por un espacio lo más amplio posible. Cuando las nuevas muestras se ponen en correspondencia con dicho modelo, en función de su proximidad pueden ser clasificadas a una u otra clase. Estas máquinas de soporte vectorial se encargan de obtener la frontera de decisión y son fáciles de manejar ya que únicamente hay que codificar la geometría de la frontera.

2.4.2. Evaluación de un clasificador.

Existen diversas formas de evaluar un mismo clasificador, entre ellas se encuentran muchas medidas cuantitativas basadas principalmente en los conceptos de verdadero/falso positivo/negativo. Para explicar estos conceptos, se emplea una Tabla de contingencia o Matriz de confusión de 2x2, mostrada en la Figura 16:

		-- REALIDAD --	
		POSITIVO	NEGATIVO
-- PREDICCIÓN --	POSITIVO	VERDADERO POSITIVO	FALSO POSITIVO
	NEGATIVO	FALSO NEGATIVO	VERDADERO NEGATIVO

Figura 16 – Tabla de contingencia o Matriz de confusión.

Considerando un problema de predicción de clases binario, en la que los resultados se etiquetan como positivos (p) o negativos (n), existen cuatro posibles resultados con el clasificador binario propuesto. Si el resultado de una exploración es p y el valor real también es p , entonces se conoce como un Verdadero Positivo (VP); sin embargo, si el valor real es n , entonces se conoce como un Falso Positivo (FP). De igual modo, se tiene un Verdadero Negativo (VN) cuando, tanto la exploración como el valor real son n y un Falso Negativo (FN) cuando el resultado de la predicción es n , pero el valor real es p .

Una vez aclarados los conceptos verdadero/falso positivo/negativo, a continuación se muestran algunos de las medidas que evalúan a un clasificador:

- Medida XOR [28], definida por la Ecuación 5 – XOR.:

$$XOR = \frac{FP + FN}{VP + VN} \times 100\%$$

Ecuación 5 – XOR.

- Medida del valor predictivo positivo [28] [36], calculada con la Ecuación 6:

$$\text{Valor predictivo positivo} = \frac{VP}{VP + FP} \times 100\%$$

Ecuación 6 – Valor predictivo positivo.

- Medida del valor predictivo negativo [36], descrita por la Ecuación 7:

$$\text{Valor predictivo negativo} = \frac{VN}{VN + FN} \times 100\%$$

Ecuación 7 – Valor predictivo negativo.

Capítulo 3. CAPTURA Y ACONDICIONAMIENTO DE LA IMAGEN.

En este capítulo se explica cómo se obtienen las imágenes de las lesiones que constituyen la base de datos de este documento, nombradas como imagen original de la lesión. Además, se detallan los pasos a seguir para acondicionar la imagen original de la lesión, para obtener una imagen adecuada para segmentar, llamada imagen original recortada. Lograr ambas imágenes es importante, puesto que son empleadas en los análisis realizados a lo largo del proyecto.

3.1. CAPTURA DE LA IMAGEN DE LA LESIÓN.

En este proyecto, las imágenes aportadas por el Departamento de Dermatología de la Universidad de Missouri – Rolla se realizaron empleando un dermatoscopio en inmersión por aceite. Este proceso permite que las estructuras subsuperficiales de la piel sean visualizadas y que las imágenes captadas sean comparables y reproducibles. Las imágenes obtenidas a partir de las lesiones mediante este procedimiento, constituyen la base de datos empleada en este proyecto y dichas imágenes son nombradas como imagen original de la lesión.

A continuación, se exponen algunas de las imágenes que componen la base de datos del proyecto, indicando que la Figura 18, la Figura 19 y la Figura 22 corresponden a melanomas; mientras que la Figura 17, la Figura 20 y la Figura 21 corresponden a no melanomas:



Figura 17 – Lesión 1: No melanoma.



Figura 18 – Lesión 2: Melanoma.

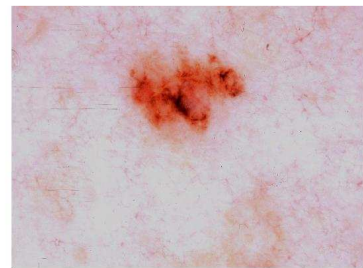


Figura 19 – Lesión 3: Melanoma.

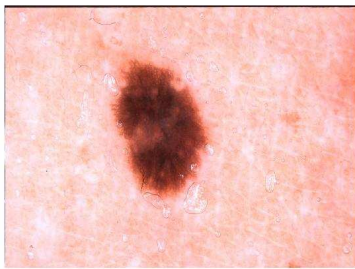


Figura 20 – Lesión 4: No melanoma.



Figura 21 – Lesión 5: No melanoma.



Figura 22 – Lesión 6: Melanoma.

A simple vista, una persona que no esté muy especializada o simplemente que no esté acostumbrada a observar melanomas, no puede diferenciar las lesiones entre melanomas y no melanomas tan fácilmente como un dermatólogo especializado.

3.2. ACONDICIONAMIENTO DE LA IMAGEN DE LA LESIÓN.

Tras la digitalización, algunas imágenes presentan bordes o márgenes que no deben tenerse en cuenta durante los análisis. Algunas de estas imágenes se presentan en las figuras anteriores, es decir, la Figura 20, la Figura 21 y la Figura 22. Por ello, se procura que la lesión, normalmente situada en el centro de la imagen, permanezca centrada y para conseguir un ajuste genérico, se quiere eliminar un margen lo más pequeño posible en cada imagen. La imagen resultante del ajuste, es llamada imagen original recortada de la lesión.

En las siguientes figuras se puede observar el cambio producido en las imágenes, tras realizar el ajuste eliminando los bordes:

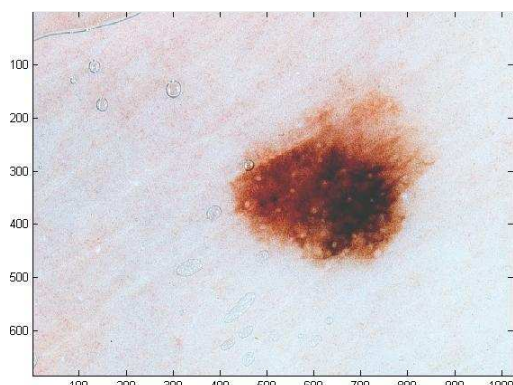


Figura 23 – Lesión 7: Imagen original.

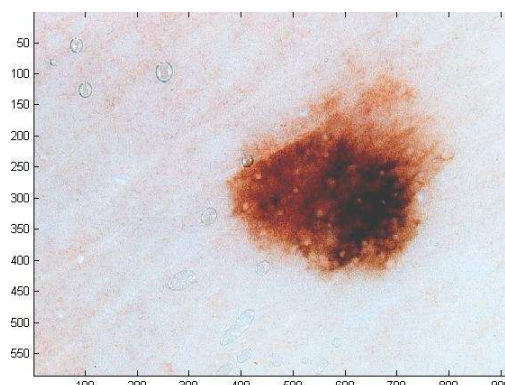


Figura 24 – Lesión 7: Imagen original recortada.

En esta ocasión, para poder apreciar la variación, se debe examinar el valor numérico de los ejes mostrado en las figuras. En donde se observa que la Figura 23 tiene unas dimensiones de 684x1024 píxeles y la Figura 24 tiene unas dimensiones menores a la anterior, correspondiente a unas dimensiones de 585x925 píxeles, aproximadamente de 50 píxeles menos en cada margen.

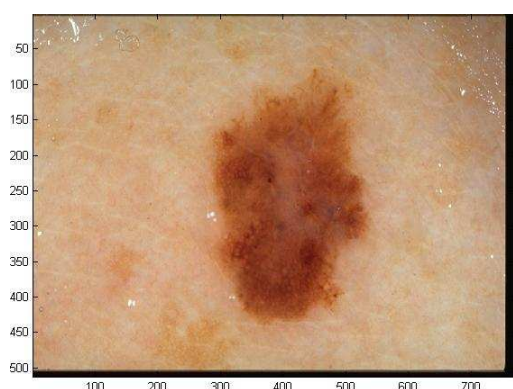


Figura 25 – Lesión 8: Imagen original.

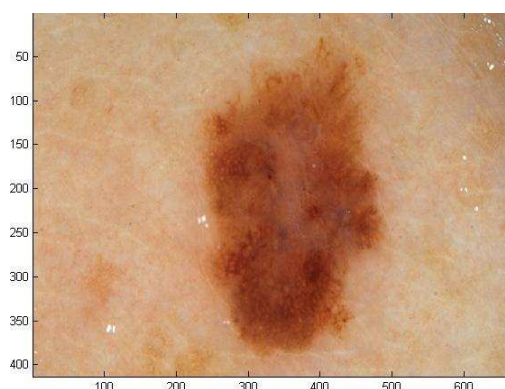


Figura 26 – Lesión 8: Imagen original recortada.

En este otro caso, el cambio es apreciable en el margen inferior y derecho de la imagen, además de notarse en el valor numérico de los ejes. Se puede observar que la Figura 25 tiene unas dimensiones de 512x768 píxeles y la Figura 26 tiene unas dimensiones menores, es decir, de 413x669 píxeles, aproximadamente de 50 píxeles menos en cada margen. Este valor de 50 píxeles, fue seleccionado de forma que elimina los márgenes de las imágenes influyendo lo mínimo posible en cada lesión.

Para los siguientes capítulos o análisis, se debe tener en cuenta que la imagen original es aquella imagen que puede contener bordes o márgenes, mientras que la imagen original recortada es aquella que tiene eliminado dicho margen o borde.

Además, para poder reducir el tiempo durante algunos análisis, como sucede en el primer análisis de la segmentación, se emplea también la imagen sampleada. Que es aquella imagen de la que se toman píxeles cada cierto espacio, es decir, que no se toma la imagen completa con todos los píxeles, si no que se toma la imagen quitando algunos píxeles.

A continuación, se muestra un ejemplo para comprender mejor esta explicación. En donde se observa la imagen original recortada y su imagen sampleada, puesto que primeramente se realiza el recorte y luego el muestreo de la imagen. Este ejemplo es una de las pocas lesiones en la que se distingue el muestreo:

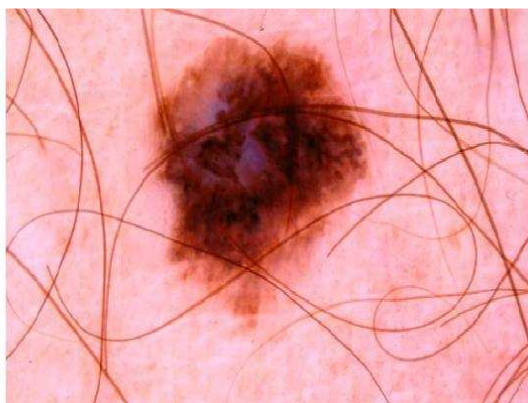


Figura 27 – Lesión 9: Imagen original recortada.

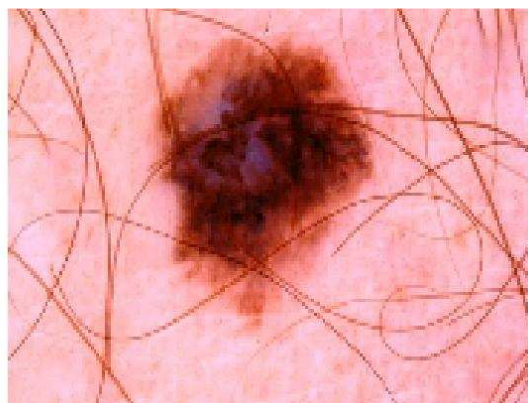


Figura 28 – Lesión 9: Imagen original recortada y sampleada.

Como se puede observar, el muestreo solo es apreciable en los artificios exteriores a la lesión, en este caso, el pelo. En la Figura 27, se puede advertir que el pelo es un trazo continuo, mientras que en la Figura 28, el trazo del pelo es discontinuo, o lo que coloquialmente se conoce por pixelado (en esta ocasión, levemente). Como se ha mencionado anteriormente, el muestreo en la lesión apenas es apreciable visualmente. El muestreo empleado para cada imagen de lesión, fue realizado por cada tres píxeles en cada banda de la imagen (1:3).

Capítulo 4. TÉCNICAS DE SEGMENTACIÓN DE LA IMAGEN.

La segmentación de una imagen consiste en separar el objeto principal del resto, en este caso, se quiere separar la lesión de la piel sana.

Existen varios métodos para segmentar una imagen, de entre todos ellos, se emplea el método por agrupación (clustering). Pero, para realizar una correcta segmentación de la imagen, se debe realizar un preprocesado, la propia segmentación y un postprocesado que se explican a lo largo de este capítulo.

4.1. COMBINACIÓN DE IMÁGENES RGB.

Una imagen está formada por tres bandas o imágenes monocromáticas. En este proyecto, la imagen viene dada en representación RGB (cuyas siglas en inglés corresponden a Red, Green, Blue, que se refieren a rojo, verde, azul).

A continuación se muestran dos lesiones, indicando de cada una de ellas, la imagen original recortada (Figura 29 y Figura 33) y las respectivas bandas RGB que la componen (Figura 30, Figura 31 y Figura 32, para una lesión y para la otra, Figura 34, Figura 35 y Figura 36):



Figura 29 – Lesión 10: Imagen original recortada.

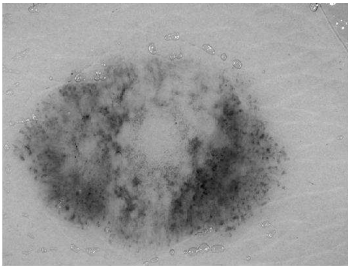


Figura 30 – Lesión 10: Banda roja.

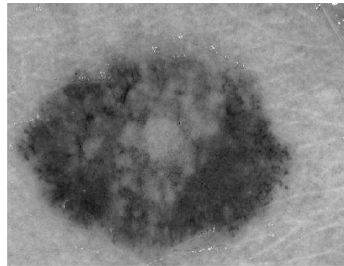


Figura 31 – Lesión 10: Banda verde.

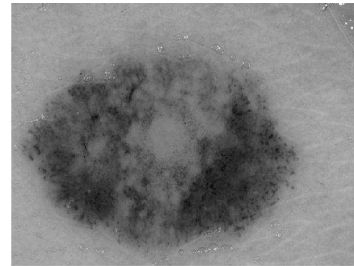


Figura 32 – Lesión 10: Banda azul.

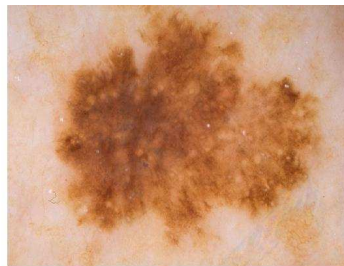


Figura 33 – Lesión 11: Imagen original recortada.

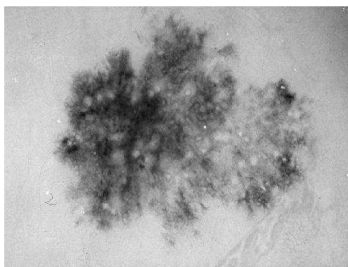


Figura 34 – Lesión 11: Banda roja.

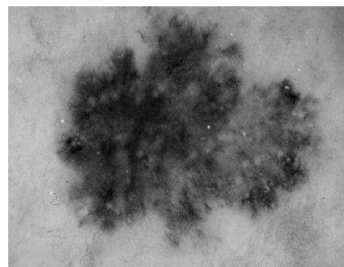


Figura 35 – Lesión 11: Banda verde.

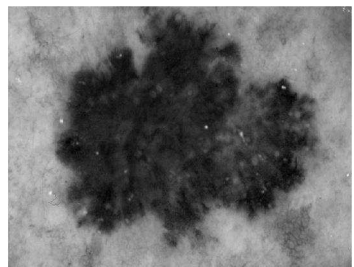


Figura 36 – Lesión 11: Banda azul.

