



UNIVERSIDAD CATÓLICA
SAN ANTONIO
UCAM

ESCUELA INTERNACIONAL DE DOCTORADO

Programa de Doctorado en Ciencias de la Salud

“Clasificación automática de tumores de ovario y
citologías cervicovaginales a partir de imágenes
ecográficas y microscópicas mediante su análisis con
técnicas de aprendizaje automático ”

Autor: D. José Martínez Más

Directores:

Dr. D. Andrés Bueno Crespo

Dr. D. Juan Pedro Martínez Cendán

Dr. D. Manuel Remezal Solano

Murcia, 30 de Marzo de 2020



UNIVERSIDAD CATÓLICA
SAN ANTONIO
UCAM

AUTORIZACIÓN DE LO/S DIRECTOR/ES DE LA TESIS PARA SU PRESENTACIÓN

.....

El Dr. Andrés Bueno Crespo, el Dr. Juan Pedro Martínez Cendán y el Dr. Manuel Remezal Solano, como Directores de la Tesis Doctoral titulada “Clasificación automática de tumores de ovario y citologías cervicovaginales a partir de imágenes ecográficas y microscópicas mediante su análisis con técnicas de aprendizaje automático” realizada por Don José Martínez Más, autorizan su presentación a trámite dado que reúne las condiciones necesarias para su defensa. Lo que firmamos, para dar cumplimiento a los Reales Decretos 99/2011, 1393/2007, 56/2005 y 778/98,

en Murcia a 31 de marzo de 2020.

Dr. A. Bueno Crespo | Dr. M. Remezal Solano | Dr. J. P. Martínez Cendán

RESUMEN

RESUMEN

OBJETIVOS

Aplicar técnicas de Aprendizaje Automático para la clasificación diagnóstica de imágenes ecográficas de tumores de ovario e imágenes microscópicas de citologías cervicovaginales teñidas con la técnica de Papanicolaou.

Aplicar técnicas de Aprendizaje Automático para la clasificación de descriptores geométricos obtenidos de imágenes ecográficas de tumores ováricos mediante *Fast Fourier Transform* y comparar con los resultados obtenidos con otros trabajos que utilizan la misma base de datos.

Crear una base de datos mediante fusión de imágenes para mezclar y superponer las células originales y aumentando el número de imágenes disponibles para clasificación.

Aplicar Deep Learning para la clasificación de las imágenes microscópicas de la base de datos de nueva creación y comparar los resultados con los obtenidos en la clasificación de la base de datos original.

MÉTODO

Se realizará el análisis de resultados de la aplicación de clasificadores basados en Aprendizaje Automático clásicos sobre una base de datos de descriptores extraídos mediante Transformada de Fourier a partir de una colección de 187 imágenes ecográficas de tumores de ovario, 112 benignos y 75 malignos, cedidos por la Universidad de Buckingham, que a su vez obtuvo las imágenes originales de la Universidad Católica de Leuven. Las características que se clasifican son Histogramas de Intensidad y Descriptores de Patrón Binario Local.

También se realizará la clasificación de imágenes microscópicas de células escamosas cervicovaginales teñidas mediante la técnica de Papanicolaou aplicando una red neuronal convolucional. La muestra es creada a partir de 10 citologías cervicovaginales procedentes del Servicio de Anatomía Patológica del

Complejo Hospitalario Universitario Santa Lucía-Santa María del Rosell (Cartagena, Murcia), de donde se extrajeron 1405 células. 450 normales, 323 ASC-US, 213 L-SIL, 419 H-SIL. Además, a partir de estas células se desarrolla un sistema de fusión de imágenes para aumentar el número de imágenes de la muestra, obteniendo 20.000 imágenes por cada categoría, 80.000 en total.

RESULTADOS

La clasificación de características descriptivas de imágenes ecográficas de tumores de ovario ofrece resultados similares a los obtenidos por un observador experimentado cuando se aplica los métodos *Linear Discriminant*, *Support Vector Machine* y *Extreme Learning Machine*.

La clasificación de células escamosas cervicales mediante *Deep Learning* ofrece resultados interesantes, que mejoran al aumentar el tamaño muestral de entrenamiento aunque estas imágenes sean más complejas por la fusión, obteniendo resultados comparables a los obtenidos por patólogos expertos.

CONCLUSIONES

Los métodos basados en Inteligencia Artificial pueden tener utilidad para el diseño de sistemas de ayuda al diagnóstico médico asistidos por ordenador, para la clasificación de imágenes ecográficas de tumores de ovario, así como para la detección de células escamosas cervicales atípicas procedentes de citologías cervicovaginales teñidas mediante la técnica de Papanicolaou.

Los descriptores Geométricos obtenidos mediante *Fast Fourier Transform* aportan información útil y relevante para la clasificación de ecografías de tumores de ovario, mejorando los resultados obtenidos en nuestro estudio en comparación con otros estudios anteriores realizados sobre la misma base de datos.

Se ha generado una base de datos mediante fusión de imágenes con transparencia a partir de la base de datos original de células escamosas cervicovaginales, obteniendo una colección de 80.000 imágenes de nueva creación, con mayor

complejidad.

Al aplicar las técnicas de *Deep Learning* sobre la base de datos de nueva creación, se comprueba que aumentando el número de imágenes de muestra, independientemente de su complejidad, se mejora los resultados y la estabilidad de la clasificación.

PALABRAS CLAVE: Aprendizaje automático, ecografía, tumor ovárico, Papanicolaou, Deep Learning, SVM, ELM, LD, KNN.

SUMMARY

SUMMARY

AIM

To apply Machine Learning techniques for classification of ovarian tumor ultrasound scan images and microscopic images taken from Papanicolau smears.

To apply Machine Learning techniques for classification of Fast Fourier Geometric Features obtained from ovarian tumor ultrasound scan images, and compare the results with other publications made with the same sample dataset.

To create a new dataset by a cell merging approach to increase the number and complexity of the images sample.

To apply Deep Learning techniques to classify microscopic images from the new created dataset and compare the results with the previous experiment over the original images dataset.

METHOD

A feature collection extracted from 187 ultrasound scan images obtained from ovarian tumors will be classified by Machine Learning classic techniques. In this collection, 112 images are benign and 75 are malignant tumors and it was lent by Buckingham University, who borrowed the original images from Catholic University of Leuven. Extracted features to be classified are Intensity Histograms and Local Binary Pattern features.

A classification of microscopic images of squamous cervical cells stained by Papanicolau's method will be performed using a convolutional neural network. The images collection is created from 10 cervical smears from the Pathology Service of Complejo Hospitalario Universitario Santa Lucía-Santa María del Rosell (Cartagena, Murcia), obtaining 1405 cells: 450 normal, 323 ASC-US, 213 L-SIL and 419 H-SIL. A cell merging approach is created to increase the number and complexity of the sample, obtaining 20,000 images for each category, 80,000 in total.

RESULTS

The classification of extracted features from ultrasound scan images from ovarian tumors give similar results to those obtained by an experienced human observer when Linear Discriminant, Support Vector Machine and Extreme Learning Machine are applied.

The classification of squamous cervical cells by Deep Learning techniques give interesting results, that improve when the cell merging approach is classified although its higher complexity, reaching similar results to those obtained by experienced pathologists.

CONCLUSIONS

Machine Learning techniques could be effectively applied to diagnostic images classification at Gynecology field, specially to classify ultrasound scan images from ovarian tumors and microscopic images from squamous cervical cells stained by Papanicolaou's technique.

Fast Fourier Geometric Features provide useful and relevant information to classify ovarian tumor ultrasound-scan images, improving the obtained results from other studies performed over the same dataset.

A new images dataset has been created by a cell merging approach, from the original squamous cervical cells, getting a set of 80,000 new images, with higher complexity.

It is suggested that increasing the number of sample images using the new created dataset to apply over it Deep Learning techniques, improves the results and stability of the performance, independently of its higher complexity.

KEYWORDS: Machine learning, Ultrasound scan, ovarian tumor, Papanicolaou, Deep Learning, SVM, ELM, LD, KNN.

*A Marisol, mi amor
a nuestro pequeño José María
y a mi padre y mi tío, que nos cuidan desde el cielo*

AGRADECIMIENTOS

Al Dr. Pedro García Laencina, por ser el impulsor inicial de este trabajo.

Al Dr. Dirk Timmerman, cuya disposición a colaborar con nuestro grupo ha sido fundamental para realizar el análisis de clasificación de imágenes ecográficas, por haber sido el mediador con otras Universidades y ayudarnos a continuar con nuestro trabajo.

A la Dra. Shan Khazendar, cuya colaboración con nuestro equipo supuso una aventura desde su regreso desde Londres al Kurdistán Iraquí, y cuyos trabajos sirvieron de punto de partida inicial al nuestro. Siempre dispuesta a ayudar.

A la Dra. Marisa Sánchez Ferrer y al Dr. José Ramón Rodríguez Hernández, por su apoyo y consejos durante el proceso de publicación de los artículos.

Al Dr. Sebastián Ortiz Reina y todo el Servicio de Anatomía Patológica del Hospital Universitario Santa María del Rosell, por su apoyo y ayuda para la creación y estudio de la base de datos de citologías cervicovaginales.

A mi tutora de la Escuela Internacional de Doctorado, la Dra. Débora Villaño Valencia, cuyo trabajo y apoyo han sido fundamentales para solventar las dificultades burocráticas y de gestión con la Universidad.

A mis Directores de Tesis, el Dr. Andrés Bueno Crespo, el Dr. Manuel Remezal Solano y el Dr. Juan Pedro Martínez Cendan, quienes han sido mi guía para descubrir la Inteligencia Artificial, mis maestros durante la residencia de ginecología, compañeros en el trabajo diario y sobre todo, amigos en los buenos

y en los malos momentos que hemos pasado.

A mi padre y mi tío, que fueron siempre mi ejemplo de superación y constancia para conseguir mis objetivos, sin quienes hoy yo no sería quien soy.

A Marisol, mi mujer, por apoyarme siempre en mis proyectos, siendo la luz de nuestro hogar.

A José María, nuestro pequeño, por esperarme cada noche cuando llego a casa con la ilusión de verme y jugar juntos. Te prometo que, a partir de ahora, jugaremos aún más.

*“No cuentes los días:
haz que los días cuenten”.*

Muhammad Ali

ÍNDICE GENERAL

Índice general

Lista de abreviaturas	33
Índice de figuras	39
Índice de tablas	45
1. INTRODUCCIÓN	55
1.1. Anatomía e histología del aparato genital femenino: útero, ovarios y cuello uterino	55
1.2. Cáncer de ovario	65
1.2.1. Incidencia y mortalidad mundial del Cáncer de Ovario . .	66
1.2.2. Incidencia y mortalidad en España por Cáncer de Ovario .	67
1.2.3. Etiopatogenia del Cáncer de Ovario	68
1.2.4. Tipos de Cáncer de Ovario	75
1.2.5. Manejo clínico de los tumores ováricos	76
1.3. Cáncer de Cérvix	89
1.3.1. Incidencia y mortalidad mundial del Cáncer de Cérvix . .	89
1.3.2. Incidencia y mortalidad en España del Cáncer de Cérvix .	91
1.3.3. Antecedentes de la detección precoz del Cáncer de Cérvix	92
1.3.4. Lesiones precursoras del Cáncer de Cérvix y su detección	94
1.3.5. Análisis automático de citologías cervicovaginales	104

1.4.	Inicios de la Inteligencia Artificial	106
1.5.	Aprendizaje Automático aplicado a la clasificación de tumores ováricos.	108
1.6.	<i>Deep Learning</i> y Redes Neuronales Convolucionales	110
2.	HIPÓTESIS	121
3.	OBJETIVOS	125
3.1.	Objetivo principal	125
3.2.	Objetivos secundarios	125
4.	MATERIAL Y MÉTODOS	129
4.1.	Bases de datos utilizadas	129
4.1.1.	Base de datos de ecografías de tumores ováricos	129
4.1.2.	Bases para la extracción de características de FFT:	134
4.1.3.	Base de datos de imágenes microscópicas de células esca- mosas cervicales.	137
4.1.4.	Base de datos de nueva creación mediante la aplicación de técnicas de fusión de imágenes sobre la base de datos original de células escamosas	142
4.2.	Configuración de los experimentos.	146
4.2.1.	Configuración de los experimentos para la clasificación de descriptores obtenidos de imágenes ecográficas de tumo- res de ovario.	146
4.2.2.	Evaluación de resultados y protocolo de entrenamiento y prueba para la base de datos de tumores ováricos.	152
4.2.3.	Configuración de la arquitectura de CNN para la clasifi- cación de imágenes de células escamosas a partir de cito- logías cervicovaginales mediante Deep Learning.	154

4.2.4.	Configuración de los experimentos de clasificación de imágenes microscópicas de células escamosas obtenidas a partir de citologías cervicovaginales tras la creación de la base de datos por fusión de imágenes, mediante Deep Learning.	156
4.3.	Descripción del Hardware y Software utilizados	157
5.	RESULTADOS	161
5.1.	Resultados de la clasificación automática de imágenes ecográficas de tumores de ovario	161
5.2.	Resultados de la Clasificación de Células Escamosas Cervicales con tinción de Papanicolaou.	165
5.2.1.	Resultados de la Clasificación de Células Originales.	165
5.2.2.	Resultados de la Clasificación de Células Fusionadas	168
6.	DISCUSIÓN	175
6.1.	Clasificación de tumores de ovario con técnicas clásicas de ML	175
6.2.	Clasificación de Citologías mediante DL	180
6.2.1.	Clasificación de Células Originales	182
6.2.2.	Clasificación de Células Fusionadas	186
7.	CONCLUSIONES	193
8.	LIMITACIONES DEL ESTUDIO	197
9.	ANEXO	201
	Bibliografía	207

LISTA DE ABREVIATURAS

Lista de abreviaturas

Abreviatura	Término
ACC	Accuracy, exactitud
ADNEX	<i>Assessment of Different NEoplasias in the adneXa</i> , sistema de ayuda al diagnóstico médico de tumoraciones ováricas desarrollado por el grupo IOTA
AIN	<i>Anal Intraepitaelial Neoplasia</i>
ANNs	<i>Artificial Neural Networks</i> , redes neuronales artificiales
ARID1A	<i>AT-Rich Interactive Domain-containing protein 1A</i> dominio de interacción rico en AT de la proteína A1. Esta proteína tiene función helicasa y ATPasa y regula la transcripción de genes mediante la alteración de la estructura de la cromatina
ASC-US	Atipias Celulares Escasmosas de Significado Incierto
AUC	<i>Area Under Curve</i> . Medida estadística del rendimiento diagnóstico de una prueba
BRAF	Gen que codifica la proteína B-RAF, que participa en la señalización celular para la multiplicación de estas. Su mutación se relaciona con aumento en la multiplicación y la diseminación de las células cancerosas
BRCA1	Gen localizado en el cromosoma 17 que ayuda a suprimir la división de las células. Su mutación se relaciona con un riesgo más alto de cáncer de mama, ovario, próstata y otros tipos de cáncer
BRCA2	Gen localizado en el cromosoma 13 que ayuda a suprimir la división de las células. Su mutación se relaciona con un riesgo más alto de cáncer de mama, ovario, próstata y de otros tipos de cáncer

Abreviatura	Término
Ca125	Glucoproteína utilizada como marcador tumoral presente en patología ovárica benigna y maligna, pero no exclusivamente
CIN I	Neoplasia Intraepitelial Cervical grado I
CIN II	Neoplasia Intraepitelial Cervical grado II
CIN III	Neoplasia Intraepitelial Cervical grado III
CMA	Cell Merging Approach, método de fusión de imágenes celulares
CNN	Redes Neuronales Convolucionales
CTNNB1	Gen localizado en el cromosoma 22, que codifica la proteína β -Catenina, que forma parte del citoesqueleto celular. Su mutación se relaciona con ciertos tipos de cáncer, como el intestinal, gástrico, glioblastoma, y otros síndromes hereditarios
DFT	Discrete Fourier Transform, Transformada de Fourier Discreta
DL	Deep Learning
ELM	Extreme Learning Machine
FFGF	<i>Fast Fourier Geometrical Feature</i> , Características descriptivas geométricas extraídas mediante procesamiento con Transformada de Fourier
FFT	<i>Fast Fourier Transform</i> , Transformada Rápida de Fourier
FIGO	Federación Internacional de Ginecología y Obstetricia
FN	Falsos Negativos
FP	Falsos Positivos
FT	Fourier Transform, Transformada de Fourier
GPU	<i>Graphic Proccesing Unit</i> , Unidad de procesamiento gráfico
H-SIL	Lesión intraepitelial escamosa de alto grado, es el hallazgo citológico concordante con un CIN II o III histológico
IA	Inteligencia Artificial

Abreviatura	Término
IC	Intervalo de confianza
IOTA	<i>International Ovarian Tumor Analysis</i> , grupo internacional dedicado al estudio de la estandarización de las exploraciones ecográficas de tumores ováricos, así como al desarrollo de herramientas de asistencia al diagnóstico médico
KNN	K-Nearest Neighbours
LARS	Least Angle Regression System, Sistema de regresión del mínimo ángulo
LBC	Liquid-Based Cytology, es el sistema de toma de citologías en medio líquido, que requiere realizar el muestreo con un cepillo especial de recogida de células, el LBC Specimen brush
LD	Linear Discriminant
LOO-CV	Leave One-Out Cross Validation, técnica de entrenamiento y test que consiste en utilizar todos los elementos de una base de datos excepto uno para el entrenamiento del sistema de clasificación, y utilizar el elemento restante para el test. El procedimiento se repite una vez por cada elemento del conjunto de datos existente, dejando fuera del entrenamiento cada vez un elemento distinto
LR	<i>Linear Regression</i> , regresión lineal
L-SIL	Lesión intraepitelial escamosa de bajo grado, es el hallazgo citológico concordante con un CIN I histológico
ML	Machine Learning o Aprendizaje Máquina
MLP	<i>MultiLayer Perceptron</i> , perceptrón multicapa
OCE	Orificio Cervical Externo
OCI	Orificio Cervical Interno

Abreviatura	Término
PO-ELM	<i>Optimally-Pruned Extreme Learning Machine</i> , método optimizado de poda para Extreme Learning Machine
p53	Protooncogén también conocido como el “guardián del genoma” Localizado en el cromosoma 17. Es un gen supresor tumoral que codifica un factor de transcripción nuclear que controla el ciclo celular e induce apoptosis en caso de detectar anomalías genéticas en la célula
RMI	Risk of Malignancy Index. Fórmula matemática para calcular el riesgo de que una tumoración ovárica sea maligna o benigna. Desarrollada por Jacobs y cols. (1990)
RMN	Resonancia Magnética Nuclear
ROC	<i>Receiver Operating Characteristic</i> , representación gráfica de la sensibilidad de una prueba diagnóstica
SD	Desviación estándar
SEN	Sensibilidad
SPE	Especificidad
SVM	Support Vector Machine
VAIN	<i>Vaginal Intraepitaelial Neoplasia</i>
VIN	<i>Vulvar Intraepitaelial Neoplasia</i>
VN	Verdaderos Negativos
VP	Verdaderos Positivos
VPH	Virus del Papiloma Humano

ÍNDICE DE FIGURAS

Índice de figuras

1.1. Ilustración en detalle del ovario, irrigación y relaciones, tomada de Franz Netter [1].	57
1.2. Ilustración de la pelvis femenina, tomada de Franz Netter [1]. . .	61
1.3. Imagen de un espéculo vaginal desechable.	62
1.4. Sala de exploraciones ginecológicas, con ecógrafo de altas prestaciones.	76
1.5. Ejemplos de imágenes ecográficas de tumoraciones ováricas procedentes de la base de datos creada por el Servicio de Ginecología del Hospital KU de Leuven, Bélgica. (a) Tumoración ovárica benigna; (b) Tumoración ovárica maligna.	79
1.6. Tumoración ovárica benigna, de difícil caracterización, procedente de la base de datos creada por el Servicio de Ginecología del Hospital KU de Leuven, Bélgica.	80
1.7. Esquema del desarrollo de las lesiones cervicales, de izquierda a derecha, desde epitelio normal hasta cáncer invasivo, pasando por CIN I, II y III. tomada de Carrasco Colomer (2008) [2].	98
1.8. Material utilizado habitualmente para la toma de muestras cervicovaginales para realizar citologías (a) Espátula de Ayre, (b) cepillo citológico LBC, y (c) lavado del cepillo en líquido preservativo BD Surepath®.	100

-
- 4.1. Ejemplos de imágenes ecográficas procedentes de la base de datos original creada por el Servicio de Ginecología de la Universidad KU de Leuven (Bélgica), utilizada para la extracción de descriptores: (a) Tumoración ovárica benigna y (b) Tumoración ovárica maligna. 131

 - 4.2. Comparación del espectro de las imágenes ecográficas de tumores ováricos benignos y malignos, tomada de Martínez Más y cols (2019). [3]. 136

 - 4.3. Ejemplos procedentes de la base de datos utilizada, obtenida tras la fragmentación de imágenes procedentes del archivo de imágenes del Servicio de Anatomía Patológica el Complejo Hospitalario de Cartagena Santa Lucía-Santa María del Rosell (Cartagena, Murcia): (a) Célula escamosa normal (b) ASC-US (c) L-SIL y (d) H-SIL. 139

 - 4.4. Distintos ejemplos donde se muestra la complejidad para la clasificación de las células debido a superposicionamientos, plegamientos, borrosidad por defectos de enfoque y artefactos de tinción en (a) Célula normal, (b) ASC-US, (c) L-SIL y (d) H-SIL. Imágenes procedentes de la base de datos utilizada, obtenida tras la fragmentación de imágenes procedentes del archivo de imágenes del Servicio de Anatomía Patológica el Complejo Hospitalario de Cartagena Santa Lucía-Santa María del Rosell (Cartagena, Murcia) 141

4.5.	Ejemplos de distintas células procedentes de la base de datos del estudio, sin procesamiento de imagen: (a) Célula escamosa benigna, (b) células glandulares, (c) leucocitos polimorfonucleares, (d) células aglomeradas de distintas estirpes, o (e) célula escamosa borrosa, por defecto de enfoque. Imágenes procedentes de la base de datos utilizada, obtenida tras la fragmentación de imágenes procedentes del archivo de imágenes del Servicio de Anatomía Patológica el Complejo Hospitalario de Cartagena Santa Lucía-Santa María del Rosell (Cartagena, Murcia).	143
4.6.	Descripción gráfica esquematizada del proceso de fusión de imágenes procedentes de la base de datos utilizada, obtenida tras la fragmentación de imágenes procedentes del archivo de imágenes del Servicio de Anatomía Patológica el Complejo Hospitalario de Cartagena Santa Lucía-Santa María del Rosell (Cartagena, Murcia). Figura tomada de Ortiz-Reina (2019) [4].	145
4.7.	El método de clasificación realiza una medición de ACC en dos niveles: Una clasificación en 4 grupos diagnósticos (Normal, ASC-US, L-SIL y H-SIL) y otra clasificación respecto a si requiere o no una revisión por el patólogo. Figura tomada de Ortiz-Reina y cols. (2019) [4].	155
5.1.	Los aciertos del modelo CNN en la clasificación binaria de “No requiere revisión” y “Requiere revisión” se muestran separados para cada categoría celular utilizando las imágenes originales como entrada. Figura tomada de Ortiz-Reina y cols. (2019) [4]. . . .	166

- 5.2. Presentación de los aciertos de las células clasificadas dentro de las categorías “Requiere revisión” y “No requiere revisión”, en relación al tipo de célula clasificada, utilizando la base de datos creada por CMA como entrada. 171

ÍNDICE DE TABLAS

Índice de tablas

5.1. Resultados Obtenidos (Exactitud (en %) -ACC-, Área bajo la curva ROC -AUC-, Sensibilidad (en %) -SEN- y Especificidad (en %) -SPE-) por los clasificadores KNN, LD, SVM y ELM utilizando Características Geométricas obtenidas por FFT. Evaluación realizada mediante procedimiento LOO-CV.	162
5.2. Resultados obtenidos (Exactitud -ACC-, Área bajo la curva ROC -AUC- y neuronas ocultas -HN-) por el clasificador ELM de distintos Kernels (Media y Desviación estándar) utilizando Características Geométricas obtenidas por FFT. Evaluación realizada mediante procedimiento LOO-CV.	164
5.3. ACC media, SD, SEN y SPE de la clasificación de la base de datos imágenes originales, en “No requiere revisión” y “Requiere revisión”.	165
5.4. Media y SD de ACC en % para clasificación multiclase en categorías individuales de las imágenes originales en H-SIL, L-SIL, ASC-US y Normal.	167
5.5. Matriz de confusión para la clasificación multiclase de categorías individuales (H-SIL, L-SIL, ASC-US y Normal) utilizando las imágenes originales como entrada, resaltando en negrita las clasificaciones correctas en cada categoría.	167

-
- 5.6. Media y SD de la ACC, sensibilidad y especificidad para la muestra generada mediante CMA. 168
 - 5.7. Media y SD de la ACC en la clasificación multiclase de la muestra generada mediante CMA en las categorías individuales H-SIL, L-SIL, ASC-US y Normal. 169
 - 5.8. Los aciertos del modelo de CNN para la clasificación binaria en “No requiere revisión” y “Requiere revisión” se muestran separados para cada categoría celular individual utilizando como entrada la muestra obtenida mediante CMA. 170

PREÁMBULO

PREÁMBULO

Como tantos proyectos importantes en la vida, la idea de realizar este trabajo surgió cuando, por casualidad, al poco de comenzar mi Residencia en Ginecología y Obstetricia, conocí a un ingeniero muy creativo, entusiasta y con una capacidad de trabajo sorprendente, docente e investigador, con quien estuve comentando su Trabajo de Fin de Carrera, cuyo objetivo era detectar microcalcificaciones en mamografías digitalizadas, en el contexto de programas de cribado de cáncer de mama, con una metodología y unos resultados muy interesantes. Este hombre, el Dr. Pedro García Laencina, fue el responsable de que nuestro proyecto comenzase a dar sus primeros pasos, que me llevaron a contactar con el Dr. Dirk Timmerman, pues figuraba como autor en la gran mayoría de trabajos encaminados a la caracterización de imágenes ecográficas de ovario que encontramos en la bibliografía, quien me citó para una reunión en la cafetería de un hotel de Granada, pues era ponente invitado en un Curso Internacional de Ecografía, donde le planteé nuestro proyecto inicial, y accedió a colaborar con nosotros, poniéndonos en contacto con la Universidad de Buckingham, que a su vez colaboraba con él en un proyecto similar al nuestro y le pareció interesante que pudiéramos trabajar ambos grupos sobre la misma base de datos, de forma independiente y paralela.