En ambos casos, se observa que las estructuras internas de cada lesión son más definidas en la imagen en banda azul, que en las otras dos bandas. Además, el contorno de la lesión se define correctamente en las bandas verde y azul, intensificándose de nuevo, en la banda azul. Por estos motivos, como preprocesado de la segmentación, se podría escoger la banda azul de cada imagen original RGB de cada lesión.

Pero, como se ha podido apreciar, dependiendo de la banda en la que se vaya a trabajar, la lesión puede definirse mejor o peor. Por tanto, se decide mejorar el contraste de la imagen, optando por calcular la mejor combinación de las bandas RGB de forma que aumente el contraste entre las diferentes estructuras de la lesión.

4.1.1. Mezcla de Gaussianas.

Para encontrar la mejor combinación de las imágenes RGB se aplica una modificación del método que emplea la proyección del espacio original RGB que maximiza la bimodalidad de los datos proyectados mediante un histograma.

Este método es el Independent Histogram Pursuit (IHP) [27] [28], basado en un algoritmo lineal que estima un conjunto de combinaciones en las que proyectar las bandas RGB de la imagen. Transformando la imagen RGB a un espacio, en muchos casos de componente independiente y adaptativa, donde la lesión y la piel sana presentan una gran discriminación.

La mejor combinación lineal, es aquella que muestra la separación de las dos medias del histograma obtenido por la proyección de píxeles en esa línea, como una separación máxima. Por tanto, en lugar de aplicar una combinación general a todas las imágenes, se debe calcular la combinación propia de cada imagen y para ello, se emplean histogramas basados en la propia imagen. Siendo el umbral de cada histograma, el valor del histograma mínimo entre los dos picos de cada histograma [40].

Este método mejora el contraste de cada imagen analizada, sin emplear parámetros genéricos, si no que emplea para cada imagen sus propios parámetros. En el caso que se está estudiando, para cada tipo de piel, se mejora el contraste con respecto a la lesión, ya que para una piel clara, el contraste es alto, mientras que para una piel oscura, se necesitaría un mayor contraste. Por lo que este método, mejora el contraste dependiendo de la imagen que se analice y diferencia la lesión dependiendo del color de piel en la que se encuentre [28].

Sin embargo, este método es muy complejo ya que al obtener el histograma de la imagen, para alcanzar un mejor resultado se debe suavizar el histograma y maximizar el espacio que se obtiene entre las dos curvas [27] [28].

Por tanto, en este proyecto, para encontrar la mejor combinación lineal de las imágenes RGB proyectadas en esa línea, se analizan cada una de las imágenes originales recortadas y sampleadas de cada lesión para reducir el análisis y se aplica la Ecuación 8:

$$\text{Matriz combinada} = x \cdot R + y \cdot G + z \cdot B$$

Ecuación 8 – Análisis para encontrar la mejor combinación lineal para proyectar las bandas RGB.

En donde:

La *Matriz combinada* representa la combinación lineal de las imágenes RGB. Los parámetros R, G, B corresponden a cada una de las bandas roja, verde y azul, respectivamente. Los otros parámetros x, y, z proceden de la relación de coordenadas esféricas con coordenadas cartesianas [10], aplicadas para encontrar el eje en donde la proyección de la imagen sea la mejor. Dicha relación de coordenadas corresponde a la Ecuación 9:

$$\begin{aligned}x &= r \cdot \sin \theta \cdot \cos \varphi \\y &= r \cdot \sin \theta \cdot \sin \varphi \\z &= r \cdot \cos \theta\end{aligned}$$

Ecuación 9 – Relación entre coordenadas esféricas y coordenadas cartesianas.

En donde:

El parámetro r representa el radio de una esfera cualquier, la variable θ representa al ángulo polar o colatitud y la variable φ representa el azimuth. Con estos tres parámetros se pueden definir todos los posibles puntos de una esfera, como se muestra en la Figura 37:

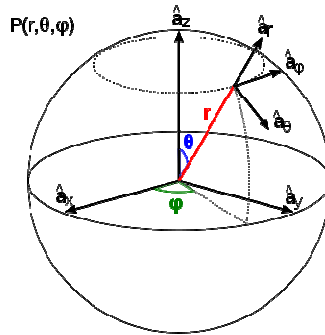


Figura 37 – Representación gráfica de la relación entre coordenadas esféricas y coordenadas cartesianas [10].

De la figura anterior, se observa que cuando el radio es constante, los valores de cada variable esférica son las mostradas en la Ecuación 10:

$$\begin{aligned}r &= cte \\0 &\leq \theta \leq \pi \\0 &\leq \varphi \leq 2\pi\end{aligned}$$

Ecuación 10 – Valores que toman las variables esféricas.

En nuestro caso, el radio no es significativo ya que definiría la misma proyección, por lo que es fijado a 1 y además, por similitud se puede reducir el rango de θ y φ . Obteniendo los valores indicados en la Ecuación 11, para cada variable esférica:

$$\begin{aligned}r &= 1 \\0 &\leq \theta \leq \pi \\0 &\leq \varphi \leq \pi\end{aligned}$$

Ecuación 11 – Valores simplificados que toman las variables esféricas en el programa.

Por lo tanto, la Ecuación 8 representa todas las posibles combinaciones a partir de las bandas de la imagen original recortada y sampleada. A nosotros nos interesa, obtener los parámetros x, y, z que aporten la mejor discriminación de la lesión y la piel en la imagen proyectada.

Para adquirir un mejor entendimiento del problema, a continuación, se observan algunos de los tipos de histograma que podría tener una imagen. Donde se muestran los efectos a consecuencia de modificar los parámetros asociados a las bandas R , G , B .

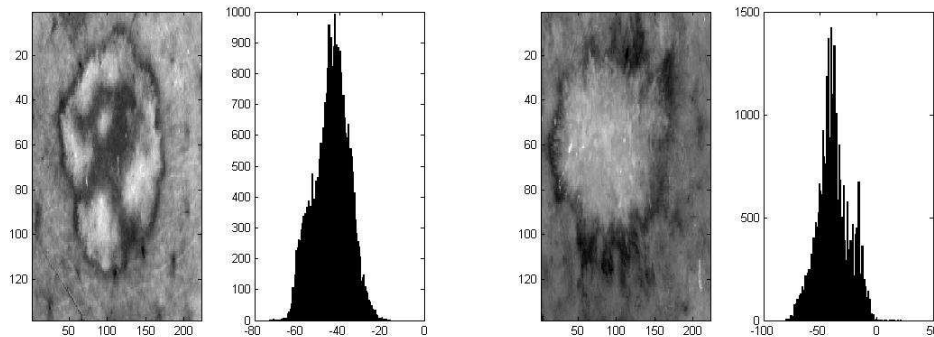


Figura 38 – Lesiones 12 y 13: Combinación cualquiera cuyo histograma diferencia poco la lesión de la piel.

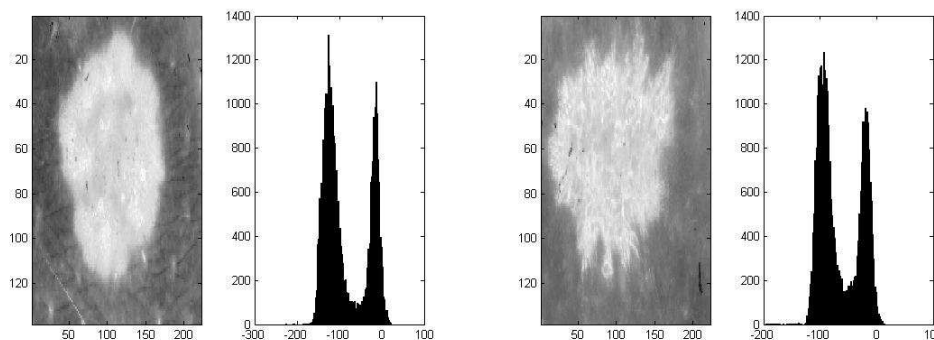


Figura 39 – Lesiones 12 y 13: Combinación cualquiera cuyo histograma diferencia muy bien la lesión de la piel.

Como se puede apreciar, existen dos tipos de histogramas, un histograma en donde la lesión se diferencia poco de la piel, como muestra la Figura 38 y un histograma en donde la lesión y la piel están muy bien diferenciadas, como muestra la Figura 39, el cual es bimodal.

De los histogramas en donde la lesión y la piel están muy bien diferenciadas se puede observar que se tienen aproximadamente dos Gaussianas, una de ellas correspondiente a la piel y la otra a la lesión. Por ello, de la Mezcla de Gaussianas se puede obtener la media y la varianza de cada una de cada una de las Gaussianas.

En nuestro caso, Matlab tiene una instrucción que realiza el cálculo de la media y la varianza de una Mezcla de Gaussianas. Dicha instrucción es: "gmdistribution".

4.1.2. Obtención de los parámetros de la combinación adecuada de la imagen.

Con la media y la varianza de cada una de las Gaussianas obtenidas mediante un algoritmo de Mezcla de Gaussianas de cada combinación de la imagen, se calcula el índice de Fisher [10], que corresponde a la Ecuación 12:

$$S = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2}$$

Ecuación 12 – Coeficiente de asimetría de Fisher o índice de Fisher.

En donde:

El parámetro S corresponde al índice de Fisher, μ_1 y μ_2 corresponden a las medias de las Gaussianas y σ_1^2 y σ_2^2 corresponden a sus respectivas varianzas. El índice es grande cuando la diferencia entre las medias es grande (indicando que las Gaussianas están muy separadas entre sí) y la suma de sus varianzas muy pequeña (presentando una media de distancias de los datos lo más pequeña posible con respecto a la media aritmética).

Los valores x, y, z asociados a la combinación en el que el índice de Fisher fue máximo, se consideraron como los valores óptimos que discriminan la lesión para una imagen dada.

Los índices asociados a las diferentes combinaciones se pueden expresar en una gráfica tridimensional, como por ejemplo la representada en la Figura 40. En la que se observa como todos los posibles valores convergen en un punto máximo.

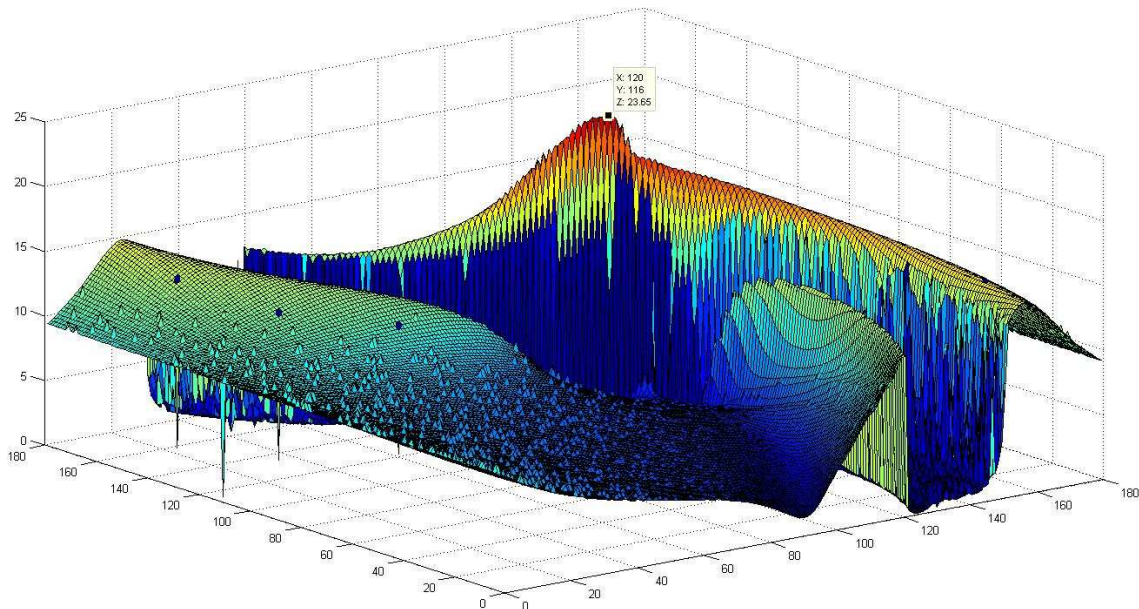


Figura 40 – Representación gráfica tridimensional de las posibles combinaciones.

El punto máximo de la figura anterior, se encontró cuando se tenían los valores mostrados a continuación en la Ecuación 13:

$$\begin{aligned} \varphi &= 116^\circ \\ \theta &= 120^\circ \\ S &= 23.6539 \end{aligned}$$

Ecuación 13 – Valores φ, θ, S de una lesión cualquiera.

En donde:

Los valores de φ y θ atienden a la Ecuación 11 y el valor de S corresponde al valor máximo obtenido aplicando el índice de Fisher, es decir, empleando la Ecuación 12.

Por último, aplicando los resultados obtenidos hasta ahora, en la Ecuación 9 y la Ecuación 8, en ese orden, se obtiene para cada imagen, los valores de los parámetros x , y , z y la imagen con mejor contraste.

Para el caso anterior, los valores x , y , z asociados a las variables θ y φ , son los indicados por la Ecuación 14:

$$\begin{aligned}x &= -0.3796 \\y &= 0.7784 \\z &= -0.5000\end{aligned}$$

Ecuación 14 – Valores x , y , z de una lesión cualquiera.

Por tanto, a continuación, se muestran varios ejemplos de lesiones, en donde se presenta la imagen original (Figura 41 y Figura 44), la imagen original recortada (Figura 42 y Figura 45) y la imagen con mejor combinación (es decir, con mayor índice de Fisher, Figura 43 y Figura 46), en donde se observa una mejora en el contraste:



Figura 41 – Lesión 14: Imagen original.



Figura 42 – Lesión 14: Imagen recortada.

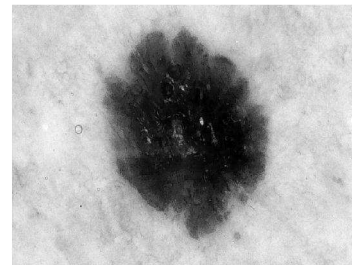


Figura 43 – Lesión 14: Imagen con mejor contraste.



Figura 44 – Lesión 15: Imagen original.



Figura 45 – Lesión 15: Imagen recortada.

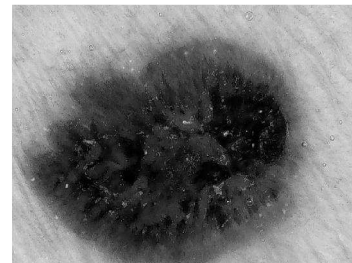


Figura 46 – Lesión 15: Imagen con mejor contraste.

Una vez encontrados los valores de la mejor combinación de la imagen de cada lesión, se aplican dichos valores a la imagen original recortada y se procede a realizar la segmentación a partir de la mejor combinación de la imagen original recortada.

4.2. IMAGEN SEGMENTADA (PREVIA Y DEFINITIVA).

La segmentación se debe realizar a partir de la mejor combinación obtenida a partir de la imagen original recortada.

En este proyecto, se realiza la comparación de dos métodos de segmentación:

- El algoritmo de Otsu.

El algoritmo de Otsu descrito previamente en el capítulo 2, proporciona un valor óptimo o umbral que utilizamos para separar la lesión de la piel sana. Para esta separación se asignan los píxeles con un valor inferior al umbral, a un grupo y los píxeles con valor superior al umbral, al otro grupo.

- Clustering mediante Mezcla de Gaussianas.

Para realizar este segundo método de clasificación utilizamos la función de Matlab llamada "gmdistribution.fit", la cual proporciona un campo vectorial denominado "cluster" que indica la pertenencia de un píxel a una clase u otra. Este vector es redimensionado al tamaño original de la imagen utilizando la instrucción "reshape".

Aplicando ambos métodos a dos lesiones distintas, representadas en la Figura 47 y Figura 50, se obtienen los siguientes resultados:



Figura 47 – Lesión 16: Imagen original recortada.