Cuando nos fue cedida la base de datos, nos planteamos que sería intere-

sante contar en la dirección del Proyecto con otro ingeniero informático y un patólogo, momento en el que se le propuso tanto al Dr. Andrés Bueno Crespo como al Dr. Manuel Remezal Solano formar parte como co-directores de la Tesis. Así, comenzamos nuestra aventura, quedando en el Bar Príncipe (en Cartagena) como punto habitual de encuentro y definiendo, entre cafés solos y asiáticos, el germen de nuestro trabajo, garabateado en servilletas de bar, que guardo como reliquias del inicio de este camino. Pero, debido a las dificultades que surgieron en el proceso, en parte causadas por la colaboración entre varias Universidades y hospitales, con diferentes puntos de vista e intereses en distintos proyectos, el nuestro se fue frenando tras obtener los primeros resultados, retrasando así la publicación de los mismos.

En este momento, sufrimos el fallecimiento repentino de Pedro, que era nuestro mayor impulsor, viéndonos entonces frenados tanto de forma académica por los impedimentos comentados como de forma anímica, siendo a la vez, el momento de formalizar el proyecto en la Universidad Católica de Murcia. Dado que nos quedó vacante un puesto en la Dirección del proyecto, decidimos invitar a participar en él al Dr. Juan Pedro Martínez Cendán, cuya visión y experiencia como ginecólogo y docente podía ayudarnos a continuar adelante, aportando otro punto de vista diferente.

Cuando tuvimos listos los resultados iniciales de la clasificación de tumores ováricos, nos encontramos de nuevo con trabas ajenas a nuestro equipo para su publicación, que volvieron a retrasar otra vez el proyecto, de forma que nos planteamos la posibilidad de cambiar de ámbito de estudio y definir, a nuestro pesar, una nueva línea de investigación, para trabajar sobre una base de datos diferente, de creación propia, con los mismos objetivos originales. Decidimos

crear esta base de datos con imágenes microscópicas de citologías cervicovaginales y estudiar la posibilidad de realizar una clasificación automática de las mismas tras realizar una exhaustiva búsqueda bibliográfica. Para este nuevo proyecto, contamos con la colaboración desinteresada del Dr. Sebastián Ortiz Reina y parte del equipo del Servicio de Anatomía Patológica del Complejo Hospitalario Universitario de Cartagena, y especialmente, con la Dra. Ana Ortiz González. A partir de aquí, comienza una aventura apasionante, de donde nace una relación de amistad entre el equipo, que nos permite conocer y colaborar con distintas Instituciones internacionales y equipos de investigación de primera línea, en la que se mezcla la Física con la Informática y la Medicina de forma armónica y en su justa proporción para dar como resultado este trabajo.

INTRODUCCIÓN

1. INTRODUCCIÓN

1.1. Anatomía e histología del aparato genital femenino: útero, ovarios y cuello uterino

El aparato genital femenino interno se encuentra localizado en el interior de la cavidad pélvica, formada por la unión de ambos huesos coxales (a su vez formados por íleon, isquion y pubis) a través de la sínfisis púbica por delante y la articulación con el sacro por detrás, según Latarjet y cols. (2004) [5].

En el eje central de la cavidad pélvica se encuentra el útero, y anterior a él se localiza la vejiga urinaria con la uretra, y posterior a él, se sitúa el recto. Estos órganos se encuentran recubiertos por su cara exterior por peritoneo, que forma un receso anterior al útero, entre éste y la vejiga, y otro posterior al mismo entre el útero y el recto, estructuras conocidas genéricamente como fondos de saco. El fondo de saco anterior es el denominado vesicouterino y el posterior rectouterino o de Douglas, según Latarjet y cols. (2004) [5].

El útero, en sus extremos superiores y laterales se continúa con dos estructuras tubulares que comunican la cavidad uterina con la cavidad abdominal: las trompas de Falopio, que finalizan en una dilatación final, la porción ampular, de la que parten las fimbrias, que son las prolongaciones terminales digitiformes,

encargadas de desplazarse hacia el ovario en el momento de la ovulación para así recoger el ovocito, según Latarjet y cols. (2004) [5].

Los ovarios son dos órganos situados a cada lado del útero, suspendidos por el ligamento uteroovárico y ligamento infundibulopélvico, ambos vascularizados con la función de nutrir el ovario, así como también se encuentran suspendidos de las trompas de Falopio mediante sendos mesoovarios. La vascularización principal de los ovarios se realiza por medio de las arterias ováricas, que son ramas procedentes directamente desde la aorta abdominal. Estas arterias cruzan sobre los vasos iliacos externos y entran a los ligamentos infundibulopélvicos, para dirigirse hacia los ovarios, emitiendo ramas que irrigarán el mesosálpinx y el mesoovario. Las ramas ováricas y tubáricas terminales se anastomosan con las ramas homólogas terminales procedentes de las arterias uterinas que llegan en sentido ascendente desde la zona medial, formando así un arco anastomótico. El drenaje venoso se realiza a través de las venas del plexo pampiniforme cercano al ovario y a la trompa de Falopio, que después se unirán para formar las venas ováricas. La vena ovárica derecha asciende hasta desembocar en la vena cava inferior mientras la izquierda lo hace en la vena renal izquierda, según Latarjet y cols. (2004) [5]. Estos detalles se pueden identificar en la **Figura 1.1**.

Los vasos linfáticos discurren en paralelo a los vasos sanguíneos ováricos, y se unen a los procedentes de las trompas de Falopio y del fondo uterino mientras ascienden hacia los ganglios linfáticos paraaórticos, según Latarjet y cols. (2004) [5].

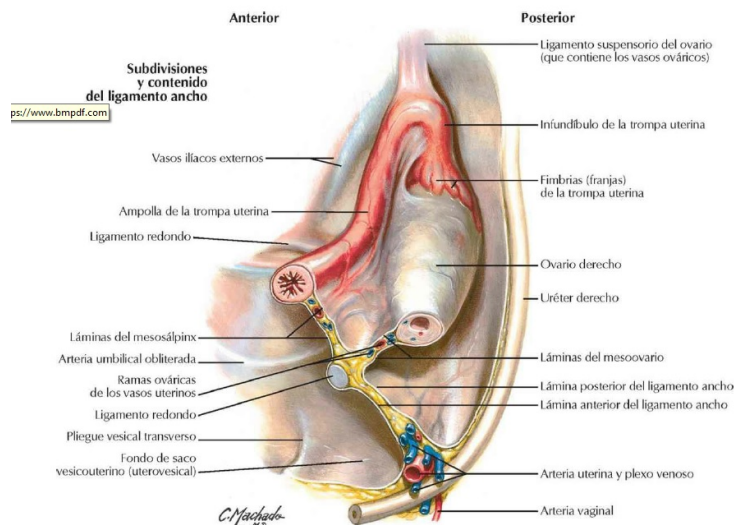


Figura 1.1: Ilustración en detalle del ovario, irrigación y relaciones, tomada de Franz Netter [1].

Los ovarios reciben inervación que discurre paralela a los vasos, procedente del plexo ovárico, que comunica con el plexo uterino. Las fibras parasimpáticas en el plexo derivan de los nervios esplácnicos pélvicos, y las fibras aferentes procedentes del ovario entran en la médula espinal mediante las raíces T10 y T11, según Latarjet y cols. (2004) [5].

Los ovarios tienen una doble función: endocrina y reproductiva.

La función endocrina consiste en la síntesis y liberación de estrógenos y progesterona fundamentalmente, acción regulada por el efecto de las hormonas FSH y LH liberadas desde la hipófisis que a su vez está controlada por la secreción pulsátil de la hormona GnRH hipotalámica, según Pallardo Sánchez y Cols. (2013) [6].

La función reproductiva se basa en la maduración de los ovocitos dentro de los folículos antrales localizados en la corteza ovárica externa, gracias a la acción estimulante de la hormona FSH, que serán recogidos en cada ciclo ovulatorio por las trompas de Falopio en espera de ser fecundados y entonces, transportar el cigoto hacia el endometrio que tapiza el interior de la cavidad uterina, según Pallardo Sánchez y Cols. (2013) [6].

Histológicamente, en el ovario se distinguen tres áreas: La más externa o corteza, la médula y la red ovárica, según Ross y cols. (2013) [7].

La corteza ovárica está compuesta por el epitelio de recubrimiento ovárico, los folículos ováricos y el tejido conectivo adyacente que conforma el estroma. El epitelio de recubrimiento es de tipo cúbico simple, conocido como epitelio germinal pues clásicamente se pensaba de forma equivocada que era en él donde se originaban las células germinales, aunque en realidad únicamente se trata de un epitelio de revestimiento, según Ross y cols. (2013) [7].

Bajo el epitelio germinal se identifica la túnica albugínea, formada por tejido conjuntivo denso irregular, debajo del que se encuentran los folículos ováricos, en distinta fase de desarrollo en función de la fase del ciclo en que se encuentre. Un folículo es la estructura formada por un ovocito más las células que lo rodean, y se encuentran a su vez rodeados de estroma, según Latarjet y cols. (2004) [5].

La médula es la zona más profunda del ovario, y está rodeada por la corteza. Los límites entre ambas estructuras no están muy definidos. La médula está formada por tejido conjuntivo fibroelástico laxo y contiene numerosos va-

sos sanguíneos, linfáticos y nervios, que entran y salen del ovario por el hilio. En la médula no hay folículos ováricos, según Latarjet y cols. (2004) [5].

El aparato genital se encuentra suspendido en la pelvis mediante una serie de fascias y ligamentos que se fijan en las paredes del mismo y en la pared abdominal, pelvis y sacro, con la función de mantener la estática y correcto funcionamiento del contenido de la cavidad pélvica, según Latarjet y cols. (2004) [5].

El cuello uterino, o también llamado cérvix, es la porción inferior del útero, que comunica la cavidad uterina con el interior de la vagina, a modo de cuello de botella, mediante el canal que lo atraviesa. Habitualmente, aunque puede haber grandes variaciones interpersonales, tiene una longitud de unos 3 a 4 centímetros y un diámetro alrededor de 3 cm, variando en función de la edad de la paciente, el número de gestaciones y partos que haya tenido, incluso en la fase del ciclo menstrual en que se examine, pues se puede ver afectado por cambios hormonales, según Sellors y cols. (2003) [8].

El útero es un órgano estructuralmente conformado por una capa serosa externa, una capa muscular denominada miometrio, y una capa mucosa interna, el endometrio. La estructura del miometrio es de tejido fibromuscular, que contiene en su interior las estructuras vasculares, linfáticas y nerviosas que lo nutren, eliminan desechos e inervan. La irrigación del útero procede directamente de las arterias ilíacas internas, a través de las arterias uterinas, que discurren en paralelo a los laterales del útero, dando sus ramas que penetran en la estructura del órgano, de entre ellas, las arterias cervicales y vaginales destacan por ser las que irrigarán el cuello uterino y el tercio superior de la vagina dirigiéndose

en sentido caudal. El cérvix tiene una estructura similar al útero, dado que es la continuación caudal de este órgano, y de forma análoga, dispone del canal endocervical central, rodeado por una estructura fibromuscular con una vascularización importante, y en su exterior un tejido mucoso que se continúa con la mucosa vaginal. El retorno venoso del cérvix se realiza a través de las venas cervicales que desembocan en las venas hipogástricas, y los vasos linfáticos se dirigen habitualmente hacia la cadena ganglionar iliaca común, iliaca externa, iliaca interna, obturatriz y parametrial. Las fibras nerviosas que lo inervan proceden de ramas originadas en el plexo hipogástrico. La inervación del exocérvix es muy escasa, por lo que su sensibilidad está muy disminuída y por ello, los procesos que lo afectan pueden tolerarse prácticamente sin molestias. Sin embargo, el canal endocervical se encuentra ricamente inervado, por lo que los procesos patológicos que lo afecten, así como los procedimientos médicos que se realicen sobre él, pueden ser más molestos, según Latarjet y cols. (2004) [5] y Singer y cols (2008) [9].

El cérvix es el punto principal de sujeción del útero, y por tanto, del aparato genital interno femenino, mediante su fijación a la pelvis a través de la unión de los ligamentos uterosacos y el ligamento cardinal de Mackenrodt. El otro elemento de sujeción a pelvis del aparato genital interno femenino son los ligamentos redondos, que fijan el útero a la pelvis por su porción más craneal, la zona cornual, según Latarjet y cols. (2004) [5].

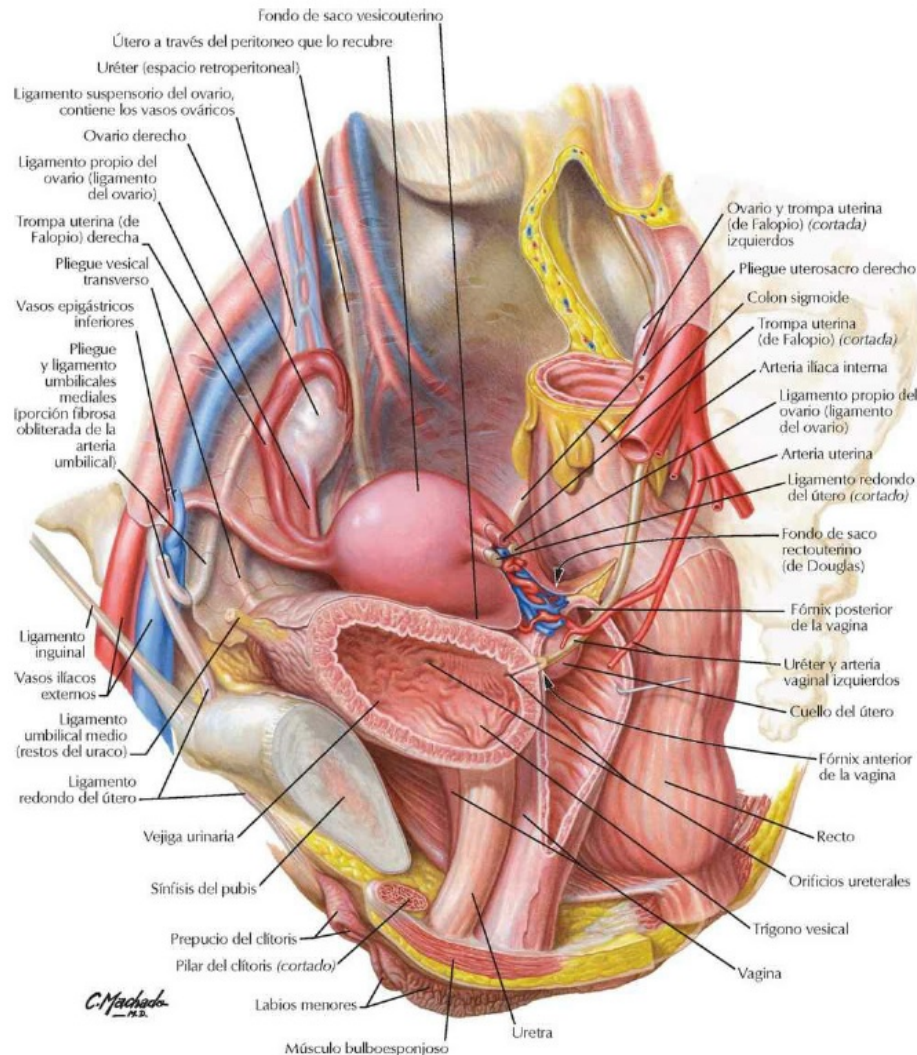


Figura 1.2: Ilustración de la pelvis femenina, tomada de Franz Netter [1].

Como se puede ver en la **Figura 1.2**, tomada de Netter [1], el extremo inferior cervical queda incluido en el interior de la cavidad vaginal, que en estado normal, se trata de una cavidad virtual, de forma que resulta accesible directamente mediante exploración vía vaginal, para su inspección manual o visualización mediante el uso de un espéculo vaginal, como el que muestra la **Figura 1.3**.

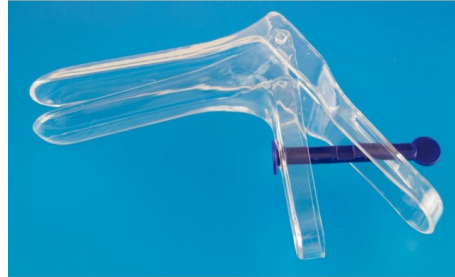


Figura 1.3: Imagen de un espéculo vaginal desechable.

La porción del cérvix accesible mediante su exploración visual con espéculo, o especuloscopia, es la mitad inferior de la estructura cervical, pudiendo visualizar la mucosa exocervical, el orificio cervical externo (OCE) que es por donde se exterioriza el canal endocervical, a través del que se comunica la vagina con la cavidad uterina mediante el orificio cervical interno (OCI), según Sellors (2003) [8].

El recubrimiento exocervical es un epitelio escamoso estratificado no queratinizado, que limita con el recubrimiento mucoso endocervical glandular cilíndrico mediante la unión escamocolumnar (UEC), según Ross y cols. (2013) [7].

El estudio histológico de la estructura del epitelio estratificado no queratinizado (o epitelio escamoso no queratinizado) cervical demuestra que en su capa basal, sobre la lámina basal del epitelio, aparece una capa de células con núcleo grande y oscuro, con un citoplasma escaso en relación al volumen del núcleo, según Ross y cols. (2013) [7]. Estas células basales proliferan y van madurando, originando otras capas celulares denominadas parabasales, que también tienen un núcleo grande y oscuro, con citoplasma escaso pero que se tiñe de color azul

verdoso cuando son teñidos con la técnica descrita por Papanicolaou (1941) [10]. Estas células parabasales también se siguen diferenciando y maduran conforme ascienden hacia la superficie del epitelio, quedando sus núcleos más pequeños y el citoplasma más grande, dando lugar a las células superficiales, que son células planas, de núcleo pequeño y denso y citoplasma claro.

Esta maduración del epitelio ocurre gracias a la influencia estrogénica, siendo fundamental la presencia del estradiol para conseguir la finalización de la misma. En la menopausia, la maduración queda detenida en las células parabasales, apareciendo de esta forma un epitelio adelgazado, con menor trofismo, y de un color pálido, según Ross y cols. (2013) [7].

El epitelio glandular cilíndrico es un epitelio monocapa que recubre el canal endocervical hasta la UEC. Dado que se trata de un epitelio monocapa, su grosor es inferior al del epitelio escamoso. Este epitelio no tiene una superficie lisa, sino que forma invaginaciones en su superficie a modo de criptas, denominadas así, criptas o glándulas endocervicales, cuya altura puede ser de hasta 8 mm desde la superficie hasta su fondo. Estas formaciones crípticas o papilares son visibles, a simple vista, lo que hace que presente un aspecto micropapilar característico, según Ross y cols. (2013) [7].

Como se ha descrito anteriormente, la UEC es la zona de unión entre el epitelio columnar endocervical y el epitelio escamoso exocervical, y es visible habitualmente a simple vista mediante la especuloscopia, identificable por ser la línea divisoria entre un epitelio cilíndrico micropapilar rojizo y el epitelio escamoso de superficie lisa y de color rosado. Su localización varía en función de la influencia hormonal predominante, por lo que cambia dependiendo de la edad,

de la toma de anticonceptivos hormonales o la gestación. En la edad infantil, pubertad e inicio de la edad fértil se localiza muy cercano al OCE, siendo esta su localización embriológica. Posteriormente, debido a la influencia estrogénica, ocurre un crecimiento en longitud del cérvix, con desplazamiento del epitelio endocervical hacia el exterior del OCE llegando a recubrir parcialmente el exocérvix, en un fenómeno denominado ectropión o ectopia, dando un color rojizo característico al cuello uterino. Este fenómeno de ectropion es un proceso fisiológico en la mujer, que conlleva la exposición del epitelio cilíndrico glandular endocervical al medio ácido vaginal, cuyo efecto directo es la pérdida del moco cervical que protege dicho epitelio, recibiendo así un daño directo por el pH vaginal, ocurriendo procesos de reparación celular mediante metaplasia, sustituyendo así este epitelio cilíndrico por el epitelio escamoso poliestratificado. Esta metaplasia se inicia en la UEC y se va desplazando de forma centrípeta hacia el OCE, desplazando así la UEC hacia el canal cervical, de forma que entre la localización inicial del OCE y la localización donde se encuentra finalmente queda una zona epitelial escamosa denominada como zona de transformación (ZT). De este modo, la UEC se va desplazando sobre la superficie del cérvix durante las distintas etapas de la vida de la mujer en función de las influencias hormonales, localizándose por norma general en el exocérvix, siendo visible durante la premenopausia en dicha zona, pero, debido al descenso de influencia estrogénica en la menopausia, va siendo desplazada hacia el canal endocervical, pudiendo quedar incluida en su interior en la etapa menopáusica de la mujer, según Sellors (2003) [8].

Característicamente, es en la zona cercana a la UEC donde se origina y localiza el cáncer de cérvix y sus lesiones precursoras, de ahí la importancia de la correcta identificación de la ZT, según Moscicki y cols. (2006) [11].

1.2. Cáncer de ovario

El cáncer de ovario ha sido, desde la Época Clásica, una patología muy agresiva en su presentación y con muy poca esperanza de vida, dado que su diagnóstico era tardío y que el acceso quirúrgico al interior del abdomen era limitado hasta finales del siglo XIX y principios del XX. Con las mejoras técnicas y científicas que permitieron abarcar quirúrgicamente esta patología, así como con el descubrimiento de fármacos quimioterápicos efectivos, la esperanza de vida global del cáncer de ovario ha mejorado ostensiblemente, junto con el aumento del número de pacientes diagnosticadas en estadios precoces gracias a exploraciones ecográficas de la pelvis, aunque sigue siendo la principal causa de mortalidad por cáncer ginecológico en la actualidad.

El cáncer de ovario se puede originar a partir de cualquiera de las células que forman parte de la estructura del órgano, pero aproximadamente el 90 % de los cánceres ováricos diagnosticados son de tipo epitelial, que deriva de células pluripotenciales del epitelio celómico de recubrimiento. Entre el 10 y el 20 % de las neoplasias son de bajo potencial de malignidad, también conocidas como "borderline", y se caracterizan por una proliferación celular alta sin invasión del estroma, según Darai y cols. (1996) [12].

De las neoplasias invasivas epiteliales, entre el 75 y el 85 % son serosas, originadas en el epitelio de superficie o en quistes de inclusión sobre el estroma, el 10 % son mucinosas simulando células endocervicales o intestinales y el 10 % son endometrioides que imitan al tejido endometrial y se originan a partir de la malignización en células de quistes endometriósicos. Las neoplasias de célu-

las claras son minoritarias, también originadas en quistes endometriósicos, así como también es raro encontrar tumores de Brenner, microcíticos e indiferenciados. Además de las neoplasias epiteliales, hay otros tipos de neoplasias derivadas de otras estirpes celulares, como las derivadas de los cordones sexuales que corresponden al 6 %, de las células germinales que son el 3 % o bien tumores indeterminados y metástasis que corresponden al 1 %, según Eifel y cols. (2006) [13].

1.2.1. Incidencia y mortalidad mundial del Cáncer de Ovario

La estimación aproximada de incidencia de cáncer en términos generales en el mundo es de 18,1 millones de nuevos casos, y 9,6 millones de muertes por cáncer en 2018, según los datos publicados por GLOBOCAN (2018) [14], que recoge los datos de incidencia y mortalidad por cáncer en 2018 en 185 países, según Bray y cols. (2018) [15].

La incidencia global anual del cáncer de ovario es de 9,9 casos cada 100.000 personas, siendo el sexto cáncer en frecuencia en mujeres, la quinta causa de muerte por cáncer en mujeres y la primera causa de mortalidad por cáncer ginecológico en países desarrollados. La incidencia del cáncer de ovario es más elevada en el Norte y Oeste de Europa, en Estados Unidos e Israel, mientras que es menor en Japón y países en vías de desarrollo, datos publicados por Roldán (1998) [16], Zubik y cols. (2008) [17] y la Asociación Española Contra el Cáncer (2018) [18].

La razón principal por la que presenta una mortalidad tan alta es que se trata de una patología que habitualmente se diagnostica en estadios avanzados, esto es, estadios III y IV de clasificación de la Federación Internacional de Gi-

necología y Obstetricia (FIGO), en los que la enfermedad se encuentra en localizaciones extrapélvicas o ya diseminada a distancia, y la tasa de supervivencia a largo plazo es inferior al 20%. Sin embargo, alrededor del 90% de pacientes diagnosticadas en estadio I pueden ser curadas, por lo que sería ideal disponer de una herramienta capaz de detectar esta enfermedad de forma precoz, según Eifel y cols. (2006) [13].

La mortalidad por cáncer de ovario es un indicador complejo que depende no solo de la incidencia del mismo y sus determinantes, sino también de la posibilidad de diagnóstico precoz, disponibilidad de tratamiento óptimo y cuidados del mismo.

En Europa, la posibilidad de tratamiento óptimo es mayor, y, gracias a los avances en su tratamiento en las últimas décadas, la supervivencia a 5 años para el cáncer de ovario aumentó ligeramente en el periodo 1988-1999 (38-41%). La tendencia no presenta mucha variación en relación a la edad pero sí entre diferentes países (29-43%), en posible relación con la disponibilidad de tratamientos y detección en estadios más iniciales de la enfermedad. El mayor aumento en la supervivencia en dicho periodo ocurrió en Finlandia con un 6%, según Bray y cols. (2005) [19].

1.2.2. Incidencia y mortalidad en España por Cáncer de Ovario

En España, en el periodo entre 1980 y 2006, el cáncer de ovario fue causa de 36.157 muertes en España. El análisis de tendencias, en ese periodo, muestra un acusado aumento de la mortalidad hasta 1998 (4,3% anual; IC 95% = 3,9-4,8), estabilizándose posteriormente. Se observan diferencias importantes con la edad:

en menores de 50 años, aumentó la mortalidad un 1,6 % anual, mientras que las mayores de 50 años lo hicieron un 4,4 % anual, y en mayores de 65 años aumentó a 5,8 % anual, según Cabanes y cols. (2009) [20]. Para 2019, el número estimado de fallecimientos por cáncer de ovario es de 3.548 mujeres, según la Red Española de Registros de cáncer (2019) [21].