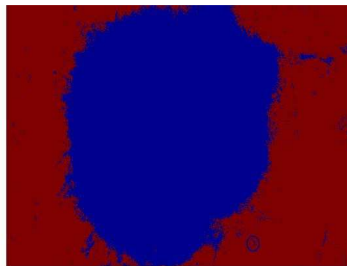


Figura 48 – Lesión 16: Imagen segmentada (algoritmo de Otsu).

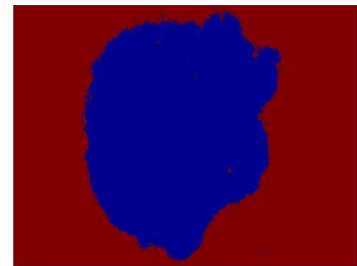


Figura 49 – Lesión 16: Imagen segmentada (Matlab, función "cluster").



Figura 50 – Lesión 17: Imagen original recortada.

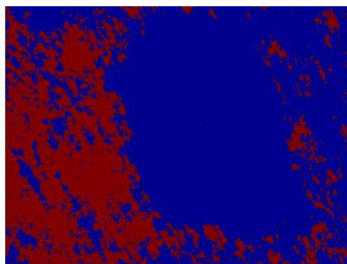


Figura 51 – Lesión 17: Imagen segmentada (algoritmo de Otsu).

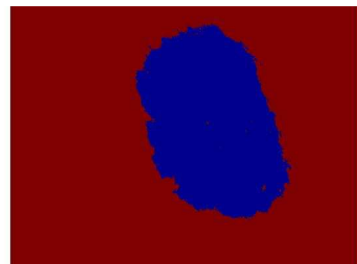


Figura 52 – Lesión 17: Imagen segmentada (Matlab, función "cluster").

En estas imágenes se puede observar que en algunas lesiones el método de Otsu proporciona una correcta segmentación, como muestra la Figura 48, aunque existe un pequeño número de lesiones cuya segmentación no fue satisfactoria, como se observa en la Figura 51. Por otro lado, el método de agrupación o clustering dado por la Mezcla de Gaussianas de Matlab, proporcionó un mejor resultado en las segmentaciones de las distintas lesiones, como exponen la Figura 49 y la Figura 52. Por lo que en este proyecto, se emplea este segundo método de segmentación. La imagen obtenida tras realizar la segmentación, es nombrada como imagen segmentada previa.

Dependiendo de la orientación de la línea de proyección, el resultado de la segmentación puede dar lugar a dos posibilidades, o bien la piel o bien la lesión, es indexada con el valor 1.

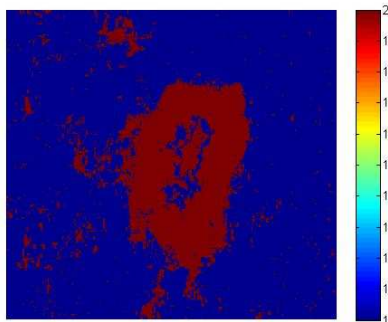


Figura 53 – Lesión 18: Segmentada previa.

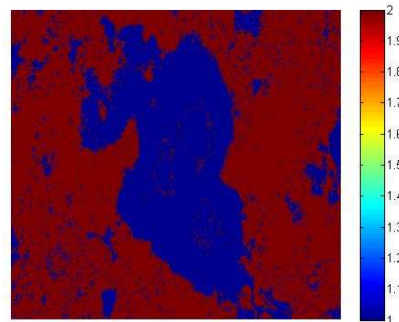


Figura 54 – Lesión 19: Segmentada previa.

Este hecho se observa en las figuras anteriores, en donde se aprecia que la diferencia se encuentra en el valor de la máscara binaria resultante. A nosotros nos interesa por consistencia la máscara binaria, en donde la lesión tenga un valor final de 1 y la piel tenga un valor final de 0. Es decir, que en la Figura 53, nos interesa simplemente reducir el valor de la imagen segmentada previa, mientras que en la Figura 54, nos interesa invertir y reducir los valores de la imagen segmentada previa.

4.2.1. Definición de la máscara binaria.

La imagen segmentada con máscara binaria es la imagen que diferencia los píxeles pertenecientes a la piel de los pertenecientes a la lesión, haciendo que los valores de la piel correspondan a ceros y que los valores de la lesión correspondan a unos.

Para conseguir que el análisis sea automático y obtener una imagen segmentada con máscara binaria consistente, es decir, que los píxeles asociados al valor 1 correspondan a la lesión en todas las imágenes, dividimos la imagen segmentada previa en nueve bloques aproximadamente del mismo tamaño. Este proceso se puede observar en la Figura 55 y en la Figura 56 mostradas a continuación:

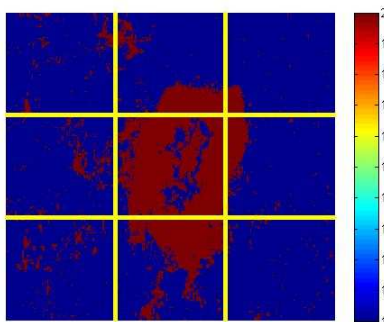


Figura 55 – Lesión 18: Segmentada previa dividida en nueve partes.

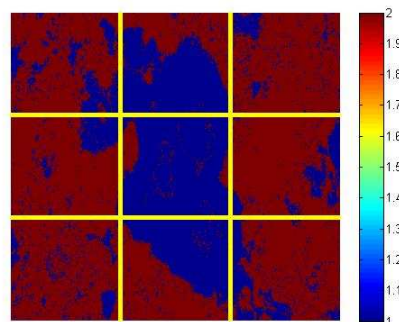


Figura 56 – Lesión 19: Segmentada previa dividida en nueve partes.

De esos nueve bloques en los que se ha dividido la imagen, sólo nos importan los bloques de las esquinas y el bloque central de la imagen ya que la lesión en nuestras imágenes está en el centro de las mismas.

La media calculada del conjunto de los cuatro bloques de las esquinas se compara con la media del bloque central y si la diferencia entre ambas es:

- Mayor que cero, significa que estamos en el caso en que la lesión tiene valor 2 y la piel tiene valor 1, por tanto para obtener la máscara binaria de ceros y unos, simplemente es necesario restarle la unidad a la imagen. Como se puede observar en la Figura 57, que es el resultado de la segmentada con máscara binaria de la Figura 55 correspondiente a la segmentada previa de una lesión.
- Mientras que si es menor que cero, quiere decir que estamos en el caso en que la lesión tiene valor 1 y la piel tiene valor 2, por tanto para obtener la máscara binaria de unos y ceros es necesario sustraer 1 e invertir la imagen. Como en el caso de la Figura 58, que es el resultado de la segmentada con máscara binaria de la Figura 56 correspondiente a la segmentada previa de otra lesión.

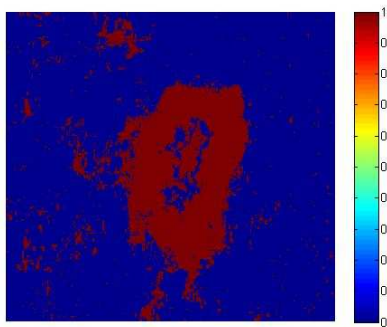


Figura 57 – Lesión 18: Segmentada con máscara binaria.

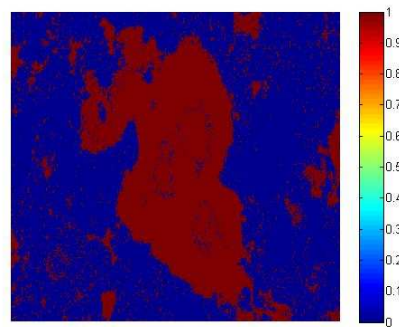


Figura 58 – Lesión 19: Segmentada con máscara binaria.

Una vez conseguida la imagen segmentada con máscara binaria de ceros y unos se mejoró la precisión de la segmentación. Para obtener la imagen segmentada definitiva, se deben eliminar los artificios que no pertenezcan a la lesión y rellenar los huecos que se encuentren en el interior de la lesión, empleando para ello la imagen segmentada con máscara binaria.

4.2.2. Eliminación de los artificios que no pertenecen a la lesión.

Para este cálculo, la imagen segmentada con máscara binaria, es considerada como un conjunto de islas o artificios, como por ejemplo las lesiones mostradas en la Figura 57 y la Figura 58. Siendo cada isla un conjunto de píxeles conectados, en donde la lesión corresponde al islote de mayor tamaño. Para determinar automáticamente cuál de estas islas corresponde a la lesión, se emplea la función morfológica de Matlab llamada "bwlablel" que asigna un valor numérico idéntico a los píxeles de una misma isla y diferente a píxeles de distintas islas. De esta forma, se determina la lesión calculando mediante un bucle el número de píxeles

asociados a cada isla (identificados por un mismo valor numérico) y seleccionando la isla de mayor tamaño.

De este modo, en las siguientes figuras, correspondiente a dos lesiones diferentes, se puede comparar la imagen segmentada con máscara binaria que posee artificios (Figura 59 y Figura 61) con su respectiva imagen segmentada con máscara binaria en la que los artificios están numerados (Figura 60 y Figura 62):

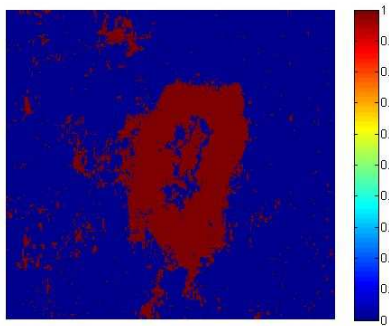


Figura 59 – Lesión 18: Segmentada con máscara binaria.

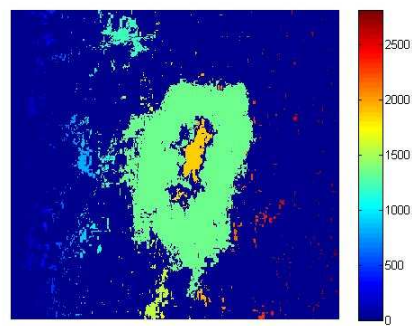


Figura 60 – Lesión 18: Segmentada con máscara binaria y artificios.

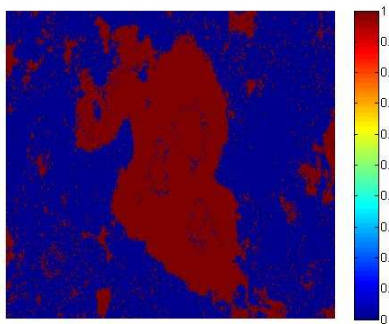


Figura 61 – Lesión 19: Segmentada con máscara binaria.

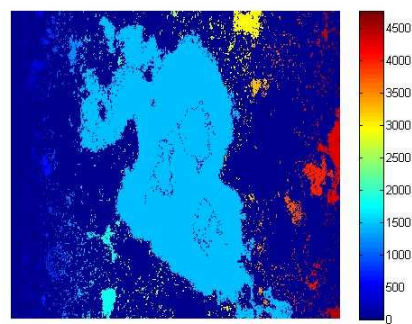


Figura 62 – Lesión 19: Segmentada con máscara binaria y artificios.

Como se puede observar de las figuras anteriores, cada isla es asociada a un valor numérico. Al comparar la barra de colores con los colores de la imagen de la lesión se observan varias tonalidades, debido a que cada color significa una isla diferente. Cabe apreciar que se suelen encontrar más de 2000 islas o artificios.

Para eliminar las islas que no pertenecen a la lesión, simplemente se toman los valores de la isla de mayor tamaño, ya que corresponde a la lesión. La nueva imagen obtenida, es decir, la imagen segmentada con máscara binaria y sin artificios, será la nueva imagen segmentada.

4.2.3. Rellenado de los huecos que se encuentran en el interior de la lesión.

A partir de la última imagen segmentada, es decir, tras haberse eliminado los artificios que se encontraban en la periferia de la lesión, se debe completar o rellenar la lesión resultando justamente la segmentación definitiva de la imagen original recortada. En este caso con Matlab, se emplea una instrucción que rellena automáticamente los huecos de cualquier área de una imagen. Dicha instrucción es “imfill”.

A continuación, en las siguientes figuras se muestran dos ejemplos de lesiones. En las que se puede comparar la imagen segmentada con máscara binaria, sin artificios, pero con huecos (Figura 63 y Figura 65) y su correspondiente imagen segmentada definitiva (Figura 64 y Figura 66):

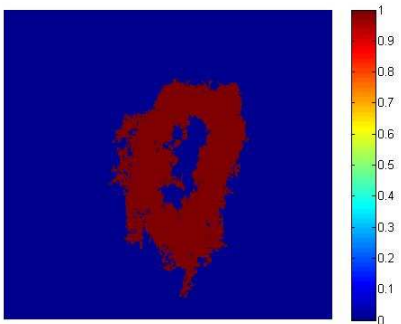


Figura 63 – Lesión 18: Segmentada con máscara binaria, sin artificios y con huecos.

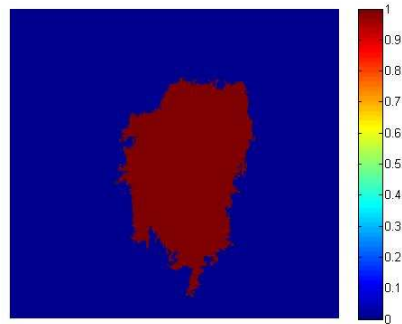


Figura 64 – Lesión 18: Segmentada definitiva.

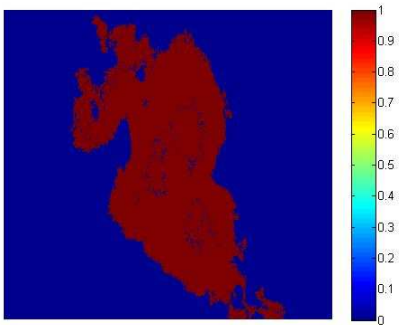


Figura 65 – Lesión 19: Segmentada con máscara binaria, sin artificios y con huecos.

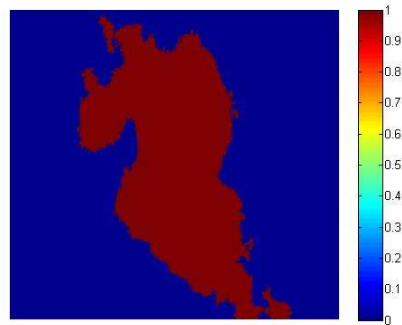


Figura 66 – Lesión 19: Segmentada definitiva.

Tras realizar todos estos pasos se obtiene la segmentada definitiva a partir de la imagen original recortada de la lesión. A continuación, se muestran de cada ejemplo analizado durante este capítulo, la imagen original recortada, representada en la Figura 67 y la Figura 69, y su respectiva imagen segmentada definitiva, expuesta en la Figura 68 y la Figura 70:

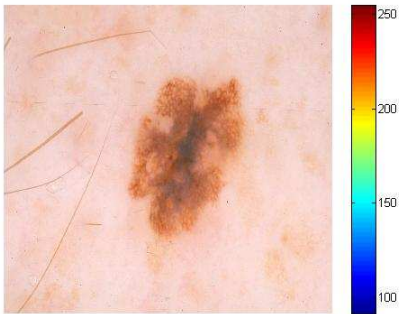


Figura 67 – Lesión 18: Imagen original recortada.

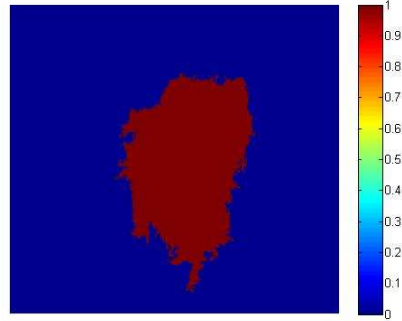


Figura 68 – Lesión 18: Imagen segmentada definitiva.