Aunque los cambios en la incidencia han sido menores, las cohortes de mujeres más jóvenes presentan menor riesgo de cáncer de ovario, probablemente, debido al mayor uso de anticonceptivos orales, según la publicación de *Epidemiological Studies of Ovarian Cancer Collaborative Group* (2008) [22]. Las mujeres de más edad tienen menos probabilidad de recibir el tratamiento estándar del cáncer de ovario por la mayor comorbilidad que habitualmente presentan, así como por la peor tolerancia a la toxicidad de los quimioterápicos. El tratamiento incompleto o subóptimo en este grupo de edad es una de las razones para el peor pronóstico que estas pacientes presentan con respecto a las más jóvenes, que no se explica por las características propias del tumor, según la publicación de *Epidemiological Studies of Ovarian Cancer Collaborative Group* (2008) [22].

En resumen, la mortalidad por cáncer de ovario descendió en España a finales del siglo XX, de igual forma que en Europa, aunque la magnitud de este cambio varió con la edad y los diferentes países.

1.2.3. Etiopatogenia del Cáncer de Ovario

La edad media en la que se diagnostica el carcinoma de ovario es de 60 años, y el riesgo de padecer un cáncer de ovario a lo largo de la vida para una mujer en los países desarrollados es de 1 entre 70.

La etiología parece ser multifactorial, asociando factores genéticos predispo-

nentes, ambientales y hormonales entre otros. Se han desarrollado varias teorías para explicar la génesis del cáncer ovárico, según Kurman y Shih (2010) [23]:

- Teoría de la ovulación incesante: Desarrollada por Fathalla en 1971, plantea que la neoplasia epitelial de ovario está relacionada con la continua disrupción de la superficie epitelial que ocurre con cada ovulación, que conlleva rotura y reparación epitelial en cada ciclo. Esta teoría se fundamenta en la observación de la mayor probabilidad de presentar cáncer de ovario en mujeres nulíparas, con menopausia tardía y en el menor riesgo en pacientes multíparas, con la toma de anovulatorios y en madres que han dado lactancia materna, según Bailey y cols. (1998) [24]. Sin embargo, esta teoría no explica el mayor riesgo de presentar cáncer de ovario en mujeres infértiles anovuladoras.
- Teoría del exceso de gonadotropinas: Intenta explicar la limitación de la teoría anterior, estableciendo que las gonadotropinas aumentan el riesgo de cáncer de ovario, según Cramer y Welch (1983) [25], aunque trabajos posteriores refutan la teoría argumentando que en estas pacientes el riesgo de cáncer parece estar aumentado en relación con el hiperandrogenismo, a expensas principalmente de la hormona Dehidroepiandrosterona (DHEA), en lugar de por el aumento de gonadotropinas, según Helzlsouer y cols. (1995) [26].
- Teoría de Parmley y Woodruff: Plantea que el epitelio donde se origina el cáncer de ovario es susceptible de recibir agresiones externas por vía ascendente, de forma que contaminantes como el polvo de talco, la inflamación crónica y otros factores pueden favorecer el desarrollo de cáncer de ovario, de forma similar al comportamiento de los mesoteliomas. Esto explica la razón de por qué las pacientes con ligadura tubárica o histerectomía presentan un menor riesgo, según Parmley y Woodruff (1974) [27].

- Teoría del origen Müllleriano: Esta teoría plantea que el cáncer de ovario no se desarrolla a partir del epitelio de superficie ovárico, sino a partir de restos del sistema müllleriano, según Dubeau (1999) [28], aunque no hay estudios posteriores de calidad que apoyen esta hipótesis.

Según avanza el conocimiento sobre el cáncer de ovario, vemos que las tumoraciones epiteliales tanto benignas como malignas tienen etiologías similares. También se pone de manifiesto la necesidad de realizar un seguimiento de determinadas lesiones benignas en ciertas circunstancias, pero según la evidencia histopatológica, se ha demostrado que la progresión a cáncer desde lesiones benignas ocurre cuando se trata de tumoraciones de tipo mucinoso y en el seroso de bajo grado, pero no ocurre este fenómeno en el cáncer epitelial seroso de alto grado. Generalmente, se trata de una enfermedad monoclonal, originada a partir de una célula del ovario, según Jacobs y cols (1992) [29], aunque algún autor ha propuesto que el tipo seroso papilar derivado del peritoneo en pacientes con mutación en el gen BRCA1 es un tumor policlonal, como refleja Schorge y cols. (1998) [30], siendo la mutación en el gen BRCA1 la que origina una cadena de acontecimientos en las células que derivan en otras mutaciones hasta dar lugar al cáncer con distintos clones celulares.

Por el momento hay descritos más de 60 oncogenes y moléculas relacionadas con el cáncer de ovario, según Berchuck y cols. (1990) [31] y Wan y cols. (1997) [32], lo que dificulta la identificación del momento exacto en el que se produce la transformación displásica de las células normales, según Rasila y cols. (2005) [33].

De entre los factores de riesgo descritos para el cáncer de ovario, el de mayor asociación es la historia familiar, generalmente relacionada con la presencia de

mutaciones en genes BRCA 1 y 2, que están encargados del control de mutaciones en el genoma celular. Este cáncer heredofamiliar supone el 10-15 % de todos los casos de cáncer de ovario diagnosticados. Las mutaciones en estos genes, representan aproximadamente el 25 % de los casos de cáncer de mama y de ovario familiar, lo que sugiere que el resto de casos con susceptibilidad genética serían de naturaleza poligénica, con varios genes de baja penetrancia implicados. Estos casos se asociarían a un riesgo menor de cáncer de mama u ovario, pero en combinación con otros factores ambientales o genéticos podrían aumentar significativamente este riesgo.

En la actualidad, se estima que el riesgo acumulado a los 70 años de edad para el cáncer de ovario es de entre un 20 % y un 40 % para mujeres portadoras de mutaciones en el gen BRCA1, y entre un 15 % y un 20 % para las portadoras de mutaciones en el gen BRCA2, pero esta probabilidad puede verse modificada por la influencia de otros genes y factores ambientales. En relación a la edad, el riesgo para pacientes con mutación en el gen BRCA1 se empieza incrementar a partir de los 40 años, y para las portadoras de mutación en el BRCA2 a partir de los 50, persistiendo el aumento hasta la séptima década de la vida. Así mismo, el riesgo de presentar cáncer de ovario después de haberlo presentado en mama es del 12,7 % en pacientes con mutación en el gen BRCA1 y el 7 % con mutación en BRCA2, según Eifel y cols. (2006) [13].

Otros factores de riesgo asociados al cáncer de ovario son los factores reproductivos, con un riesgo de asociación menor que los hereditarios, pero parecen ser los responsables de las variaciones en incidencia y mortalidad por cáncer ovárico en los países desarrollados, en las últimas décadas. El uso de anticonceptivos orales durante al menos 5 años parece ser un factor protector

para cáncer de ovario, con estudios que reportan una reducción de hasta el 40 % de riesgo en pacientes usuarias de anticonceptivos frente a las que nunca los han utilizado, independientemente de la dosis, formulación y edad, y, aunque el efecto protector disminuye progresivamente al dejar su uso, parece que se mantiene durante al menos 30 años, según Stanford (1991) [34] y el *Epidemiological Studies of Ovarian Cancer Collaborative Group* (2008) [22]. De igual forma, la multiparidad es un factor protector, dado que en 12 estudios de caso control frente a nulíparas en Estados Unidos, se demuestra un 40 % de disminución de riesgo tras el primer parto y un 14 % adicional con cada parto posterior, aunque no en todos los estudios valorados se confirma, según Whittemore y cols. (1992) [35]. Respecto al tipo histológico, esta protección se refiere a los de tipo epitelial, borderline e invasores, sin embargo, para el tipo mucinoso no se ha demostrado asociación con factores reproductivos. El efecto de los abortos es controvertido debido a la falta de acuerdo en su definición y a que no siempre se comunican las interrupciones legales, pero en general, parece ser que los abortos también se asocian a un ligero efecto protector para el cáncer de ovario.

Respecto a la lactancia, las madres lactantes presentan una Odds Ratio de 0,9-0,6 frente a las que nunca han dado pecho. Ha habido estudios que han relacionado la duración de la lactancia con la reducción del riesgo, pero el estudio de tendencia no ha sido consistente, por otra parte, en cambio, otros estudios no encuentran asociación o incluso presentan una asociación positiva, según Whittemore y cols. (1992) [35].

La edad de la menarquia y la menopausia han sido valoradas también, demostrando una asociación débil, con una relación inversamente con la edad de la menarquia, y una asociación positiva con la edad de la menopausia, con unos

valores de riesgo relativo de 1,5-2,9 para el grupo de mayor edad en la menopausia frente a las más jóvenes.

La edad del primer embarazo no se ha podido relacionar como factor de riesgo o protector para el cáncer de ovario, pues hay estudios contradictorios.

El tratamiento hormonal sustitutivo en la menopausia, a diferencia del tratamiento hormonal anticonceptivo en edad fértil, parece asociarse en algunos estudios a un ligero aumento del riesgo de cáncer de ovario. En un estudio colaborativo en el que se incluyó a casi un millón de mujeres, según Beral y cols. (2007) [36], el riesgo relativo para usuarias de terapia hormonal sustitutiva durante 5 años o más para presentar un cáncer de ovario de tipo seroso fue de 1,2, con IC (1,09-1,38), sin encontrar asociación para el resto de tipos histológicos. Este riesgo se demostró independiente del tipo de formulación, y descendió rápidamente tras finalizar los tratamientos. Otros estudios observan un mayor riesgo asociado al uso de estrógenos solos frente a la terapia hormonal combinada, o incluso reportan ausencia de asociación de la terapia hormonal combinada y el riesgo de padecer cáncer de ovario.

La endometriosis es un factor de riesgo leve, con un riesgo relativo de 1,6 (IC 95 % de 1.12-2.09) para el cáncer de ovario en general, pero se ha demostrado una asociación positiva para el tipo histológico endometriode y de células claras, según Ñiguez Sevilla y cols. (2018) [37], en relación con una mutación en el gen supresor ARID1A, que se observa tanto en las lesiones endometriosis y cáncer de ovario cuando se presentan en la misma paciente, según Wiegand y cols. (2010) [38].

La diferente tasa de incidencia entre distintos países hizo pensar en la influencia de factores ambientales y hábitos de vida como factores implicados en la génesis del cáncer de ovario, prestando especial atención a los factores dietéticos, con la publicación de varios estudios de cohortes y tamaños muestrales grandes, para finalmente, realizar metaanálisis de los mismos: El consumo de lácteos se ha relacionado desde hace tiempo con un mayor riesgo de cáncer de ovario, pero no ha sido posible demostrar esta relación, incluso con un metaanálisis de 12 estudios de cohortes con más de medio millón de mujeres como muestra Genkinger y cols. (2006) [39], por lo que hoy en día, se acepta que no hay evidencia epidemiológica de relación entre el consumo de lácteos y el riesgo de sufrir cáncer de ovario. También se ha estudiado de similar forma la posible influencia de otros factores dietéticos, como la ingesta de dietas hipercalóricas, colesterol, huevos, sin demostrar relación con el cáncer de ovario, según Genkinger y cols. (2006) [40]. De igual manera, se ha estudiado el posible efecto protector de antioxidantes como los carotenoides y el consumo de fruta y vegetales, sin hallar ninguna relación según Koushik y cols. (2006) [41]. Los folatos sólo parecen ejercer como factor protector en mujeres consumidoras de más de 20g de alcohol al día según Larsson y cols. (2004) [42]. El ajo y la cebolla por el contenido en flavonoides, parecen ser protectores de múltiples tumores, y entre ellos, del cáncer de ovario, según Galeone y cols. (2006) [43].

El tabaco se ha demostrado como factor de riesgo para el tipo histológico mucinoso, tanto en su presentación borderline como en la invasora, con leve efecto dosis dependiente.

La influencia de la obesidad y el sedentarismo tampoco demuestra influencia en el cáncer de ovario por el momento.

1.2.4. Tipos de Cáncer de Ovario

La tendencia general es a considerar el cáncer de ovario como un grupo homogéneo de cáncer, con una única estrategia de cribado, prevención y tratamiento, pero en realidad se trata de un grupo de cánceres de tipos histológicos muy diferentes, con comportamientos también diferentes. En el 75-90% de los casos, se tratará de tumores ováricos epiteliales, pero en el 10-25% restante puede tratarse de tumores ováricos de los cordones sexuales (de células del estroma de la granulosa, androblastomas, de células esteroideas, gynandroblastoma, o indiferenciado), tumores de las células germinales (disgerminoma, tumor del saco vitelino, carcinoma embrionario, poliembrioma, coriocarcinoma, teratomas, tumores mixtos o gonadoblastomas), tumores de tejidos blandos no específicos del ovario, tumores inclasificables o tumores metastásicos [13].

Si se atiende al comportamiento biológico de los distintos tumores de ovario, se puede diferenciar por un lado, los tumores de bajo grado, que presentan un comportamiento indolente, de crecimiento lento, que en el momento del diagnóstico suelen estar confinados en el ovario o pueden expandirse fuera del mismo, que corresponden a los tumores borderline. Estos tumores son genéticamente estables, con mutaciones de distintos genes que se diferencian según el tipo histológico. Por otro lado, se puede diferenciar los tumores de alto grado, que son los más frecuentes y presentan un comportamiento mucho más agresivo, por lo que en el momento del diagnóstico suelen encontrarse extendidos fuera del ovario, invadiendo órganos pélvicos o incluso fuera de la pelvis. Estos tumores de alto grado son genéticamente más inestables que los anteriores.

1.2.5. Manejo clínico de los tumores ováricos

En la práctica clínica habitual de un ginecólogo, la prueba complementaria por excelencia es la ecografía pélvica, siendo el estudio básico y fundamental que se utiliza para detectar anomalías estructurales pélvicas. En la **Figura 1.4** se muestra una sala de exploraciones ginecológicas, donde, como se puede ver, el ecógrafo cuenta con un protagonismo especial. Entre las anomalías detectadas ecográficamente, las tumoraciones ováricas son de las más frecuentemente manejadas en la práctica clínica diaria, y a veces en su presentación pueden ofrecer grandes dudas diagnósticas.

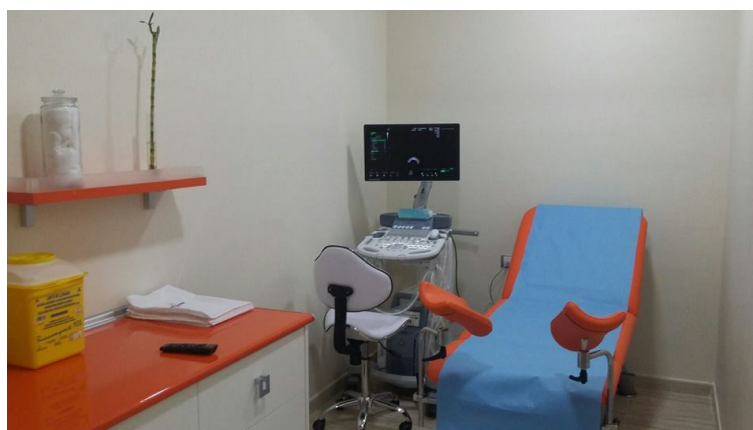


Figura 1.4: Sala de exploraciones ginecológicas, con ecógrafo de altas prestaciones.

La exploración ecográfica se ha convertido en la técnica principal para estudiar, categorizar y plantear el seguimiento de la patología anexial, dado que en manos de un examinador experimentado, es la prueba que presenta la mejor correlación diagnóstica entre benignidad o malignidad de forma preoperatoria, según Meys y cols (2016) [44]. Se debe resaltar que se trata de una exploración no invasiva, en la que la paciente no recibe radiación alguna, siendo por tanto inocua. Esta técnica permite valorar con gran definición el tamaño de la tumo-

ración y la presencia de signos de sospecha de malignidad, tales como la existencia de tejido sólido dentro de la tumoración, heterogeneidad estructural del tumor, presencia de estructuras papilares y su número, así como la presencia de ascitis. La exploración ecográfica puede realizarse vía abdominal o vía transvaginal, siendo esta segunda vía la preferente si es posible, utilizando para ello un transductor a una frecuencia de 5 a 7,5 MHz.

El motivo por el que se recomienda la utilización de la vía transvaginal para realizar esta exploración es porque de esta forma, el transductor se localiza en la posición más cercana posible a las estructuras que deseamos estudiar, obteniendo así imágenes de mayor resolución, ofreciendo una señal Doppler más sensible y disminuyendo la posibilidad de aparición de artefactos en la imagen, aunque presenta la limitación de que el campo estudiado no es muy grande y no puede abarcar masas de gran tamaño o que se encuentren en una localización alta de la pelvis. La vía abdominal sólo se recomienda en caso de mujeres nulíparas, o lesiones ováricas de gran tamaño que no puedan ser abarcadas por la vía vaginal, permitiendo también examinar lesiones asociadas en abdomen superior, sugestivas de enfermedad diseminada tales como tumores metastásicos en peritoneo, epiplon o incluso en hígado o bazo, así como la posibilidad de valorar adenopatías en cadenas ganglionares linfáticas, según Bajo Arenas y cols (2009) [45]. Para realizar la exploración ecográfica de la pelvis vía abdominal se recomienda utilizar un transductor a una frecuencia de 3,5 MHz, y realizar la técnica con la vejiga urinaria repleta, pues de esta forma se mantiene a las asas intestinales fuera del campo de estudio, además de que la vejiga, con su contenido líquido, hace efecto ventana, aumentando la resolución de las imágenes obtenidas, sirviendo además como estándar de comparación para valorar la densidad y naturaleza de los contenidos de los quistes anexiales. Sin embargo,

esta vía no ofrece buenos resultados cuando se trata de pacientes incapaces de mantener la vejiga llena, obesas y en pacientes con útero en retroposición, en las que el fondo uterino se encuentra fuera de la zona focal del transductor, según Eifel y cols. (2006) [13].

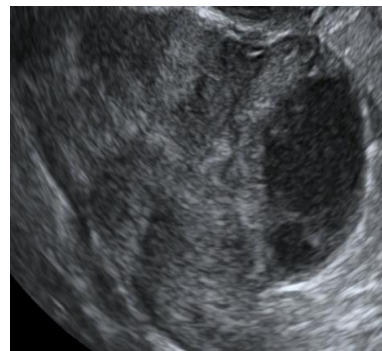
Una vez que se detecta una tumoración ovárica durante una exploración, la prioridad del ginecólogo es determinar si esa masa corresponde a un tumor benigno o maligno, con lo que planteará el manejo clínico más adecuado para cada caso. El interés principal del ginecólogo por caracterizar correctamente la benignidad o malignidad de una tumoración ovárica radica en que el abordaje terapéutico difiere mucho si se trata de una u otra entidad. Tratar una tumoración como maligna supone que la cirugía debe ser realizada por un ginecólogo oncólogo especializado, movilizándose unos recursos asistenciales muy importantes, tales como listas de espera prioritarias, técnicas de diagnóstico por imagen como tomografía axial computerizada o resonancia magnética, reserva de quirófanos para cirugías de mayor duración, realización de biopsias intraoperatorias, entre otros, por lo que es fundamental discriminar de forma previa a la intervención quirúrgica cuáles son las pacientes que se beneficiarán de la movilización de todos estos recursos, frente a las que, por presentar una patología benigna, no requieren de este despliegue de medios.

El análisis de la imagen ecográfica por parte del especialista consiste en un proceso subjetivo, en el que influye tanto la experiencia del mismo profesional, como la disponibilidad de otros datos clínicos y de exploraciones complementarias que pueden ayudar a clasificarla de forma adecuada. Afortunadamente, la mayoría suelen ser tumores benignos (80-85 %), como la tumoración de la **Figura 1.5a**, y presentan una incidencia máxima entre los 20 y los 44 años de edad,

según Berek y cols. (2015) [46]. El riesgo de presentar una tumoración maligna aumenta cuando el tumor se detecta en mujeres fuera de edad reproductiva, esto es, en prepúberes y postmenopáusicas, dado que durante el periodo fértil, entre la menarquia y la menopausia, los ovarios se encuentran bajo la influencia de las hormonas del eje hipotálamo-hipofisario-gonadal que estimulan su funcionamiento, y por tanto el desarrollo folicular y ovulación, pudiendo determinar la aparición de quistes funcionales, y otras tumoraciones benignas secundarias al funcionamiento de los ovarios tales como los cuerpos lúteos, cuerpos lúteos hemorrágicos, o endometriomas, entre otros.



(a)



(b)

Figura 1.5: Ejemplos de imágenes ecográficas de tumoraciones ováricas procedentes de la base de datos creada por el Servicio de Ginecología del Hospital KU de Leuven, Bélgica. (a) Tumoración ovárica benigna; (b) Tumoración ovárica maligna.

Los marcadores analíticos por sí mismos, tales como el marcador tumoral Ca125 han demostrado tener un gran valor en el seguimiento de las posibles recurrencias tras el tratamiento de la enfermedad, pero presentan una baja sensibilidad y especificidad como herramienta de screening, según Eifel y cols. (2006) [13]. Además, en la mitad de los casos de cáncer de ovario detectados

en estadio I aparece el nivel de Ca125 dentro de los rangos de la normalidad, mientras que otras patologías ginecológicas benignas tales como la endometriosis, leiomiomas o enfermedad inflamatoria pélvica pueden presentar este marcador elevado, según Eifel y cols. (2006) [13]. Diversos estudios poblacionales han realizado ecografías transvaginales como prueba de screening poblacional, demostrando una especificidad limitada para la detección del cáncer ovárico, por lo que en mujeres sin un riesgo aumentado de presentar cáncer de ovario no se recomienda realizar ningún tipo de cribado poblacional, según Eifel y cols. (2006) [13] y Menon y cols. (2009) [47]. La mayor limitación del estudio ecográfico es que al tratarse de un proceso subjetivo, su resultado es altamente dependiente del examinador y de su experiencia, por lo que presenta una gran variabilidad diagnóstica interobservador, según Yazbek y cols. (2010) [48], dado que en muchas ocasiones, las tumoraciones no presentan una imagen ecográfica claramente benigna o maligna como podemos ver al comparar la **Figura 1.6** con las imágenes de la **Figura 1.5**.



Figura 1.6: Tumoración ovárica benigna, de difícil caracterización, procedente de la base de datos creada por el Servicio de Ginecología del Hospital KU de Leuven, Bélgica.

Esta variabilidad en la presentación se basa, como se ha expuesto con anterioridad, en que el cáncer de ovario no es una entidad diagnóstica única, sino

que son diversas las estirpes celulares las que pueden dar origen a un cáncer ovárico y por ello tener diferente comportamiento. Esta dificultad para la clasificación es debida a la heterogeneidad de las posibles presentaciones clínicas y en pruebas de imagen, y es el motivo por el que diversos grupos de investigación han desarrollado distintas herramientas para intentar ayudar a los profesionales a interpretar y clasificar imágenes ecográficas de tumores ováricos.

Como se ha dicho, no existe un cribado poblacional para el cáncer de ovario, debido a que es una patología que, aunque su morbimortalidad es alta, presenta una baja incidencia y su periodo preclínico es muy corto, dificultando su diagnóstico en estadios iniciales en la mayoría de los casos, siendo este su mayor inconveniente para desarrollar programas dirigidos a su detección precoz en estadios iniciales.

Sin embargo, hay ciertos grupos poblacionales que presentan unas características que los hacen más susceptibles que el resto de presentar un cáncer ovárico, por lo que en estos grupos sí se recomienda una estrategia de seguimiento para detección precoz en caso de no realizar cirugía profiláctica según las circunstancias clínicas y sociales individuales de cada caso. Dichos grupos poblacionales son aquellas mujeres que presentan historia familiar de cáncer de ovario u otros cánceres relacionados (como cáncer de mama, endometrial o de colon), o portadoras de mutaciones genéticas con una conocida relación con el cáncer de ovario, como son los genes BRCA1 y BRCA2 entre otros. Los criterios de sospecha para pensar en la posibilidad de la existencia de una mutación en uno de estos genes son la presencia de varios individuos en la misma línea familiar con antecedentes de cáncer de ovario o mama, la presencia de individuos en la familia con carcinomatosis peritoneal o cáncer de trompa de Falopio

a cualquier edad, la presencia de un individuo en la familia con cáncer ovárico y de mama sincrónicos o no, antecedentes familiares de cáncer de mama en mujer premenopáusica o cáncer de ovario en mujer con ascendencia judía Ashkenazi, según Eifel y cols. (2006) [13]. En los casos en los que se sospeche la posibilidad de la existencia de una mutación en uno de estos genes, estaría indicado realizar una determinación genética, a poder ser del caso guía, es decir, de la paciente que padece el cáncer, y de ser positiva la mutación, se realizará el análisis genético a toda la familia. En estas pacientes, además de realizar ecografías transvaginales semestrales, una vez hayan cumplido su deseo genésico, se recomienda plantear la realización de una anexectomía bilateral para reducir el riesgo de presentar cáncer de ovario, según Kauff y cols. (2002) [49] y Rebbeck y cols. (2002) [50], encontrando en hasta el 5% de las pacientes que se realizan esta intervención de forma profiláctica la presencia de un cáncer tubárico u ovárico oculto según Eifel y cols. (2006) [13]. En estos casos, la realización de seguimiento ecográfico periódico no aporta una reducción de la mortalidad en este grupo de riesgo concreto, por lo que debe mantenerse únicamente hasta cumplir el deseo genésico y plantear la doble anexectomía una vez cumplido, o bien, a partir de los 40 años.

El fenotipo de cáncer de ovario que aparece en portadoras de estas mutaciones tiene tendencia a ser el adenocarcinoma invasor de alto grado con mayor frecuencia, que sugiere mal pronóstico. Sin embargo, estos casos tienen mayor supervivencia a los 5 años que el cáncer de ovario esporádico, debido en parte a la mejor respuesta que presentan a la quimioterapia con cisplatino.