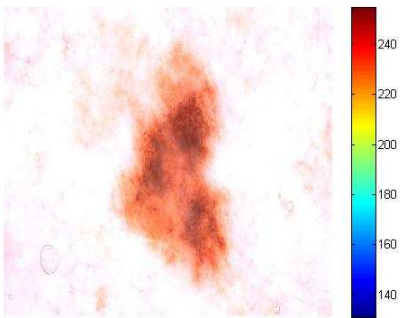


Figura 69 – Lesión 19: Imagen original recortada.

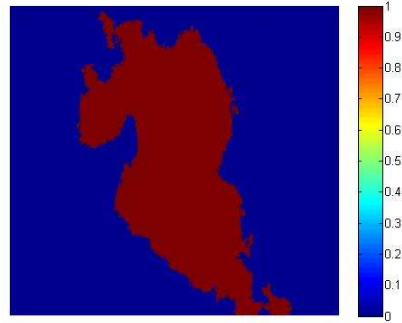


Figura 70 – Lesión 19: Imagen segmentada definitiva.

Capítulo 5.

EXTRACCIÓN DE

CARACTERÍSTICAS.

El objetivo de este capítulo consiste en analizar cada una de las lesiones y obtener una serie de características que posteriormente ayuden a determinar si una lesión es un melanoma o no.

Para ello, en este capítulo, se explica cómo se calculan cada una de las características que se emplean a lo largo del proyecto. Estas características se obtienen siguiendo la regla ABCD de la dermatoscopia, en la medida de lo posible, y a partir de otras características propuestas de estudios o artículos mencionados anteriormente en el capítulo 2.

Para que los datos extraídos sean coherentes y las características se extraigan del mismo patrón de imagen, hay que tener en cuenta tanto en las bandas de la imagen como en la imagen combinada, que la lesión debe poseer valores correspondientes al color oscuro de la imagen y la piel debe poseer valores correspondientes al color claro de la misma imagen. Por este motivo, se deben estandarizar dichas imágenes de manera que sus valores se encuentren entre 0 (color más oscuro) y 1 (color más claro).

A continuación, se explican los pasos que se deben realizar para desarrollar la estandarización de estas imágenes. Dada una imagen (ya sea una de las tres bandas o la imagen combinada de la imagen original recortada de una lesión), se debe dividir dicha imagen en nueve bloques aproximadamente del mismo tamaño, como ya se realizó con la imagen segmentada previa en el anterior capítulo. De los nueve bloques en los que se ha dividido la imagen, sólo nos interesan los bloques de las esquinas y el bloque central de dicha imagen puesto que la lesión en las imágenes se encuentra en el centro de las mismas.

La media calculada del conjunto de los cuatro bloques de las esquinas, a la que llamaremos media de la piel, se compara con la media del bloque central, a la que llamaremos media de la lesión. Existiendo dos posibilidades:

- La primera, que la media de la piel sea mayor a la media de la lesión. Esto quiere decir que la piel tiene valores mucho más altos que la lesión, y simplemente se tienen que ajustar los valores de los píxeles a valores entre 0 y 1 para realizar la estandarización de la imagen. Para ello, se resta a cada valor de píxel, el valor mínimo del conjunto de píxeles que forman la imagen y los nuevos valores obtenidos se dividen entre el máximo del nuevo conjunto. La imagen estandarizada será el resultado del proceso y tendrá valores entre 0 y 1, correspondientes a la lesión y la piel, respectivamente.
- La segunda posibilidad, es que la media de la piel sea menor a la media de la lesión. Esto significa que la piel tiene valores mucho más bajos que la lesión, y por ello se deben ajustar los valores de los píxeles a valores entre 0 y 1 y luego invertir la imagen para realizar la estandarización de la imagen. Para ello, como en el caso anterior, se resta a cada valor de píxel, el valor mínimo del conjunto de píxeles que forman la imagen y los nuevos valores obtenidos se dividen entre el máximo del nuevo conjunto. La imagen estandarizada será el resultado de haber invertido la imagen obtenida, es decir, al sustraer de 1 la imagen que se quiere invertir. Así la imagen estandarizada tendrá valores entre 0 (correspondientes a la lesión) y 1 (correspondientes a la piel).

Las características que se mencionan durante este capítulo, se obtienen partiendo de:

- La imagen estandarizada de cada una de las bandas y de la imagen combinada o proyectada, procedentes de la imagen original recortada de la lesión que se desea analizar, puesto que contienen el valor de los píxeles de la imagen de la lesión estandarizados a valores entre 0 y 1.
- Y/o de la imagen de la segmentada definitiva de la imagen original recortada de la misma lesión, porque contiene el valor de las posiciones en las que se encuentra la lesión, por ser una máscara binaria.

Nota: A partir de este momento y para el resto del proyecto, cuando se mencionen a cada una de las bandas de la imagen original recortada de una lesión o a su imagen combinada o proyectada, nos estaremos refiriendo a sus respectivas imágenes estandarizadas.

5.1. ANÁLISIS DE LA ASIMETRÍA DE LA LESIÓN.

Para analizar la asimetría de una lesión, se caracterizan una serie de parámetros:

- La excentricidad [35].
- Y el coeficiente de asimetría sobre los ejes principales de la lesión [42].

Para calcular la excentricidad de una lesión, se utiliza la instrucción “regionprops” de Matlab. Una de sus propiedades es “eccentricity” cuyo resultado corresponde a un círculo cuando la excentricidad de la elipse es próxima a cero y corresponde a una línea cuando la excentricidad

de la elipse es próxima a uno. Esta propiedad se aplica a la imagen segmentada definitiva de la lesión, calculando así la excentricidad de dicha lesión.

Por otro lado, el índice de asimetría de la lesión se calcula mediante una combinación de instrucciones. Primeramente, a partir de la imagen segmentada definitiva, se trazan los ejes principal y secundario desde el punto central de la región perteneciente a la lesión. El eje principal corresponde al eje que maximiza la varianza de la lesión, mientras que el eje secundario es el eje perpendicular al eje principal. Por tanto, ambos ejes se cortan en el punto central de la lesión y forman un ángulo de 90 grados entre sí. En cada eje se proyecta la lesión, obteniéndose el coeficiente de asimetría de cada eje. Este coeficiente indica cómo de asimétrica es la lesión, es decir, que cuanto mayor sea, significa que la lesión es más asimétrica y cuanto menor sea el valor, más redonda es la lesión. En este caso, con Matlab, se emplea: para encontrar el punto central de la lesión, la instrucción "regionprops" y su propiedad "centroid"; para obtener los ejes principal y secundario, la instrucción "pcacov", que calcula los autovectores de la lesión y para calcular el coeficiente de asimetría, la instrucción "skewness".

A continuación, se muestra un ejemplo de lesión, en la que se expone:

- En la Figura 71, la imagen original recortada de la lesión.
- En la Figura 72, la imagen segmentada definitiva indicándole el centro de la lesión.
- En la Figura 73, la imagen segmentada definitiva en la que se presentan los ejes principal y secundario en color verde y amarillo, respectivamente.
- Y por último, en la Figura 74 y la Figura 75, la proyección de la lesión en sus ejes principal y secundario, respectivamente.

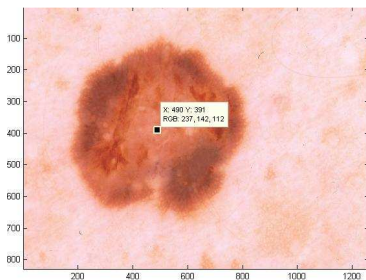


Figura 71 – Lesión 20: Imagen original recortada.

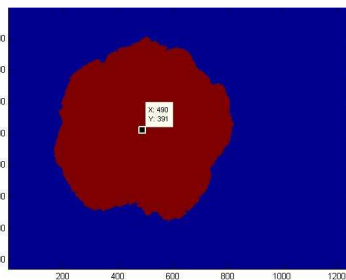


Figura 72 – Lesión 20: Imagen segmentada definitiva.

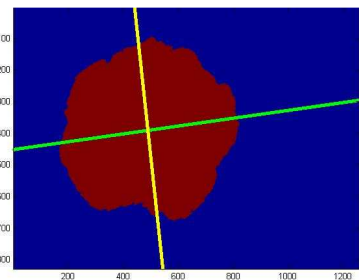


Figura 73 – Lesión 20: Imagen segmentada definitiva con ejes perpendiculares.

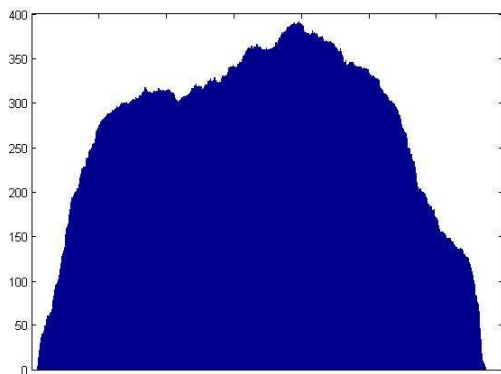


Figura 74 – Lesión 20: Histograma del eje principal.

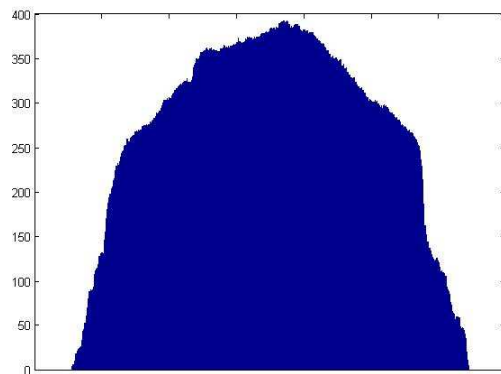


Figura 75 – Lesión 20: Histograma del eje secundario.

Por tanto, analizando la asimetría de la lesión, se obtienen tres características: ECCEN, SKEWPPAL y SKEWSEC.

5.2. ANÁLISIS DEL BORDE DE LA LESIÓN.

La hipótesis que postulamos es que el borde de la imagen, sea abrupto o suave, puede diferenciar el que una lesión sea melanoma o no. Nuestra hipótesis consiste en que las lesiones con bordes más abruptos tienen una probabilidad mayor de ser melanomas.

Para analizar el borde de una lesión, se caracteriza una serie de parámetros empleando dos aproximaciones distintas. Estas dos aproximaciones son:

- Mediante el ajuste por regresión lineal, empleando mínimos cuadrados o líneas de tendencia [26] [39].
- Mediante el ajuste por sigmoidea.

Para ambos métodos, se parte de la imagen segmentada definitiva y desde el punto central de la región de la lesión, se trazan ejes radialmente hasta el borde de la imagen. En este trabajo se calcularon un total de 36 ejes, con una variación de 10 grados cada uno, como se puede observar en la Figura 76 y en la Figura 77:

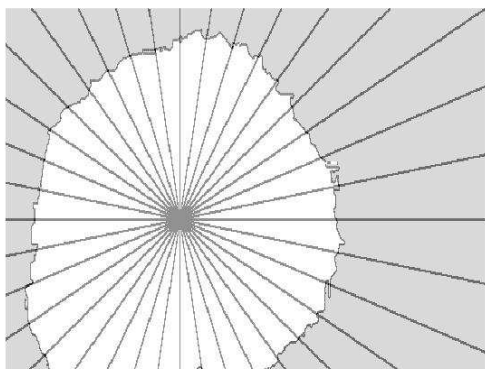


Figura 76 – Lesión 21: Imagen de la segmentada definitiva indicando los 36 ejes tomados.

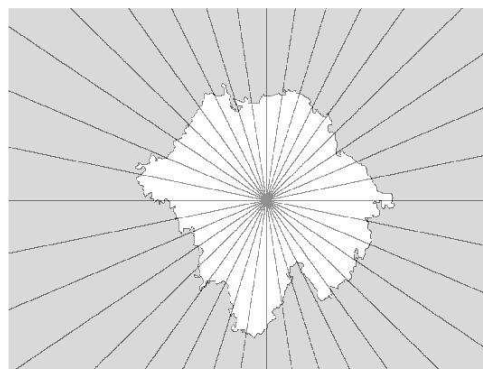


Figura 77 – Lesión 22: Imagen de la segmentada definitiva indicando los 36 ejes tomados.

En el cálculo de parámetros mediante el ajuste por regresión lineal, empleando mínimos cuadrados o líneas de tendencia, se pretende obtener de la imagen proyectada y de cada una de las tres bandas de la imagen original de la lesión, la pendiente de un conjunto de píxeles situados en cada eje. Este conjunto es obtenido previamente y está formado por N píxeles, en nuestro caso de 61 píxeles (ya que corresponden al píxel situado en la intersección del borde de la lesión con uno de los 36 ejes, y a los 30 píxeles, tanto interiores como exteriores del borde de la lesión, que pertenecen a ese mismo eje). Además, hay que tener en cuenta, que este conjunto de píxeles es tomado desde el interior de la lesión al exterior de la misma. El número de píxeles escogido se basa en el artículo desarrollado por Day *et al* [39]. Para calcular la pendiente de dicho conjunto de píxeles por una regresión lineal, se emplea la instrucción “regress” de Matlab. Dado que se tienen 36 ejes, se tendrán 36 pendientes por lesión como máximo, ya que en caso de que el eje no corte con el borde de la lesión o no existan 30 píxeles

interiores o exteriores, esa pendiente no se calculará. La siguiente lesión es un claro ejemplo de los casos mencionados:

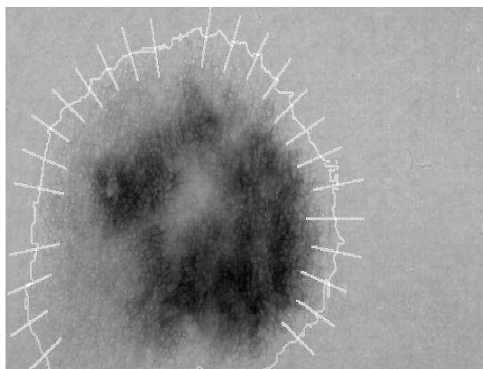


Figura 78 – Lesión 21: Banda roja de la lesión.

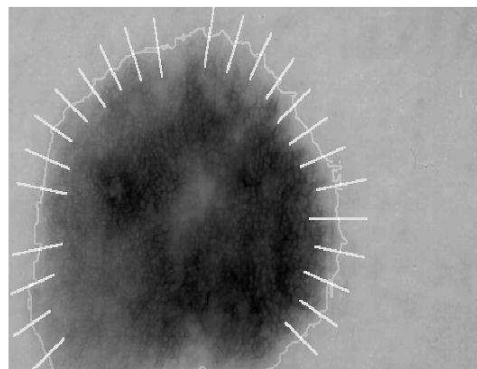


Figura 79 – Lesión 21: Banda verde de la lesión.

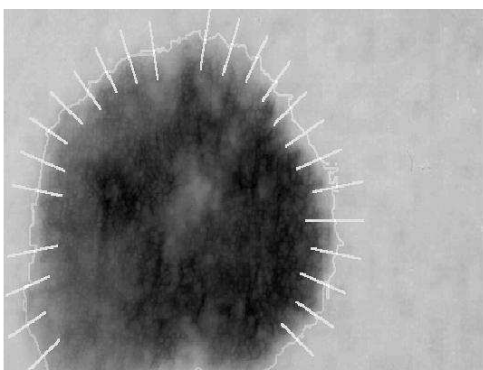


Figura 80 – Lesión 21: Banda azul de la lesión.

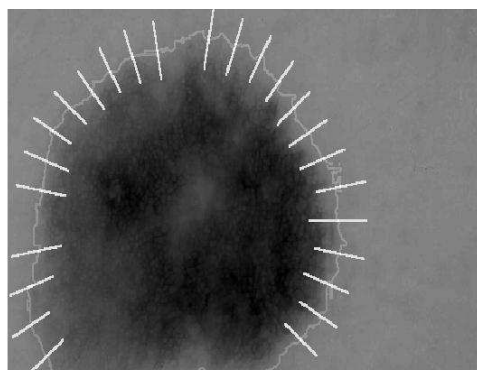


Figura 81 – Lesión 21: Mejor combinación o imagen proyectada.