Otro grupo de riesgo corresponde a familias con una enfermedad hereditaria denominada síndrome de Lynch, descrita por Lynch y cols. (2003) [51], que

representa al 3 % de los casos de los carcinomas colorrectales. Esta enfermedad tiene un patrón de herencia autosómica dominante, debida a una mutación de los genes reparadores del ADN celular. El 90 % de los casos corresponden a mutaciones de los genes MSH2 y MLH1, y menos frecuentemente en otros genes como el MSH6, que se relaciona con un síndrome de Lynch de mejor pronóstico. El carcinoma que más frecuentemente aparece en estos casos es el colorrectal, con una edad media de aparición más temprana (45 años), pero también se asocia al cáncer de ovario, adenocarcinoma de endometrio, gástrico, intestino delgado, cerebral, páncreas, y tracto urotelial superior. El riesgo de presentar un adenocarcinoma colorrectal se aproxima al 90 %, seguido del adenocarcinoma de endometrio, con un 60 % y del cáncer de ovario con un 10-12 %.

El riesgo en el síndrome de Lynch aparece desde edades tempranas, de forma que la vigilancia y cribado de cánceres en estas pacientes debe iniciarse a los 25-30 años de edad. Es importante destacar que el cáncer “centinela” para diagnosticar el síndrome de Lynch no necesariamente es un cáncer de colon, pudiendo presentarse en la mitad de los casos un cáncer de endometrio como debut. Este hecho implica que las unidades de Oncología Ginecológica, ante el diagnóstico de un cáncer de endometrio, sobre todo en pacientes jóvenes, deben investigar la posibilidad de que se trate de un síndrome de Lynch para planificar tratamientos y el seguimiento de la paciente y su familia para la detección precoz de otros cánceres relacionados.

En los casos de síndrome de Lynch se recomienda hacer una vigilancia estrecha, con colonoscopia total anual, pudiendo valorarse la colectomía subtotal profiláctica en casos seleccionados. Respecto a tumores ginecológicos, se aconseja seguimiento mediante ecografía transvaginal y determinación de Ca125

anual, pero dada su baja sensibilidad y especificidad, una vez cumplido el deseo genésico, se debe ofrecer la posibilidad de realizar histerectomía con salpingo-forectomía bilateral profilácticas, según Schmeler y cols. (2006) [52].

Entre los primeros trabajos que intentaron objetivar la discriminación entre tumoraciones ováricas benignas y malignas, destaca como uno de los pioneros el *Risk of Malignancy Index* (RMI), publicado por Jacobs y cols. (1990) [53], y que posteriormente sirvió como punto de partida para otros muchos grupos de investigación. Este sistema realiza un cálculo de riesgo para cada paciente, utilizando como variables ponderadas la edad de la paciente, una puntuación de la imagen ecográfica asignada por el observador en función de las guías establecidas por los autores a este efecto, estado de menopausia, una puntuación de la situación clínica y el nivel sérico de Ca125. Jacobs realizó distintos cálculos matemáticos, encontrando que el índice más representativo estaba compuesto por el nivel sérico de Ca125, puntuación ecográfica y estado menopáusico, con una sensibilidad del 85 % y especificidad de 97 %. Como vemos, se trata de un sistema que requiere unos datos clínicos como son la determinación analítica en sangre del nivel del marcador tumoral Ca125 que a veces no se encuentra disponible en el momento de la exploración inicial de la paciente, y requiere una segunda evaluación del caso tras realizar la determinación analítica, lo que puede demorar la toma de decisiones en función de la disponibilidad del resultado y de la saturación de la consulta del clínico, desde varios días hasta incluso semanas.

Varios años después de la publicación del RMI, el grupo IOTA (*International Ovarian Tumor Analysis*) desarrolló un documento de consenso para estandarizar los Términos y Definiciones utilizados en el estudio y valoración de las

imágenes ecográficas de ovario, según Timmerman y cols. (2000) [54], como son la existencia de proyecciones papilares, irregularidad en paredes internas de lesiones quísticas, presencia de ascitis o flujo vascular anormal. Este trabajo del grupo IOTA, publicado por Timmerman y cols. (2000) [55] y (2008) [55] ha tenido una gran repercusión internacional, pues se trata de la primera estandarización clínica en el examen de las imágenes ecográficas de tumores ováricos que se extiende y acepta mundialmente.

En 1999, este mismo grupo, IOTA, encabezado por Timmerman (1999) [56], también publicó la aplicación de Redes Neuronales Artificiales, en inglés *Artificial Neural Networks* (ANNs) al problema de la caracterización de las imágenes ecográficas de tumores ováricos, y comparó sus resultados con la impresión diagnóstica subjetiva de expertos, dos modelos de Regresión Logística, en inglés *Logistic Regression*, desarrollados por ellos (LR 1 y LR2) y el RMI, obteniendo una sensibilidad del 95,9 % y especificidad del 93,5 % con ANNs, aunque había un cierto número de imágenes en las que el resultado de su análisis resultaba inconcluyente. Los modelos de Regresión Logística ofrecen un resultado probabilístico, entre 1 y 100, y permiten fijar un punto de corte a partir del que se obtiene una detección mayor de tumores malignos de ovario con un índice de falsos positivos aceptable. Este grupo de trabajo ha continuado proponiendo y validando lo que denominaron Reglas Simples basadas en características ecográficas, trabajos publicados por Timmerman y cols. (2008) [55] y (2010) [57], y han determinado que los mejores resultados diagnósticos se obtienen aplicando las Reglas Simples como prueba de cribado inicial y después, en un segundo escalón diagnóstico, cuando las Reglas Simples ofrecen un resultado inconcluyente, es la opinión subjetiva del examinador la que finalmente clasificará la imagen como posiblemente benigna o maligna, obteniendo una sensibilidad del

91 % y especificidad del 93 % siguiendo esta metodología. Como vemos, este sistema continúa dependiendo de la valoración subjetiva de un experto, que en el ámbito hospitalario de alto nivel suele estar disponible, pero cabe la posibilidad de que, en clínicas de menor envergadura, el caso no pueda ser valorado por un especialista experimentado, requiriendo por ello, o bien una derivación a otro centro asistencial, o bien la posibilidad de un diagnóstico erróneo. Un metaanálisis reciente publicado por Kaijser y cols. (2014) [58] sugiere que la caracterización prequirúrgica de cualquier masa anexial debería incorporar el uso de las Reglas Simples o el modelo de regresión logística LR2, ambos sistemas, desarrollados por el grupo IOTA, especialmente en mujeres en edad fértil, pues esta estrategia ha demostrado los mejores resultados en estudios de validación con una sensibilidad y especificidad en mujeres premenopáusicas para la LR2 del 85 % y del 91 % respectivamente, y para las Reglas Simples del 93 % y el 83 % respectivamente.

Las Reglas Simples son un método que intenta simplificar el análisis del riesgo de malignidad de una tumoración ovárica, para que pudiera ser aplicable de forma universal. Consiste en un grupo de características ecográficas, y dependiendo de la presencia o ausencia de ellas, se clasificará la imagen como benigna o maligna. Las imágenes benignas son aquellas que presentan al menos una característica de benignidad y ninguna de malignidad. Si la tumoración presenta alguna característica de malignidad y ninguna de benignidad, será clasificada como maligna, pero si aparecen características de ambas categorías o ninguna de ellas, el caso se define como inconcluyente y debe ser valorado por un ecografista experto. Las características que definen benignidad son lesión unilocular, presencia de componentes sólidos de diámetro menor a 7mm, presencia de sombra acústica, tumor multilocular de bordes regulares y tamaño menor a 100mm,

y ausencia de captación Doppler (puntuación Doppler 1). Las características que definen malignidad son irregularidad de un tumor sólido, presencia de ascitis, presencia de al menos cuatro estructuras papilares, tumor sólido multilocular de bordes irregulares y tamaño mayor a 100mm, captación Doppler intensa (puntuación Doppler 4) según Timmerman y cols. (2008) [55].

Los modelos de regresión logística hacen un cálculo matemático utilizando la edad de la paciente, el diámetro máximo del componente sólido hasta un máximo de 50mm, y la presencia de características ecográficas, ponderadas, asignando el valor numérico 1 a cada una si están presentes, tales como presencia de ascitis, papilas vascularizadas, irregularidad en las paredes y sombra acústica. El resultado del cálculo se ofrece en porcentaje, fijando como sospecha de malignidad cuando supera el 10 %, según Huerta Sáenz (2019) [59].

En el año 2014, este mismo grupo IOTA desarrolló un nuevo modelo predictivo denominado ADNEX, *Assessment of Different Neoplasias in the Adnexa*, publicado por Van Calster y cols. (2014) [60] y posteriormente, se realizó una prueba de validación externa, publicada por Sayasneh y cols. (2016) [61]. Este modelo, dado el gran número de pacientes participantes (más de 5.000), además de determinar la probabilidad de que se tratase de un tumor benigno o maligno, aporta la posibilidad de predecir que se tratase de una lesión borderline, tumor metastásico, cáncer de ovario invasor en estadio I o en estadios II a IV de la clasificación vigente de la Federación Internacional de Ginecología y Obstetricia (FIGO), superando en fiabilidad a los modelos anteriormente descritos, con una eficacia dentro del grupo de pacientes estudiadas muy alta. La sensibilidad publicada para la detección global de malignidad es del 98 % y 96,4 % dependiendo del punto de corte seleccionado, y una especificidad global del 61,1 % y 73,2 % respectivamente para los distintos puntos de corte, con un área bajo

la curva global de 0,94, pero este método ha sido diseñado para ser aplicado únicamente en pacientes que serán operadas para valorar la necesidad o no de abordaje oncológico, pero no se ha diseñado para el manejo expectante de las lesiones [60].

Al valorar la aplicabilidad de estos modelos en centros hospitalarios, vemos que todos tienen sus diferencias, no siendo generalizables ninguno de ellos. Como ejemplo, el modelo LR2 tiene en cuenta la presencia de vascularización en las tumoraciones, por lo que requiere el empleo de ecógrafos equipados con Doppler. El modelo ADNEX, por el contrario, no requiere del uso del Doppler, pero sí de la determinación del nivel sérico de Ca 125, al igual que el RMI, por lo que tras la detección ecográfica de una lesión, habitualmente será necesario realizar un análisis de sangre a la paciente y esperar a tener dicho resultado, de forma que tendremos que realizar un cálculo en diferido en otro momento, cuando dispongamos de los datos necesarios para ello. Aunque el modelo ADNEX puede ofrecer una predicción sobre el tipo histológico de la lesión, no parece que sea una información de tanta relevancia de forma prequirúrgica para el clínico, pues con la información binaria de benignidad o malignidad, es suficiente para plantear el tipo de abordaje quirúrgico y los recursos necesarios para atender a la paciente de forma óptima, sabiendo así qué pacientes se beneficiarían de una aproximación oncológica y cuáles no. Otro factor a tener en cuenta es que el modelo ADNEX fue desarrollado sobre una población en concreto, y aún requiere la realización de más estudios de validación externa como el publicado por Sayasneh (2016) [61] que demuestren su aplicabilidad en distintos centros y con ecografías realizadas por especialistas de distinto nivel de capacitación, así como con diferentes equipos ecográficos.

1.3. Cáncer de Cérvix

1.3.1. Incidencia y mortalidad mundial del Cáncer de Cérvix

El cáncer de cérvix es uno de los cánceres más frecuentes e incapacitantes en el mundo, siendo la cuarta neoplasia en frecuencia en mujeres, con 570.000 nuevos casos diagnosticados en 2018, representando así el 3,2 % de todos los diagnósticos de cáncer y el 6,6 % de todos los cánceres femeninos. Es el cuarto cáncer diagnosticado en el mundo en mujeres, el séptimo si tenemos en cuenta ambos sexos, sexto cáncer en frecuencia en Europa, y es el cáncer más frecuente en 28 países en vías de desarrollo. El 83 % de los casos acontecen en países en vías de desarrollo, donde tiene una mortalidad cercana al 90 %, diagnosticando el 50 % de los casos en mujeres menores de 35 años, datos publicados por GLOBOCAN (2018) [14], Organización Mundial de la Salud (2018) [62] y Bray y cols. (2018) [15].

Esta diferencia de incidencia y mortalidad entre unos países y otros se fundamenta en las diferencias existentes entre los servicios sanitarios, diferencias sociales y socioeconómicas, pero sobre todo, la clave se encuentra en la presencia o no de estrategias dirigidas a la prevención.

Estas estrategias preventivas se basan en dos pilares: el conocimiento de la historia natural del desarrollo del cáncer de cérvix con sus lesiones precursoras, y la existencia de métodos diagnósticos y terapéuticos dirigidos a evitarlo. Afortunadamente, disponemos de una herramienta de screening efectiva, incluida en algunos programas de screening, que ha demostrado disminuir la incidencia y mortalidad del cáncer de cérvix en países en los que estos programas de screening se encuentran implementados, según presentan Hakama y cols. (1976) [63]

y Sasloy y cols. (2012) [64]. Estos programas consisten en identificar individuos sanos en riesgo de desarrollar cáncer de cérvix, mediante la detección de lesiones premalignas en estadios subclínicos según Hakama y cols. (1976) [63], lo que conlleva una supervivencia del 92 % a 5 años según Saslow y cols. (2012) [64], dado que aumenta el diagnóstico de lesiones preinvasoras, de forma que se reduce la incidencia y la mortalidad producida por el cáncer de cérvix, según publica García Asenjo y cols. (2013) [65]. Por ello, este cáncer es especialmente problemático en aquellos países que no tienen implantados estos programas de screening poblacional, puesto que en ellos, la enfermedad se diagnostica en estadios avanzados pues es cuando suelen aparecer los síntomas, tales como dolor o sangrado genital, cuando los tratamientos necesarios son muy agresivos y no se encuentran disponibles en todos los países, por lo que la tasa de supervivencia es muy baja y se asocia habitualmente a graves secuelas.