Como se puede observar en las figuras anteriores, correspondientes a cada una de las tres bandas (Figura 78, Figura 79, Figura 80) y a la imagen proyectada (Figura 81) de una lesión, sólo se pueden calcular 25 pendientes de los 36 ejes, en cada caso. Esto se debe a que el resto de ejes o no corta con la lesión o los píxeles no son suficientes. Además, cada uno de los segmentos visibles en las figuras, corresponden al conjunto de píxeles tomados para el ajuste por regresión lineal.

Por tanto, los parámetros obtenidos para caracterizar el borde de una lesión, mediante el ajuste por regresión lineal, corresponden a:

- La media de las pendientes de la regresión lineal creada a partir del conjunto de los 61 píxeles explicados anteriormente, en la banda roja, verde, azul (REDPTEREGMN, GREENPTEREGMN, BLUEPTEREGMN) y en la combinación con mejor contraste calculada previamente (COMBPTEREGMN).
- La desviación típica de las pendientes de la regresión lineal creada a partir del conjunto de los 61 píxeles explicados anteriormente, en la banda roja, verde, azul (REDPTEREGSD, GREENPTEREGSD, BLUEPTEREGSD) y en la combinación con mejor contraste calculada previamente (COMBPTEREGSD).

- El primer cuartil (Q_1) de las pendientes de la regresión lineal creada a partir del conjunto de los 61 píxeles explicados anteriormente, en la banda roja, verde, azul (REDPTEREGQ1, GREENPTEREGQ1, BLUEPTEREGQ1) y en la combinación con mejor contraste calculada previamente (COMBPTEREGQ1).
- El segundo cuartil (Q_2) de las pendientes de la regresión lineal creada a partir del conjunto de los 61 píxeles explicados anteriormente, en la banda roja, verde, azul (REDPTEREGQ2, GREENPTEREGQ2, BLUEPTEREGQ2) y en la combinación con mejor contraste calculada previamente (COMBPTEREGQ2).
- El tercer cuartil (Q_3) de las pendientes de la regresión lineal creada a partir del conjunto de los 61 píxeles explicados anteriormente, en la banda roja, verde, azul (REDPTEREGQ3, GREENPTEREGQ3, BLUEPTEREGQ3) y en la combinación con mejor contraste calculada previamente (COMBPTEREGQ3).
- El rango intercuartílico (RQ) de las pendientes de la regresión lineal creada a partir del conjunto de los 61 píxeles explicados anteriormente, en la banda roja, verde, azul (REDPTEREGRQ, GREENPTEREGRQ, BLUEPTEREGRQ) y en la combinación con mejor contraste calculada previamente (COMBPTEREGRQ).
- El coeficiente de variación (CV) de las pendientes de la regresión lineal creada a partir del conjunto de los 61 píxeles explicados anteriormente, en la banda roja, verde, azul (REDPTEREGCV, GREENPTEREGCV, BLUEPTEREGCV) y en la combinación con mejor contraste calculada previamente (COMBPTEREGCV).

Es decir, que simplemente analizando el borde de la lesión, mediante el ajuste por regresión lineal, se obtienen 28 características:

REDPTEREGMN,	GREENPTEREGMN,	BLUEPTEREGMN,	COMBPTEREGMN,
REDPTEREGSD,	GREENPTEREGSD,	BLUEPTEREGSD,	COMBPTEREGSD,
REDPTEREGQ1,	GREENPTEREGQ1,	BLUEPTEREGQ1,	COMBPTEREGQ1,
REDPTEREGQ2,	GREENPTEREGQ2,	BLUEPTEREGQ2,	COMBPTEREGQ2,
REDPTEREGQ3,	GREENPTEREGQ3,	BLUEPTEREGQ3,	COMBPTEREGQ3,
REDPTEREGRQ,	GREENPTEREGRQ,	BLUEPTEREGRQ,	COMBPTEREGRQ,
REDPTEREGCV,	GREENPTEREGCV,	BLUEPTEREGCV	y COMBPTEREGCV.

En el cálculo de parámetros mediante el ajuste por sigmoidea, el procedimiento es análogo al anterior. Este método tiene la ventaja con respecto al anterior, de que no es necesario determinar un número fijo de píxeles, por lo que se toman todos los píxeles contenidos en cada eje desde el punto central de la lesión hasta el borde de la imagen (siempre en este sentido). En las siguientes figuras, se muestra una lesión con los ejes tomados para este ajuste, correspondiente a cada una de las tres bandas (Figura 82, Figura 83, Figura 84) y a la imagen proyectada (Figura 85):

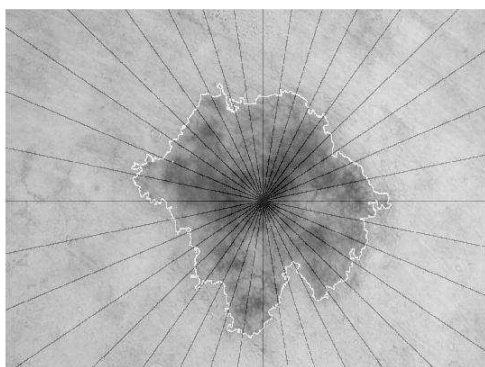


Figura 82 – Lesión 22: Banda roja de la lesión.

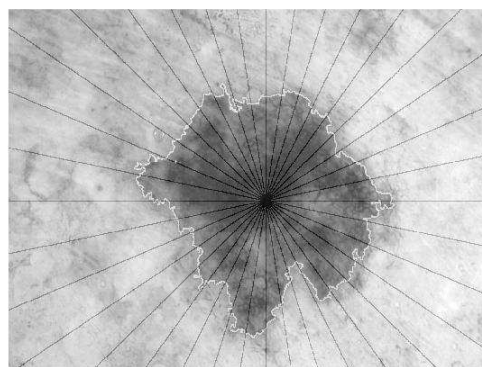


Figura 83 – Lesión 22: Banda verde de la lesión.

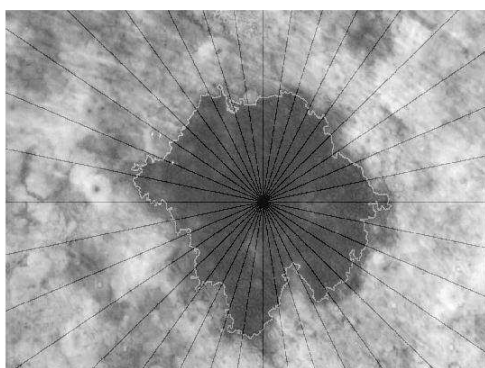


Figura 84 – Lesión 22: Banda azul de la lesión.

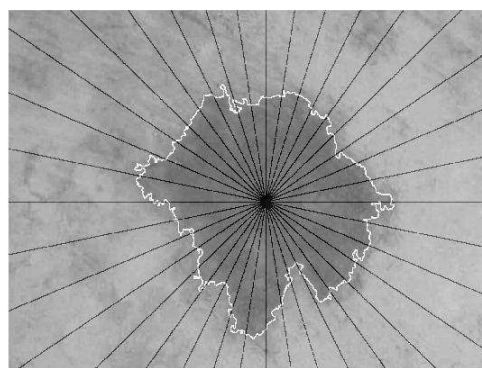


Figura 85 – Lesión 22: Mejor combinación o imagen proyectada.

Una sigmoidea genérica se define con la Ecuación 15:

$$P(t) = \frac{1}{1 + e^{-(a \cdot t - b)}}$$

Ecuación 15 – Sigmoidea genérica.

En donde:

La sigmoidea está representada por $P(t)$, el valor de a es el valor de la máxima pendiente y el valor de b es la posición en la que se encuentra dicha pendiente. Su representación gráfica corresponde a la Figura 86:

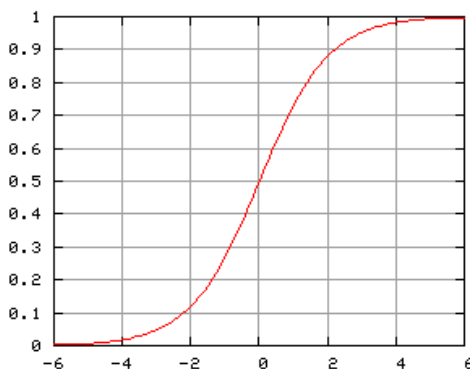


Figura 86 – Representación gráfica de la sigmoidea genérica [10].

Para realizar el ajuste por sigmoidea se recogen los valores de píxeles de cada una de las bandas de la imagen original de la lesión y de la imagen proyectada, y se aplica la Ecuación 15 directamente, ya que los valores de los píxeles se encuentran en un rango entre 0 y 1.

Para calcular los parámetros de la sigmoidea, se emplea la función “fminsearch” de Matlab. Esta función calcula automáticamente los parámetros que minimizan una función dada por el usuario. En nuestro caso, buscamos minimizar el error cometido entre la sigmoidea y los valores de los píxeles. Por tanto, la función “fminsearch” nos proporciona los valores a y b que mejor ajusta la sigmoidea a los datos. En particular, estamos interesados en el parámetro a de la Ecuación 15, correspondiente a la inclinación de la sigmoidea.

A continuación, se muestra el conjunto de píxeles perteneciente a uno de los ejes, en cada una de las bandas de la imagen original de la lesión (Figura 87, Figura 88, Figura 89) y de la imagen proyectada (Figura 90), ajustadas por su sigmoidea correspondiente:

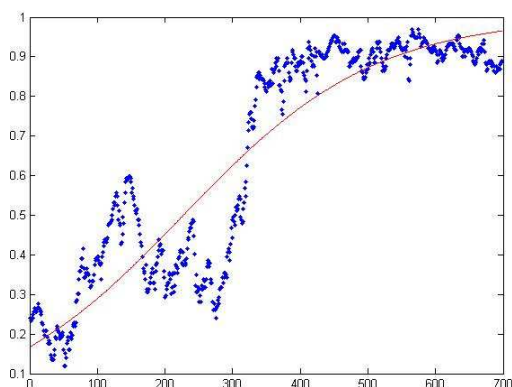


Figura 87 – Lesión 22: Píxeles en banda roja.

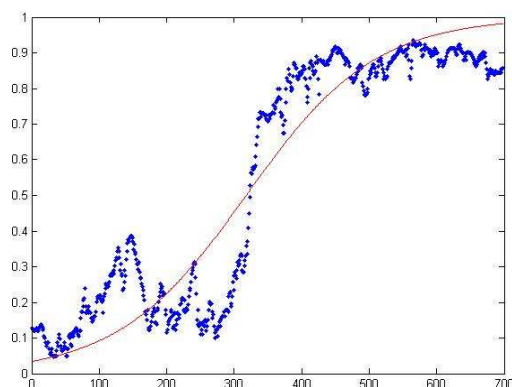


Figura 88 – Lesión 22: Píxeles en banda verde.

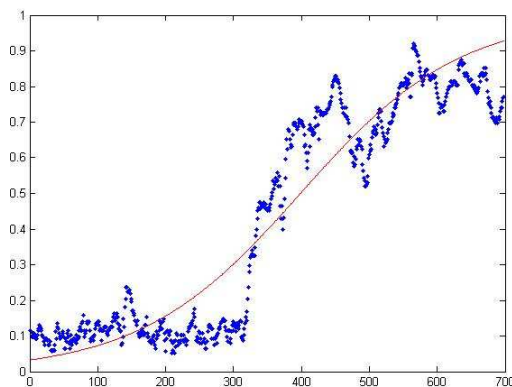


Figura 89 – Lesión 22: Píxeles en banda azul.

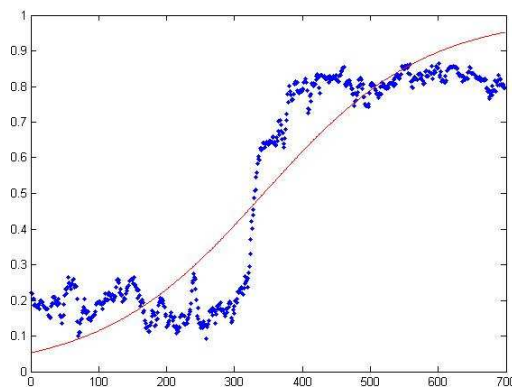


Figura 90 – Lesión 22: Píxeles de la mejor combinación o imagen proyectada.

Por tanto, los parámetros obtenidos para caracterizar el borde de una lesión, mediante el ajuste por sigmoidea, corresponden a:

- La media de las pendientes máximas de las sigmoideas creadas a partir de los píxeles tomados desde el centro de la lesión hasta el borde de la imagen, en la banda roja, verde, azul (REDPTESIGMN, GREENPTESIGMN, BLUEPTESIGMN) y en la combinación con mejor contraste calculada previamente (COMBPTESIGMN).
- La desviación típica de las pendientes máximas de las sigmoideas creadas a partir de los píxeles tomados desde el centro de la lesión hasta el borde de

la imagen, en la banda roja, verde, azul (REDPTESIGSD, GREENPTESIGSD, BLUEPTESIGSD) y en la combinación con mejor contraste calculada previamente (COMBPTESIGSD).

- El primer cuartil (Q_1) de las pendientes máximas de las sigmoideas creadas a partir de los píxeles tomados desde el centro de la lesión hasta el borde de la imagen, en la banda roja, verde, azul (REDPTESIGQ1, GREENPTESIGQ1, BLUEPTESIGQ1) y en la combinación con mejor contraste calculada previamente (COMBPTESIGQ1).
- El segundo cuartil (Q_2) de las pendientes máximas de las sigmoideas creadas a partir de los píxeles tomados desde el centro de la lesión hasta el borde de la imagen, en la banda roja, verde, azul (REDPTESIGQ2, GREENPTESIGQ2, BLUEPTESIGQ2) y en la combinación con mejor contraste calculada previamente (COMBPTESIGQ2).
- El tercer cuartil (Q_3) de las pendientes máximas de las sigmoideas creadas a partir de los píxeles tomados desde el centro de la lesión hasta el borde de la imagen, en la banda roja, verde, azul (REDPTESIGQ3, GREENPTESIGQ3, BLUEPTESIGQ3) y en la combinación con mejor contraste calculada previamente (COMBPTESIGQ3).
- El rango intercuartílico (RQ) de las pendientes máximas de las sigmoideas creadas a partir de los píxeles tomados desde el centro de la lesión hasta el borde de la imagen, en la banda roja, verde, azul (REDPTESIGRQ, GREENPTESIGRQ, BLUEPTESIGRQ) y en la combinación con mejor contraste calculada previamente (COMBPTESIGRQ).
- El coeficiente de variación (CV) de las pendientes máximas de las sigmoideas creadas a partir de los píxeles tomados desde el centro de la lesión hasta el borde de la imagen, en la banda roja, verde, azul (REDPTESIGCV, GREENPTESIGCV, BLUEPTESIGCV) y en la combinación con mejor contraste calculada previamente (COMBPTESIGCV).

Es decir, que simplemente analizando el borde de la lesión, mediante el ajuste por sigmoidea, se obtienen 28 características:

REDPTESIGMN,	GREENPTESIGMN,	BLUEPTESIGMN,	COMBPTESIGMN,
REDPTESIGSD,	GREENPTESIGSD,	BLUEPTESIGSD,	COMBPTESIGSD,
REDPTESIGQ1,	GREENPTESIGQ1,	BLUEPTESIGQ1,	COMBPTESIGQ1,
REDPTESIGQ2,	GREENPTESIGQ2,	BLUEPTESIGQ2,	COMBPTESIGQ2,
REDPTESIGQ3,	GREENPTESIGQ3,	BLUEPTESIGQ3,	COMBPTESIGQ3,
REDPTESIGRQ,	GREENPTESIGRQ,	BLUEPTESIGRQ,	COMBPTESIGRQ,
REDPTESIGCV,	GREENPTESIGCV,	BLUEPTESIGCV	y COMBPTESIGCV.

5.3. ANÁLISIS DEL COLOR Y DE LAS DIFERENTES ESTRUCTURAS DE LA LESIÓN.

En este proyecto, nuestro programa sólo se centra en evaluar la asimetría y el borde de una lesión, correspondientes a las dos primeras características de la regla ABCD de la dermatoscopia. Debido a que el resto de características requieren un nivel mucho más avanzado tanto de programación como de desarrollo y que para un proyecto fin de carrera de este tipo, no es necesario profundizar. Sin embargo, en esta subsección queremos indicar la importancia que estas características tienen a la hora de realizar un sistema más completo.

5.4. ANÁLISIS DE LOS PÍXELES CONTENIDOS EN LA LESIÓN.

Otras medidas propuestas en los artículos mencionados en el capítulo 2 analizan los píxeles contenidos en la lesión, mientras que otros analizan todos los píxeles de la imagen de la lesión.

Para ambos casos, se debe emplear cada una de las tres bandas y la imagen segmentada definitiva de la imagen original recortada de la lesión. Puesto que las primeras poseen el valor de los píxeles de la lesión en cada banda, mientras que la última contiene las posiciones de los píxeles que forman el conjunto de la lesión.

Analizando los píxeles contenidos en la lesión [40] [42], se obtienen las siguientes características que corresponden a:

- El área de la lesión (AREA).
- El perímetro de la lesión (PERIM).
- La relación entre $4 \cdot \pi \cdot \text{área}$ por unidad de perímetro al cuadrado de la lesión (FRAG).
- La media de los valores de los píxeles de la lesión en la banda roja, verde y azul (REDMN, GREENMN, BLUEMN).
- La desviación típica de los valores de los píxeles de la lesión en la banda roja, verde y azul (RESD, GREENSD, BLUESD).
- El primer cuartil (Q_1) de los valores de los píxeles de la lesión en la banda roja, verde y azul (REDQ1, GREENQ1, BLUEQ1).
- El segundo cuartil (Q_2) de los valores de los píxeles de la lesión en la banda roja, verde y azul (REDQ2, GREENQ2, BLUEQ2).
- El tercer cuartil (Q_3) de los valores de los píxeles de la lesión en la banda roja, verde y azul (REDQ3, GREENQ3, BLUEQ3).
- El rango intercuartílico (RQ) de los valores de los píxeles de la lesión en la banda roja, verde y azul (REDRQ, GREENRQ, BLUERQ).
- El coeficiente de variación (CV) de los valores de los píxeles de la lesión en la banda roja, verde y azul (REDCV, GREENCV, BLUECV).
- La media del gradiente de los valores de los píxeles del borde de la lesión en la banda roja, verde y azul (REDGRADMN, GREENGRADMN, BLUEGRADMN).

- La desviación típica del gradiente de los valores de los píxeles del borde de la lesión en la banda roja, verde y azul (REDGRADSD, GREENGRADSD, BLUEGRADSD).
- El primer cuartil (Q_1) del gradiente de los valores de los píxeles del borde de la lesión en la banda roja, verde y azul (REDGRADQ1, GREENGRADQ1, BLUEGRADQ1).
- El segundo cuartil (Q_2) del gradiente de los valores de los píxeles del borde de la lesión en la banda roja, verde y azul (REDGRADQ2, GREENGRADQ2, BLUEGRADQ2).
- El tercer cuartil (Q_3) del gradiente de los valores de los píxeles del borde de la lesión en la banda roja, verde y azul (REDGRADQ3, GREENGRADQ3, BLUEGRADQ3).
- El rango intercuartílico (RQ) del gradiente de los valores de los píxeles del borde de la lesión en la banda roja, verde y azul (REDGRADRQ, GREENGRADRQ, BLUEGRADRQ).
- El coeficiente de variación (CV) del gradiente de los valores de los píxeles del borde de la lesión en la banda roja, verde y azul (REDGRADCV, GREENGRADCV, BLUEGRADCV).

Por tanto, analizando los píxeles contenidos en la lesión, se obtienen 45 características:

AREA,	PERIM	FRAG,
REDMN,	GREENMN,	BLUEMN,
REDS,	GREENSD,	BLUESD,
REDQ1,	GREENQ1,	BLUEQ1,
REDQ2,	GREENQ2,	BLUEQ2,
REDQ3,	GREENQ3,	BLUEQ3,
REDRQ,	GREENRQ,	BLUERQ,
REDCV,	GREENCV,	BLUECV,
REDGRADMN,	GREENGRADMN,	BLUEGRADMN,
REDGRADSD,	GREENGRADSD,	BLUEGRADSD,
REDGRADQ1,	GREENGRADQ1,	BLUEGRADQ1
REDGRADQ2,	GREENGRADQ2,	BLUEGRADQ2
REDGRADQ3,	GREENGRADQ3,	BLUEGRADQ3,
REDGRADRQ,	GREENGRADRQ,	BLUEGRADRQ,
REDGRADCV,	GREENGRADCV,	y BLUEGRADCV.

Por otro lado, analizando todos los píxeles de la imagen de la lesión [42], se obtienen las siguientes características:

- La entropía de información en la banda roja, verde y azul de la lesión.

Para obtener estas características, es necesario emplear la ecuación de la entropía de información, que es una medida de cantidad de información [10] y cuya definición viene dada por la Ecuación 16:

$$H(x) = - \sum_i p(x_i) \cdot \log_2 p(x_i)$$

Ecuación 16 – Entropía de información.

En donde:

La entropía de información está representada por $H(x)$ y $p(x_i)$ representa la probabilidad, en este caso, de cada valor de píxel de la imagen de la lesión.

Por tanto, analizando todos los píxeles de la imagen de la lesión, se obtienen tres características: ENTROPIA_RED, ENTROPIA_GREEN y ENTROPIA_BLUE.

Capítulo 6. TÉCNICAS DE CLASIFICACIÓN Y EVALUACIÓN.

En este capítulo se explican los métodos de clasificación y de evaluación empleados. Tanto para clasificar las características obtenidas previamente en el capítulo 5, como para clasificar las lesiones a partir de dichas características. Además de la evaluación de las lesiones y del clasificador empleado.

Pero, antes de comenzar, se deben distinguir varios conceptos:

- Primero, el concepto de clasificador, que se define como el método empleado para clasificar y que debe tener un umbral con el que comparar y clasificar. En nuestro caso se clasifican tanto las características obtenidas y extraídas de la lesión, como las propias lesiones.
- Segundo, el concepto de evaluación, que se refiere a la evaluación de una lesión y a la evaluación del clasificador. La evaluación de una lesión, corresponde a las veces que se debe repetir la clasificación, puesto que para nosotros, evaluar una lesión significa extraer dicha lesión y realizar la clasificación sin ella, para conocer el resultado que se obtendría si no perteneciera a la base de datos. Mientras que evaluar el clasificador, corresponde a la consecuencia del proceso por el cual se sabe si la clasificación ha tenido éxito o por el contrario se deben encontrar mejores resultados.

6.1. CLASIFICACIÓN DE LAS CARACTERÍSTICAS MEDIANTE UN CONTRASTE DE HIPÓTESIS.

El contraste de hipótesis es un método estadístico muy conocido, que se emplea para saber si dos grupos, en este caso, llamado clases, se pueden diferenciar significativamente o no, dependiendo del resultado del parámetro P. Por tanto:

- Si P es mayor (o mucho mayor) que 0.05, quiere decir que las dos clases que se quieren comparar se diferencian mal (o muy mal).
- Si P es menor (o mucho menor) que 0.05, quiere decir que las dos clases que se quieren comparar se diferencian bien (o muy bien).

Gracias a este método podemos clasificar las características extraídas de la lesión y obtenidas en el capítulo 5, para saber cuáles de todas ellas son las que permiten diferenciar significativamente entre las dos clases (melanoma o no melanoma).

En nuestro caso, con Matlab, el contraste de hipótesis se realiza con la instrucción “ttest2”.

6.2. CLASIFICACIÓN DE LAS LESIONES MEDIANTE UN DISCRIMINANTE LINEAL [11] [30].

El clasificador o método de análisis discriminante lineal, permite clasificar las lesiones entre melanomas y no melanomas, empleando las características obtenidas y extraídas en el capítulo 5.

Para la clasificación, el método de análisis discriminante lineal emplea tres grupos:

- Un grupo llamado conjunto “sample” o de “muestra”.
- Otro grupo llamado conjunto “training” o de “entrenamiento”.
- Y el último grupo llamado conjunto de “clase”.

En nuestro caso, con Matlab, el clasificador o método de análisis llamado discriminante lineal, se puede realizar con la instrucción “classify”.

En este proyecto, se emplea el clasificador discriminante lineal de dos formas. Una forma consiste en emplear todas las características obtenidas y otra forma consiste en emplear de todas las características obtenidas, solo una selección de características.

6.2.1. Clasificación de las lesiones con todas las características obtenidas.

En esta clasificación, se emplean todas las características (obtenidas y extraídas en el capítulo 5) para clasificar las lesiones entre melanomas y no melanomas, empleando el clasificador discriminante lineal, en donde sus tres grupos son:

- El conjunto “sample” o de “muestra”, correspondiente al conjunto de todas características de la lesión que se está evaluando, es decir, de la lesión que se extrae de la base de datos.
- El conjunto “training” o de “entrenamiento”, correspondiente al conjunto de todas las características de todas las lesiones de la base de datos, excepto de la lesión que se está evaluando.
- El conjunto de “clase”, correspondiente al resultado real de las lesiones que forman el conjunto “training” o de “entrenamiento”.

6.2.2. Clasificación de las lesiones con características seleccionadas del conjunto total de características.

En esta otra clasificación, primeramente se debe obtener la combinación de características correspondientes a cada lesión que alcancen un alto porcentaje de aciertos. Para cada lesión, la combinación de características será diferente y para seleccionar dichas características, se emplea el llamado “forward selection” o “selección hacia delante” [11].

En este proyecto, el proceso seguido, se explica a continuación:

- Para obtener la combinación de características correspondientes a cada lesión, se debe evaluar dicha lesión. Para ello, se debe extraer del conjunto la lesión a evaluar y aplicar el “forward selection” o “selección hacia delante” con el resto de lesiones. Este método consiste en:
 - Clasificar las lesiones restantes empleando el clasificador discriminante lineal y aplicarlo a cada una de las características individualmente, para obtener el porcentaje de aciertos de cada característica comparando el resultado de la clasificación con el verdadero valor de las lesiones. Siendo nuestro conjunto de lesiones al aplicar este método, las lesiones restantes. Por tanto, los grupos con este método son:
 - ✓ El conjunto “sample” o de “muestra”, correspondiente a cada característica de la lesión a evaluar, es decir, de la lesión que se extrae del conjunto restante de lesiones.
 - ✓ El conjunto “training” o de “entrenamiento”, correspondiente a la misma característica de todas las lesiones restantes, excepto de la lesión que se está evaluando.

- ✓ El conjunto de “clase”, correspondiente al resultado real de las lesiones restantes que forman el conjunto “training” o de “entrenamiento”.
- Una vez obtenidos todos los porcentajes individuales de todas las características, se escoge la característica correspondiente al mayor porcentaje de aciertos y se combina con el resto de características. De nuevo, se deben clasificar las lesiones restantes empleando el clasificador discriminante lineal y aplicarlo a la combinación de características, en este caso, de dos características (posteriormente de tres y así sucesivamente). De la misma manera que antes, se obtiene el porcentaje de aciertos de la combinación de características comparando el resultado de la clasificación con el verdadero valor de las lesiones restantes.
 - ✓ El conjunto “sample” o de “muestra”, correspondiente a la combinación de características de la lesión a evaluar, es decir, de la lesión que se extrae del conjunto de lesiones restante.
 - ✓ El conjunto “training” o de “entrenamiento”, correspondiente a la misma combinación de características de todas las lesiones restantes, excepto de la lesión que se está evaluando.
 - ✓ El conjunto de “clase”, correspondiente al resultado real de las lesiones restantes que forman el conjunto “training” o de “entrenamiento”.
- El proceso de añadir características a la combinación debe continuar hasta que el porcentaje de la combinación de características deje de mejorar, ya que llegado este momento, por más que se añadan características, el porcentaje no aumentará. Por tanto, las características seleccionadas de cada lesión corresponden a dicha combinación de características.
- El proceso para obtener la combinación de características correspondiente a cada lesión, se debe repetir para todas las lesiones de la base de datos. Una vez seleccionadas las características para cada lesión, se deben clasificar las lesiones con sus correspondientes características seleccionadas aplicando el clasificador discriminante lineal. En donde los tres grupos son:
 - El conjunto “sample” o de “muestra”, correspondiente al conjunto de características seleccionadas de la lesión que se está evaluando, es decir, de la lesión que se extrae de la base de datos.
 - El conjunto “training” o de “entrenamiento”, correspondiente al conjunto de características seleccionadas de todas las lesiones de la base de datos, excepto de la lesión que se está evaluando.
 - El conjunto de “clase”, correspondiente al resultado real de las lesiones que forman el conjunto “training” o de “entrenamiento”.

- El resultado obtenido será la clasificación de todas las lesiones, evaluadas con sus respectivas características, mediante el clasificador discriminante lineal.

6.3. EVALUACIÓN DE LAS LESIONES MEDIANTE VALIDACIÓN CRUZADA.

La evaluación de las lesiones mediante el clasificador discriminante lineal, ya sea con todas las características o con la selección de características, se realiza por medio del método de evaluación de validación cruzada o “cross validation”, cuyo nombre completo es “k – fold cross validation”. Este método de evaluación indica el número de veces que se debe repetir la clasificación, es decir, el número de veces que se debe emplear el clasificador.

En este caso, el número de veces que se debe emplear el clasificador corresponde al número de lesiones que se tiene en la base de datos, puesto que ya se ha explicado que evaluar una lesión consiste en extraer dicha lesión y realizar la clasificación del resto de lesiones (a este tipo de evaluación se le conoce como “one leave out”). Por tanto, para evaluar todas las lesiones de la base de datos, habría que repetir el proceso tantas veces como lesiones haya en ella, es decir, que se tienen “100 – fold cross validation” o dicho de otra manera, se tienen “100 validaciones cruzadas”.

6.4. EVALUACIÓN DEL CLASIFICADOR.

Para evaluar al clasificador, se emplean las siguientes medidas:

- La precisión, que indica la capacidad del clasificador para predecir los casos correctamente, es decir, pronosticando los casos realmente positivos como casos positivos y los realmente negativos como casos negativos. En el caso de una enfermedad, la precisión indica la proporción de casos acertados, ya sean casos positivos (sujetos enfermos) o casos negativos (sujetos sanos). Matemáticamente, se expresa mediante la Ecuación 17:

$$\text{Precisión} = \frac{VP + TN}{VP + FP + FN + TN}$$

Ecuación 17 – Precisión.

- La sensibilidad [10] [28] [36] [43] [44], que indica la capacidad del clasificador para predecir como casos positivos los casos realmente positivos. En el caso de una enfermedad, indica la proporción de enfermos correctamente identificados. Esto significa que la sensibilidad caracteriza la capacidad del examen para detectar la enfermedad en sujetos enfermos. Matemáticamente corresponde a la Ecuación 18:

$$\text{Sensibilidad} = \frac{VP}{VP + FN}$$

Ecuación 18 – Sensibilidad.

Y cuanto mayor sea numéricamente, mejor es la capacidad de detectar a los enfermos.

- La especificidad [10] [28] [36] [43] [44], que indica la capacidad del clasificador para predecir como casos negativos los casos realmente negativos. En caso de hablar de una enfermedad, indica la proporción de sanos correctamente identificados. Esto significa que la especificidad caracteriza la capacidad del examen para detectar la ausencia de la enfermedad en sujetos sanos. Matemáticamente queda definida con la Ecuación 19:

$$Especificidad = \frac{VN}{VN + FP}$$

Ecuación 19 – Especificidad.

Y cuanto mayor sea numéricamente, mejor es la capacidad de detectar a los sanos.

Capítulo 7.

RESULTADOS DE LOS

EXPERIMENTOS.

En este capítulo se presenta la información de la base de datos, las características extraídas de las lesiones y los resultados obtenidos al emplear:

- Un contraste de hipótesis, para saber qué características son lo bastante significativas como para poder clasificar las lesiones entre melanomas y no melanomas.
- Un análisis por discriminante lineal, empleando todas las características obtenidas y extraídas del capítulo 5, para clasificar las lesiones entre melanomas y no melanomas.
- Un análisis por discriminante lineal, empleando las características seleccionadas para cada lesión, para clasificar las lesiones entre melanomas y no melanomas.
- La evaluación del clasificador discriminante lineal para ambos modos, mediante las medidas de precisión, sensibilidad y especificidad calculadas para cada caso.

Es decir, que se presentan los resultados obtenidos con los experimentos mencionados en los capítulos anteriores.

7.1. BASE DE DATOS.

La base de datos empleada durante la realización del proyecto está formada por 100 imágenes dadas por el tutor y que corresponden a lesiones de pacientes a los que se le ha realizado una biopsia para conocer el estado de la lesión y poder determinar si son melanomas o no melanomas. Gracias a la biopsia realizada, se sabe que de las 100 lesiones, 29 lesiones corresponden a melanomas y las 71 lesiones restantes corresponden a no melanomas.

7.2. RESULTADO DE LA CLASIFICACIÓN DE LAS CARACTERÍSTICAS EXTRAÍDAS.

Las características obtenidas y extraídas de las lesiones a lo largo del capítulo 5 se presentan en la primera columna de la cada una de las siguientes tablas, agrupadas según el apartado correspondiente en el capítulo 5, es decir:

- Tabla 1 – Características obtenidas analizando la asimetría de la lesión.
- Tabla 2 – Características obtenidas analizando el borde de la lesión mediante el ajuste por regresión lineal.
- Tabla 3 – Características obtenidas analizando el borde de la lesión mediante el ajuste por sigmoidea.
- Tabla 4 – Características obtenidas analizando los píxeles contenidos en la lesión.
- Tabla 5 – Características obtenidas analizando todos los píxeles de la imagen de la lesión.

La segunda columna de cada una de estas tablas, corresponde al valor numérico del parámetro P del análisis de cada característica mediante el contraste de hipótesis. Dicho parámetro indica si dicha característica permite diferenciar muy bien o muy mal las lesiones entre melanoma y no melanoma. Para poder identificar las características que diferencian significativamente dichos grupos, se resaltan en amarillo las más significativas, es decir, las menores a 0.05.

En la tercera columna de cada una de las tablas, se indica el porcentaje de lesiones correctamente clasificadas con dicha característica de forma individual, mediante el clasificador discriminante lineal

CARACTERÍSTICAS	PARÁMETRO P DEL CONTRASTE DE HIPÓTESIS	PORCENTAJE DE LESIONES CORRECTAMENTE CLASIFICADAS
ECCEN	0.2425	53.0000
SKEWPPAL	0.0471	63.0000
SKEWSEC	0.3577	52.0000

Tabla 1 – Características obtenidas analizando la asimetría de la lesión.

RESULTADOS DE LOS EXPERIMENTOS.

CARACTERÍSTICAS	PARÁMETRO P DEL CONTRASTE DE HIPÓTESIS	PORCENTAJE DE LESIONES CORRECTAMENTE CLASIFICADAS
REDPTEREGMN	0.0476	58.0000
GREENPTEREGMN	0.1110	57.0000
BLUEPTEREGMN	0.4531	58.0000
COMBPTEREGMN	0.4028	58.0000
REDPTEREGSD	0.7427	50.0000
GREENPTEREGSD	0.5353	55.0000
BLUEPTEREGSD	0.1041	58.0000
COMBPTEREGSD	0.4349	52.0000
REDPTEREGQ1	0.0194	61.0000
GREENPTEREGQ1	0.0936	54.0000
BLUEPTEREGQ1	0.2767	57.0000
COMBPTEREGQ1	0.2883	52.0000
REDPTEREGQ2	0.0434	57.0000
GREENPTEREGQ2	0.1311	54.0000
BLUEPTEREGQ2	0.4910	56.0000
COMBPTEREGQ2	0.4696	57.0000
REDPTEREGQ3	0.0852	56.0000
GREENPTEREGQ3	0.1408	56.0000
BLUEPTEREGQ3	0.6462	62.0000
COMBPTEREGQ3	0.4849	55.0000
REDPTEREGRQ	0.8598	50.0000
GREENPTEREGRQ	0.9350	31.0000
BLUEPTEREGRQ	0.1651	54.0000
COMBPTEREGRQ	0.4029	58.0000
REDPTEREGCV	0.0018	75.0000
GREENPTEREGCV	0.0272	65.0000
BLUEPTEREGCV	0.1249	64.0000
COMBPTEREGCV	0.7008	56.0000

Tabla 2 – Características obtenidas analizando el borde de la lesión mediante el ajuste por regresión lineal.

RESULTADOS DE LOS EXPERIMENTOS.

CARACTERÍSTICAS	PARÁMETRO P DEL CONTRASTE DE HIPÓTESIS	PORCENTAJE DE LESIONES CORRECTAMENTE CLASIFICADAS
REDPTESIGMN	0.6858	47.0000
GREENPTESIGMN	0.2160	52.0000
BLUEPTESIGMN	0.3975	49.0000
COMBPTESIGMN	0.0925	62.0000
REDPTESIGSD	0.1536	62.0000
GREENPTESIGSD	0.5627	50.0000
BLUEPTESIGSD	0.8020	54.0000
COMBPTESIGSD	0.3615	51.0000
REDPTESIGQ1	0.4092	52.0000
GREENPTESIGQ1	0.1620	55.0000
BLUEPTESIGQ1	0.2697	48.0000
COMBPTESIGQ1	0.0331	67.0000
REDPTESIGQ2	0.4491	49.0000
GREENPTESIGQ2	0.1829	55.0000
BLUEPTESIGQ2	0.3042	48.0000
COMBPTESIGQ2	0.0940	60.0000
REDPTESIGQ3	0.8923	36.0000
GREENPTESIGQ3	0.2382	52.0000
BLUEPTESIGQ3	0.3848	48.0000
COMBPTESIGQ3	0.1211	57.0000
REDPTESIGRQ	0.3859	57.0000
GREENPTESIGRQ	0.5076	48.0000
BLUEPTESIGRQ	0.6732	51.0000
COMBPTESIGRQ	0.3973	50.0000
REDPTESIGCV	0.0186	67.0000
GREENPTESIGCV	0.6217	53.0000
BLUEPTESIGCV	0.0741	62.0000
COMBPTESIGCV	0.0038	67.0000

Tabla 3 – Características obtenidas analizando el borde de la lesión mediante el ajuste por sigmoidea.

RESULTADOS DE LOS EXPERIMENTOS.

CARACTERÍSTICAS	PARÁMETRO P DEL CONTRASTE DE HIPÓTESIS	PORCENTAJE DE LESIONES CORRECTAMENTE CLASIFICADAS
AREA	0.0138	72.0000
PERIM	0.6349	53.0000
FRAG	0.0117	73.0000
REDMN	0.0001	66.0000
GREENMN	0.0094	65.0000
BLUEMN	0.0160	62.0000
RESD	0.6482	55.0000
GREENSD	0.4016	51.0000
BLUESD	0.1727	56.0000
REDQ1	0.0000	70.0000
GREENQ1	0.0042	67.0000
BLUEQ1	0.0185	63.0000
REDQ2	0.0001	68.0000
GREENQ2	0.0034	67.0000
BLUEQ2	0.0094	64.0000
REDQ3	0.0007	61.0000
GREENQ3	0.0197	61.0000
BLUEQ3	0.0158	65.0000
REDRQ	0.4165	54.0000
GREENRQ	0.2340	54.0000
BLUERQ	0.0720	59.0000
REDCV	0.0010	67.0000
GRENCV	0.0833	52.0000
BLUECV	0.3215	49.0000
REDGRADMN	0.0000	74.0000
GREENGRADMN	0.0000	74.0000
BLUEGRADMN	0.0000	71.0000
REDGRADSD	0.0000	77.0000
GREENGRADSD	0.0000	77.0000
BLUEGRADSD	0.0000	77.0000
REDGRADQ1	0.0000	71.0000
GREENGRADQ1	0.0001	67.0000
BLUEGRADQ1	0.0003	69.0000
REDGRADQ2	0.0000	71.0000
GREENGRADQ2	0.0000	69.0000
BLUEGRADQ2	0.0000	70.0000
REDGRADQ3	0.0000	75.0000
GREENGRADQ3	0.0000	76.0000
BLUEGRADQ3	0.0000	72.0000
REDGRADRQ	0.0000	75.0000
GREENGRADRQ	0.0000	75.0000
BLUEGRADRQ	0.0000	73.0000
REDGRADCV	0.0024	71.0000
GREENGRADCV	0.1899	63.0000
BLUEGRADCV	0.6245	60.0000

Tabla 4 – Características obtenidas analizando los píxeles contenidos en la lesión.

CARACTERÍSTICAS	PARÁMETRO P DEL CONTRASTE DE HIPÓTESIS	PORCENTAJE DE LESIONES CORRECTAMENTE CLASIFICADAS
ENTROPIA_RED	0.0107	70.0000
ENTROPIA_GREEN	0.2620	46.0000
ENTROPIA_BLUE	0.3303	52.0000

Tabla 5 – Características obtenidas analizando todos los píxeles de la imagen de la lesión.

Como se puede apreciar, se obtiene un total de 107 características, de las cuales 44 son características que diferencian significativamente lesiones que son melanomas de las que no son melanomas. Además, dichas características proporcionan un rango entre el 57% y el 77% de lesiones correctamente clasificadas individualmente.

7.3. RESULTADO DE LA CLASIFICACIÓN DE LESIONES.

Como ya se ha mencionado anteriormente, la clasificación de las lesiones se realiza mediante el clasificador discriminante lineal, con una evaluación de lesiones de “100 – fold cross validation” o “100 validaciones cruzadas”, de dos maneras diferentes. La primera empleando todas las características extraídas para todas las lesiones (a excepción de las características correspondientes al rango intercuartílico, debido a que es la combinación lineal de la diferencia del tercer cuartil y el primer cuartil) y la segunda, empleando para cada lesión sólo una selección de características escogidas a través del “forward selection” o “selección hacia delante”. Por lo tanto, al evaluar el clasificador, se determina si la clasificación empleada es adecuada para poder diferenciar melanomas de no melanomas.

7.4. RESULTADO DE LA EVALUACIÓN DEL CLASIFICADOR.

Para evaluar el clasificador discriminante lineal empleando todas las características extraídas para todas las lesiones o aplicando sólo la selección de características correspondiente a cada lesión. Se muestran los valores correspondientes a los conceptos de verdadero/falso positivo/negativo (Tabla 6), para poder calcular la precisión, la sensibilidad y la especificidad del clasificador (Tabla 7) dependiendo del número de características empleadas:

MEDIDAS CUANTITATIVAS	EMPLEANDO TODAS LAS CARACTERÍSTICAS OBTENIDAS	EMPLEANDO LAS CARACTERÍSTICAS SELECCIONADAS
VERDADEROS POSITIVOS	17	23
VERDADEROS NEGATIVOS	46	51
FALSOS POSITIVOS	25	20
FALSOS NEGATIVOS	12	6

Tabla 6 – Verdadero/Falso Positivo/Negativo del clasificador discriminante lineal.

MEDIDAS CUANTITATIVAS	EMPLEANDO TODAS LAS CARACTERÍSTICAS OBTENIDAS	EMPLEANDO LAS CARACTERÍSTICAS SELECCIONADAS
SENSIBILIDAD	58.6207%	79.3103%
ESPECIFICIDAD	64.7887%	71.8310%
PRECISIÓN	63%	74%

Tabla 7 – Sensibilidad, Especificidad y Precisión del clasificador discriminante lineal.

Por tanto, observando ambas tablas (Tabla 6 y Tabla 7) se deduce que aplicando el clasificador discriminante lineal a las lesiones, utilizando las características seleccionadas para cada una de ellas, se obtienen mejores resultados que empleando todas las características a todas las lesiones, puesto que los porcentajes son mayores. Los resultados de la evaluación del clasificador y por tanto de la propia clasificación, son:

- Respecto a la sensibilidad, un 79.31%, es decir, que aproximadamente clasifica correctamente el 79% de las lesiones del tipo melanoma.
- Respecto a la especificidad, un 71.83%, es decir, que clasifica correctamente cerca del 72% de las lesiones de tipo no melanoma.
- Respecto a la precisión, un 74%, es decir, que clasifica correctamente el 74% de las lesiones tanto de un tipo como de otro.

Además cabe destacar que las características más empleadas en la selección de características (entendiendo que se emplean en más de dos lesiones), son las mostradas en la Tabla 8:

CARACTERÍSTICAS	PARÁMETRO P DEL CONTRASTE DE HIPÓTESIS	NÚMERO DE LESIONES EN LAS QUE SE EMPLEAN
REDGRADSD	0.0000	91
BLUESD	0.1727	48
GREENSD	0.4016	43
GREENPTESIGQ2	0.1829	26
GREENPTESIGQ3	0.2382	7
GREENGRADSD	0.0000	6
AREA	0.0138	4
GREENQ3	0.0197	3
GREENCV	0.0833	3
BLUEGRADSD	0.0000	3
REDPTESIGQ1	0.4092	3

Tabla 8 – Características más empleadas en la selección de características.

La Tabla 8 indica que se emplean con frecuencia algunas de las características con parámetro P del contraste de hipótesis, con un valor por debajo de 0.02. Siendo la característica más común la llamada REDGRADSD, que como revela la tabla tiene un valor muy próximo a cero, se emplea en la mayoría de lesiones y significa que diferencia significativamente entre los grupos de melanoma y no melanoma.

Capítulo 8.

CONCLUSIONES Y TRABAJOS

FUTUROS

En este proyecto se ha desarrollado un programa que permite analizar lesiones sospechosas de ser melanoma, empleando técnicas de visión por ordenador. Ya que de esta manera el paciente no necesita una biopsia y obtiene el resultado de forma rápida.

Para cumplir el objetivo principal del proyecto y desarrollar este programa, se han seguido los siguientes pasos:

- Se han acondicionado las 100 imágenes de las lesiones dadas por el Departamento de Dermatología de la Universidad de Missouri – Rolla.
- Se ha realizado la segmentación de cada una de las imágenes mencionadas en el párrafo anterior y se han extraídos de las mismas las características que ayudan a clasificarlas entre melanomas y no melanomas, consiguiendo un conjunto de 107 características.
- Se han clasificado las lesiones de dos maneras diferentes y se ha evaluado el clasificador empleado para cada modo. Razonando del resultado obtenido que utilizando las características seleccionadas para cada lesión, los resultados son bastante favorables, puesto que la sensibilidad alcanza un 79.31%, la especificidad alcanza un 71.83% y la precisión es del 74%.

Además, este proyecto puede servir de base para nuevos proyectos similares. Donde se incluyan nuevas mejoras, como el análisis de los criterios del Color y las Diferencias estructurales de la regla ABCD de la dermatoscopia y que debido a mi nivel de programación no se llevaron a cabo. Estas mejoras se han dejado al margen para limitar el alcance del proyecto, ya que llegaría a ser muy extenso al tratarse de una herramienta tan completa

donde el nivel de programación marca la duración del proyecto y no las limitaciones del programa.

Gracias a este proyecto, he tenido la oportunidad de adquirir conocimientos sobre la herramienta de Matlab y me he dado cuenta de que no sólo es una herramienta de cálculo puesto que también permite analizar imágenes, combinarlas y mediante medidas estadísticas poder clasificar lesiones.

Capítulo 9.

PRESUPUESTO.

En este capítulo se desglosa el presupuesto que supone desarrollar este proyecto titulado: “Reconocimiento automático de melanomas mediante técnicas de visión por ordenador y reconocimiento de patrones”. Abarcando desde la captura y acondicionamiento de la imagen de la lesión, pasando por la digitalización, la segmentación y la obtención de características de la imagen, hasta llegar a la clasificación de la misma. Incluyendo tanto el coste del material como el coste del personal asociado.

9.1. COSTE DEL MATERIAL.

Como coste material de este proyecto, se tienen en cuenta los siguientes conceptos, los cuales están divididos en coste del material en hardware, desglosado en la Tabla 9, y coste del material en software, desglosado en la Tabla 10.

9.1.1. Coste del material en hardware.

CONCEPTO	MODELO	UNIDAD	PRECIO (€)
Dermoscopio digital.	DermLite II Multispectral.	1	765,00
Cámara digital.	Sony Cybershot DSC – W620.	1	155,00
Ordenador portátil.	Portátil 15,6" Pavilion G6 – 1353SS HP.	1	398,00
Almacenamiento.	Memoria USB 8 GB PinStripe VERBATIM.	1	4,99
Ratón periférico USB.	0810 KUNFT.	1	6,49
Impresora.	Impresora de Tinta 1000 HP	1	39,00
Material de oficina.	Tinta, papel, bolígrafos...		100,00
TOTAL DEL COSTE DEL MATERIAL EN HARDWARE:			1.468,48 €

Tabla 9 – Coste del material en hardware.

9.1.2. Coste del material en software.

CONCEPTO	MODELO	UNIDAD	PRECIO (€)
Licencia Matlab.	Matlab 2012 básico.	1	12.150,00
Paquete Microsoft Office 2010.	Office Hogar y Pequeña Empresa 2010.	1	249,00
Sistema Operativo Windows.	Windows 7 Home Premium.	1	199,99
TOTAL DEL COSTE DEL MATERIAL EN SOFTWARE:			12.598,99 €

Tabla 10 – Coste del material en software.

9.2. COSTE DEL PERSONAL.

Como coste del personal del proyecto, se tienen en cuenta los siguientes conceptos, los cuales se dividen en coste del personal en horas de ingeniería y coste del personal en horas de redacción.

9.2.1. Coste del personal en horas de ingeniería.

Las horas de ingeniería, corresponden al tiempo destinado a la programación, el análisis y la verificación del programa. Es decir, las horas empleadas para la realización del programa, las horas destinadas a los análisis ejecutados por el mismo (por ejemplo, el tiempo empleado por el programa para realizar la segmentación o la extracción de características) y las horas

necesarias para la verificación del programa (por ejemplo, verificando la correcta segmentación de las lesiones o el sentido correcto de recogida de datos al analizar el borde de la lesión).

El tiempo empleado para el desarrollo del proyecto ha sido durante aproximadamente 10 meses (200 días a 3h por día de media). Para desglosar el coste personal en horas de ingeniería, se debe tener en cuenta el tiempo empleado en cada fase. A continuación, se aproxima el tiempo empleado en cada fase, aunque no es muy concreto ya que las fases se han ido entremezclando a medida que surgían nuevas ideas. Las principales fases en las que se divide el proyecto son:

- Captura, digitalización y acondicionamiento de la imagen, estableciendo la base de datos del proyecto. Estas fases fueron desarrolladas en aproximadamente un mes.
- Segmentación y extracción de características de las lesiones de la base de datos. Estas fases fueron desarrolladas aproximadamente de cuatro a cinco meses.
- Obtención de los resultados y clasificación de las lesiones. Estas fases fueron desarrolladas durante los últimos meses, aproximadamente de cuatro a cinco meses.

Por tanto, proporcionalmente a lo descrito antes, el coste personal del proyecto respecto a las horas de ingeniería es desglosado en la Tabla 11:

CONCEPTO	TIEMPO APROXIMADO (meses)	UNIDAD (h)	PRECIO UNITARIO (€/h)	PRECIO (€)
Programación	4	240	30	7200
Análisis	3	180	10	1800
Verificación	3	180	10	1800
TOTAL DEL COSTE DEL PERSONAL EN INGENIERÍA:				10.800,00 €

Tabla 11 – Costes del personal en ingeniería.

9.2.2. Coste del personal en horas de redacción.

Las horas empleadas para la redacción de la memoria del proyecto, corresponden al tiempo reservado para la documentación del proyecto y la explicación del programa y de sus resultados. En este caso, el cálculo se considera como un porcentaje del 7% sobre el coste del personal en ingeniería.

Por tanto, según lo mencionado anteriormente, el coste personal del proyecto respecto a las horas de redacción es mostrado a continuación en la Tabla 12:

CONCEPTO	PRECIO (€)
Redacción	756,00
TOTAL DEL COSTE DEL PERSONAL EN REDACCIÓN:	756,00 €

Tabla 12 – Costes del personal en redacción.

9.3. COSTES TOTALES.

Por tanto, se tiene como coste total (de material, tanto hardware como software y de personal, tanto en ingeniería como en redacción), la cifra mostrada en la Tabla 13:

CONCEPTO	COSTES TOTALES (€)
Material hardware	1.468,48
Material software	12.598,99
Personal en ingeniería	10.800,00
Personal en redacción	756,00
TOTAL DE COSTES:	25.623,47 €

Tabla 13 – Costes totales.

A este total de costes, se le debe añadir también el coste del 21% en concepto de impuestos sobre el valor añadido (I.V.A.). Siendo la suma total para el desarrollo del proyecto, el valor indicado en la Tabla 14:

CONCEPTO	COSTES TOTALES (€)
Total sin I.V.A.	25.623,47
I.V.A. (21%)	5.380,93
PRESUPUESTO FINAL:	31.004,40

Tabla 14 – Presupuesto final del proyecto.

El presupuesto total de este proyecto asciende a la cantidad de
TREINTA MIL CUATRO EUROS CON CUARENTA CÉNTIMOS.

Leganés, Diciembre de 2012

El ingeniero proyectista

Fdo. Gertrudis Katerin Asunción Batugo

Capítulo 10.

REFERENCIAS.

- [1] Rebecca Siegel MPH, Deepa Naishadham MA, MS, Ahmedin Jemal DVM, PhD, "Cancer Statistics, 2012," *CA: A Cancer Journal for Clinicians*, vol. 62, pp. 10–29, January/February 2012.
- [2] American Cancer Society. [Online]. <http://www.cancer.org/Espanol/cancer/Cancerdepieltipomelanoma/index>
- [3] W. V. Stoecker, W. W. Li, and R. H. Moss, "Automatic detection of asymmetry in skin tumors," *Computerized Medical Imaging and Graphics*, vol. 16, no. 3, pp. 191 – 197, 1992.
- [4] Cáncer - un especial de mundosalud.com. [Online]. <http://www.elmundo.es/elmundosalud/especiales/cancer/piel.html>
- [5] tuotromedico.com - Melanoma. [Online]. <http://www.tuotromedico.com/temas/melanoma.htm>
- [6] Instituto Nacional del Cáncer de los Institutos Nacionales de la Salud de EE.UU. [Online]. <http://www.cancer.gov/espanol/pdq/tratamiento/melanoma/Patient>
- [7] FamilyDoctor.org información de salud para toda la familia - Cáncer de piel. [Online]. <http://familydoctor.org/familydoctor/es/diseases-conditions/skin-cancer.html>
- [8] Kids Health - Melanoma. [Online]. http://kidshealth.org/parent/en_espanol/medicos/melanoma_esp.html
- [9] R. J. Friedman, D. S. Rigel, and A. W. Kopf, "Early detection of malignant melanoma: the role of physician examination and self-examination of the skin," *CA – A Cancer Journal for Clinicians*, vol. 35, no. 3, pp. 130 – 151, May/June 1985.
- [10] Wikipedia - The free encyclopedia. [Online]. http://en.wikipedia.org/wiki/Main_Page
- [11] H. Ganster, A. Pinz, R. Röhler, E. Wildling, M. Binder, and H.Kittler, "Automated melanoma recognition," *IEEE Transactions on Medical Imaging*, vol. 20, no. 3, pp. 233 – 239, March 2001.
- [12] Yahoo México - Respuestas. [Online]. <http://mx.answers.yahoo.com>
- [13] Dibujos Fotos e Imagenes. [Online].

- <http://dibujosfotoseimagenes.blogspot.com.es/2010/05/microscopio.html>
- [14] W. Stolz, A. Riemann, A. B. Cognetta, L. Pillet, W. Abmayr, D. Hölzel, P. Bilek, F. Nachbar, M. Landthaler, and O. Braun – Falco, "ABCD rule of dermatoscopy: a new practical method for early recognition of malignant melanoma," *European Journal Dermatology*, vol. 4, pp. 521 – 527, 1994.
- [15] Manejo de la lupa binocular o estereomicroscopio. [Online]. <http://fresno.pntic.mec.es/~rruv0000/paginas/binocular.htm>
- [16] Eximpor Instrumental Científico - Estereomicroscopio con zoom. [Online]. <http://eximpor.com/productos.php?categoria=Estereomicroscopio%20con%20Zoom>
- [17] Entremujeres.com - Dermatoscopia. [Online]. http://www.entremujeres.com/belleza/dermatologia-piel-cuidado-cancer-lunares-manchas_0_301169897.html
- [18] Librerías aula médica - Dermoscopio DermLite II Hybrid M. [Online]. <http://www.libreriasaulamedica.com/ficha.asp?id=41300&fam=3&subcat=Dermoscopios&T=Dermoscopio+DermLite%AE+II+Hybrid+M>
- [19] Heine - DERMATOSCOPIO HEINE MINI3000®. [Online]. http://heine.com/esl_US/PRODUCTOS/CUADRO-DE-PRODUCTOS/Dermoscopios/Dermoscopio-HEINE-mini3000-R
- [20] S. Seidenari, M. Burrioni, G. Dell’Eva, P. Pepe, and B. Belletti, "Computerized evaluation of pigmented skin lesion images recorded by a videomicroscope: Comparison between polarizing mode observation and oil/slide mode observation," *Skin Research and Technology*, vol. 1, pp. 187 – 191, 1995.
- [21] Veo, veo. el mundo invisible. [Online]. http://www.madrimasd.org/Madridporlaciencia/feria_I/pdf/082_083.pdf
- [22] M.A.D. Apparecchiature Scientifiche ITALY. [Online]. http://www.edumad.it/es/index.php?option=com_virtuemart&page=shop.browse&category_id=77&Itemid=4&lang=es
- [23] Allbiz. [Online]. <http://www.ua.all.biz/es/g1594022/>
- [24] YouTube. [Online]. <http://www.youtube.com/>
- [25] MECAMED Equipamiento Hospitalario
- [26] Constantino Grana, Giovanni Pellacani, Rita Cucchiara and Stefania Seidenari, "A new algorithm for border description of polarized light surface microscopic images of pigmented skin lesions," *IEEE Transactions on Medical Imaging*, vol. 22, no. 8, pp. 959 - 964, August 2003.
- [27] D. Delgado, C. Butakoff, B. K. Ersbøll, and W. Stoecker, "Independent histogram pursuit for segmentation of skin lesions," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 1, pp. 157 – 161, January 2008.
- [28] M. E. Celebi, H. Iyatomi, G. schaefer, and W. V. Stoecker, "Lesion border detection in dermoscopy images," *Computerized Medical Imaging and Graphics*, vol. 33, pp. 148 – 153, 2009.
- [29] M. Emre Celebi, Gerald Schaefer, Hitoshi Iyatomi, William V. Stoecker, Joseph M. Malters, and James M. Grichnik, "An improved objective evaluation measure for border detection in dermoscopy images," *Skin Research and Technology*, vol. 15, pp. 444 - 450, November 2009.
- [30] I. Maglogiannis, S. Pavlopoulos, and D. Koutsouris, "An integrated computer supported acquisition, handling, and characterization system for pigmented skin lesions in

- dermatological images," *IEEE Transactions on Information Technology in Biomedicine*, vol. 9, no. 1, pp. 86 – 98, March 2005.
- [31] Nobuyuki Otsu, "A threshold selection method from gray - level histograms," *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC - 9, no. 1, pp. 62 - 66, January 1979.
- [32] Hiroshi Oka, M.Emre Celebi, Masahiro Hashimoto, Masafumi Hagiwara, Masaru Tanaka, Koichi Ogawa Hitoshi Iyatomi, "An improved Internet-based melanoma screening system with dermatologist - like tumor area extraction algorithm," *Computerized Medical Imaging and Graphics*, vol. 32, pp. 566 – 579, 2008).
- [33] Bulent Erkol, Randy H. Moss, R. Joe Stanley, William V. Stoecker, and Erik Hvatum, "Automatic lesion boundary detection in dermoscopy images using gradient vector flow snakes," *Skin Research and Technology*, vol. 11, pp. 17 - 26, February 2005.
- [34] Hanzheng Wang, Xiaohe Chen, Randy H. Moss, R. Joe Stanley, William V. Stoecker, M. Emre Celebi, Thomas M. Szalapski, Joseph M. Malter, James M. Grichnik, Ashfaq A. Marghoob, Harold S. Rabinovitz, and Scott W. Menzies, "Watershed segmentation of dermoscopy images using a watershed technique," *Skin Research and Technology*, vol. 16, pp. 378 - 384 , August 2010.
- [35] A. Safi, M. Baust, O. Pauly, V. Castaneda, T. Lasser, D. Mateus, N. Navab, R. Hein, and M. Ziai, "Computer – Aided diagnosis of pigmented skin dermoscopic images," *MCBR-CDS'11 Medical Content - Based Retrieval for Clinical Decision Support*, pp. 105 – 115, 2012.
- [36] F. Nachbar, W. Stolz, T. Merkle, A. B. Cagnetta, T. Vogt, M. Landthaler, P. Bilek, O. Braun – Falco, and G. Plewig, "The ABCD rule of dermatoscopy: High prospective value in the diagnosis of doubtful melanocytic skin lesions," *Journal of the American Academy of Dermatology*, vol. 30, no. 4, pp. 551 – 559, April 1994.
- [37] M. Binder, H. Kittler, A. Steiner, M. Dawid, H. Pehamberger, and K. Wolff, "Reevaluation of the ABCD rule for epiluminescence microscopy," *Journal of the American Academy of Dermatology*, vol. 40, no. 2, pp. 171 – 176, 1999.
- [38] G. Argenziano, G. Fabbrocini, P. Carli, V. De Giorgi, E. Sammarco, and M. Delfino, "Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions," *Archives of Dermatology*, vol. 134, pp. 1563 – 1570, December 1998.
- [39] G.R. Day, "How blurry is that border? An investigation into algorithmic reproduction of skin lesion border cut - off," *Computerized Medical Imaging and Graphics*, vol. 24, pp. 69 - 72, 2000.
- [40] A. Green, N. Martin, J. Pfitzner, M. O'Rourke, and N. Knight, "Computer image analysis in the diagnosis of melanoma," *Journal of the American Academy of Dermatology*, vol. 31, no. 6, pp. 958 – 964, December 1994.
- [41] N. Cascinelli, M. Ferrario, R. Bufalino, S. Zurrada, V. Galimberti, L. Mascheroni, C. Bartoli, and C. Clemente, "Results obtained by using a computerized image analysis system designed as an aid to diagnosis of cutaneous melanoma," *Melanoma Research*, vol. 2, pp. 163 – 170 , 1992.
- [42] Y. I. Cheng, R. Swamisai, S. E. Umbaugh, R. H. Moss, W. V. Stoecker, S. Teegala, and S. K. Srinivasan, "Skin lesion classification using relative color features," *Skin Research and Technology*, vol. 14, pp. 53 – 64, 2008.
- [43] SlideShare - Present Yourself. [Online]. <http://www.slideshare.net/latiatuca/enostosis-1954495#btnPrevious>
- [44] Corina Cuevas Renaud, Amalia Alejo Martínez. (2010, Octubre) VALIDEZ Y FIABILIDAD DE

LAS MEDIDAS DE EXPOSICIÓN Y MEDICIÓN. Documento.pdf. [Online].
<http://www.psicol.unam.mx/Investigacion2/pdf/SENSIBILIDAD%20Y%20ESPECIFICIDAD.pdf>