En Reino Unido se implantó un programa de cribado poblacional en 1998, y se estima que gracias a él se ha evitado una muerte por cáncer de cérvix por cada 65 mujeres nacidas desde 1950. En Australia se implantó el programa de cribado en 1991, consiguiendo dar una cobertura poblacional con el cribado cercana al 90 %, mostrando un descenso en la incidencia y mortalidad de un 13,2 y 4 inicial respectivamente, a un 6,9 y 1,9 actual por cada 100.000 habitantes, tal y como refleja Sáez y cols. (2013) [66].

Según los datos publicados en 2004 por la *International Agency for Research on Cancer* (IARC, 2004) [67], el cáncer de cérvix supuso el 7,5 % de todas las muertes por cáncer en el mundo. Sin embargo, según los datos publicados por GLOBOCAN (2018) [14], el número de muertes estimadas por esta causa es de 311.365, lo que supone un 3,3 % del total de fallecimientos por cáncer, lo que parece mos-

trar una tendencia descendente.

1.3.2. Incidencia y mortalidad en España del Cáncer de Cérvix

La incidencia en España es de 6,3 casos por 100.000 mujeres al año, según GLOBOCAN (2018) y la Organización Mundial de la Salud (2018) [14, 62].

El cribado de cáncer de cérvix en España no está incluido dentro de un programa de screening poblacional como se realiza para la detección precoz de otros cánceres como el de mama, sino que se trata de una estrategia de cribado oportunista. Esta estrategia difiere de los programas poblacionales en que no se realiza ningún test de detección a toda la población, sino que sólo a aquellos individuos que acceden a los servicios de salud por cualquier motivo, se les ofrece realizar las pruebas de detección indicadas. Según Pérez Gómez y cols. (2010) [68], la incidencia de cáncer cervical en España parece mostrar una tendencia descendente, aunque sin significación estadística, sin un demostrado aumento de la incidencia desde 1997. Sin embargo, el ligero aumento de incidencia en el segmento de población de mujeres menores de 45 años que se ha demostrado en los últimos años, pone de manifiesto que la estrategia de screening oportunista no resulta eficaz para reducir la incidencia del cáncer de cérvix, según afirma Sáez y cols. (2013) [66], siendo este el motivo por el que las Sociedades Científicas reclaman la instauración de estrategias de cribado poblacional para el cáncer de cérvix.

La incidencia en España se encuentra dentro de los países con menor incidencia de cáncer de cérvix de Europa, con cifras similares a Holanda y Grecia, ligeramente superior a Finlandia y Malta, tal como publica García Asenjo

y cols. (2013) [65]. Según las previsiones de la Red Española de Registros de Cáncer (REDECAN), publicado por Badía y cols. (2019) [69], para 2019 los tumores más prevalentes en mujeres serán por orden de frecuencia mama, colorrectal, pulmón, siendo el cáncer de cérvix uno de los menos frecuentes, con una estimación de 6.682 casos.

Respecto a la mortalidad en España, se encuentra entre las más bajas de Europa, junto con Holanda, Italia y Grecia. Esta baja mortalidad, sería esperable que disminuyera en el futuro, junto con la incidencia, debido la inclusión de la vacunación frente al Virus del Papiloma Humano (VPH), incluida en el calendario vacunal desde 2008, dado que este virus es el agente etiológico de la variedad escamosa del cáncer de cérvix, el tipo predominante de cáncer cervical, según García Asenjo y cols. (2013) [65].

1.3.3. Antecedentes de la detección precoz del Cáncer de Cérvix

En el cérvix podemos encontrar distintos tipos histológicos de cáncer, siendo el carcinoma epidermoide el tipo más frecuente.

El conocimiento de las lesiones precursoras del cáncer de cérvix se inicia con las primeras descripciones histológicas del carcinoma intraepitelial *in situ* realizadas en 1886 por Williams, según recoge Kolstad (1970) [70] y en 1900 por Cullen, según Younge (1965) [71]. En 1910, Rubin, que desconocía las descripciones de esta lesión anteriormente publicadas, describió también un caso de carcinoma *in situ*, concluyendo que esta neoplasia epitelial superficial era el estadio previo necesario para desarrollar un carcinoma escamoso cervical, según Younge (1965) [71]. Posteriormente, se realizaron diversos trabajos dirigidos a confirmar estas descripciones y afirmaciones, de entre los que destacan los rea-

lizados por Schottlaender y Kermauner en 1912 donde identifican lesiones de carcinoma *in situ* en la periferia de carcinomas cervicales invasores, siendo uno de sus discípulos, Schiller, quien comunicó no sólo estudios histológicos detallados, sino que en 1927 también ideó la técnica de tinción cervical con lugol para su detección *in vivo*, que fue denominada como test de Schiller, y que captó la atención de la Ginecología dado que gracias a ella se podía detectar de forma relativamente sencilla en las pacientes este tipo de lesiones durante la exploración visual cervical, según Younge (1965) [71]. El test de Schiller consiste en impregnar con lugol el epitelio cervical externo, basándose en la avidéz de los polisacáridos como el glucógeno presente en el interior de las células por este colorante, de forma que las células que están en estado de reposo, es decir, sanas se tiñen de color marrón, mientras que las células cuyo estado metabólico es más activo y degradan el glucógeno no se tiñen y dejan el área donde se encuentran de color blanquecino, sin teñir, siendo denominadas estas zonas como áreas Schiller positivas.

En 1924, Hinselmann desarrolla en Alemania la colposcopia, según recoge Muñoz (1972) [72], inicialmente ideada para magnificar la visualización del cérvix mediante la utilización de lentes de aumento para tomar biopsias con mayor exactitud, a la que se añadió la tinción del epitelio con ácido acético. El ácido acético tiene la propiedad de inducir una coagulación de las proteínas intracelulares de forma transitoria, con lo que las células en las que haya una alta replicación viral sufrirán esta coagulación reversible y se mostrará la zona afectada como un área blanca, denominada epitelio acetoblancos. Esta técnica, que hoy en día es fundamental para el estudio y seguimiento de pacientes con lesiones cervicales en consulta, inicialmente fue adoptada únicamente en Alemania y Suiza.

Gracias a la adopción de la colposcopia y las biopsias cervicales para la detección del cáncer cervical, las lesiones precursoras son bien conocidas en la actualidad, siendo su detección precoz el objetivo de los programas de screening, tal y como indican Luthra y cols. (1987) [73] y Holowaty y cols. (1999) [74]. Estas lesiones se originan a partir de la infección de un virus, el Virus del Papiloma Humano cuya presencia fue revelada por zur Hausen en 1974 mediante la detección del ADN viral en las muestras de cáncer de cérvix, implicándolo así en la génesis de estas lesiones, siendo galardonado con el Premio Nobel de Medicina en 2008 por este descubrimiento, según Cuestas (2008) [75]. El VPH entra en las células basales de la zona de transformación e integra su ADN en el de las células, alterando así la capacidad de división y maduración de las mismas, lo que morfológicamente se traduce en la aparición de atipias celulares. Estas lesiones precursoras con atipias, inicialmente no son consideradas aún como un cáncer, dado que no tienen capacidad de infiltración de la lámina basal del epitelio ni de diseminación a distancia.

1.3.4. Lesiones precursoras del Cáncer de Cérvix y su detección

Las lesiones precursoras son aquellas que, dejadas a su libre evolución, tienen riesgo de progresar a carcinoma invasor de cérvix. La incidencia de estas lesiones, en pacientes sometidas a screening mediante citología oscila entre el 1,1 y el 3,7 %, según Anttila y cols. (2010) [76].

En 1970 la Organización Mundial de la Salud (OMS) decidió unificar la terminología para definir estos procesos y así facilitar la comunicación entre profesionales y la correlación entre los resultados de alteraciones citológicas e his-

tológicas. Así, se propuso denominar a estas lesiones precursoras como “lesión escamosa intraepitelial”, en inglés *Squamous Intraepithelial lesion* (SIL), graduándola en tres categorías, SIL I, SIL II y SIL III. La categoría SIL III incluye displasias severas y carcinoma *in situ*.

En 1989, el Instituto Nacional del Cáncer de Estados Unidos, en su reunión de Bethesda, modificó la clasificación de las alteraciones celulares en las citologías teniendo en cuenta al VPH como agente etiológico común, según Nayar y Wilbur (2015) [77] y Pelea y González (2003) [78]. Con esta modificación, simplifican las categorías en dos grados, por un lado las lesiones de bajo grado o L-SIL, que incluye las lesiones SIL I previas y condilomas, y por otro lado, las lesiones de alto grado o H-SIL, que agrupan a las SIL II y III de la clasificación anterior. Esta clasificación tiene un mayor significado clínico, dado que las lesiones de bajo grado presentan tendencia a ser autolimitadas y regresar de forma espontánea, mientras que las de alto grado son lesiones consideradas premalignas con mayor riesgo de progresión a cáncer cervical. Cuando las anomalías celulares son de un carácter más marcado que los cambios reactivos pero no lo suficiente como para ser clasificadas como lesión intraepitelial, son clasificados como Atipias Celulares Escamosas de Significado Incierto, en inglés *Atypical Squamous Cells of Uncertain Significance* (ASC-US), según refieren Solomon y cols. (2002) [79], el *National Cancer Institute* (2002) [80] y Nayar y Wilbur (2015) [77].

Debido a que las lesiones escamosas intraepiteliales causadas por el VPH son idénticas en todos los epitelios del tracto genital inferior tanto en varones como en mujeres, en 2012 se acordó la creación de una nomenclatura histopatológica común para este tipo de lesiones, denominado Sistema LAST, del inglés *Lower Anogenital Squamous Terminology*. Según este sistema, las lesiones se denominan

Lesión Escamosa Intraepitelial (SIL). De forma análoga a como lo hace el sistema Bethesda, se dividen en L-SIL para bajo grado y H-SIL para alto grado. Para clasificar las lesiones histológicas de biopsias, este sistema utiliza la sigla de la zona afectada seguido de la terminación "IN", como neoplasia intraepitelial, del inglés *Intraepithelial Neoplasm*, y el grado 1, 2 o 3 según corresponda, de forma que para una lesión de bajo grado en cérvix su denominación sería CIN 1, para vagina VAIN 1, en vulva VIN 1 y en ano AIN 1, según publica Darragh y cols. (2012) [81].

En 2014 se actualizó la clasificación del sistema Bethesda, con cambios mínimos respecto a la revisión anterior, destacando que en los informes debe figurar la presencia de células endometriales de apariencia benigna en mujeres a partir de los 45 años de edad, para aumentar el valor predictivo de esta categoría, tal como indican Nayar y Wilbur (2015) [77].

Las lesiones cervicales intraepiteliales de bajo grado tienen tendencia a resolverse de forma espontánea por norma general, gracias a la competencia del sistema inmunitario de las pacientes, pero dependiendo de las características de la mujer tales como la edad, inicio precoz de relaciones sexuales, número alto de parejas sexuales, hábito tabáquico, paridad o inmunodeficiencias, así como del genotipo del virus implicado y la coexistencia con otras infecciones, en algunos casos, la infección puede hacerse persistente y las lesiones pueden progresar en mayor o menor tiempo, generalmente en varios años, hacia lesiones de mayor grado que pueden llevar al desarrollo de un cáncer de cérvix invasor. En la actualidad, desde las primeras décadas del siglo XX, sabemos que existe un virus de transmisión sexual responsable del desarrollo de este cáncer, el VPH, con distintos serotipos y diferentes comportamientos en función del sero-

tipo. La secuenciación del genoma de este virus así como el estudio intensivo de su estructura en la década de los años 80 ha permitido identificar los serotipos de alto riesgo oncogénico, lo que recientemente ha llevado al desarrollo y comercialización de vacunas frente a este agente etiológico, según publican la IARC (2007) [82] y McKeage y Lyseng-Williamson (2016) [83]. La infección por el VPH es factor indispensable para desarrollar la práctica totalidad de lesiones cervicales precancerosas, que finalmente pueden llevar a cáncer cervical en el transcurso de varios años o décadas si se dan las circunstancias pertinentes, tal y como detalla Hartwig y cols. (2015) [84], estando también asociado al desarrollo de cáncer en otras localizaciones anogenitales o incluso en laringe.

Didácticamente, el epitelio cervical escamoso puede ser dividido en tres tercios, desde la lámina basal hasta la superficie externa. La primera etapa en la que el VPH desarrolla una lesión cervical es conocida como CIN grado I (CIN I), en la que podemos encontrar una lesión displásica localizada en el tercio basal del epitelio. La segunda etapa, conocida como CIN grado II (CIN II) es aquella en la que tanto el tercio basal como el tercio medio se encuentran afectados por esta lesión displásica. La tercera etapa, conocida como CIN grado III (CIN III), es en la que la lesión displásica sobrepasa los tercios basal y medio, siendo sinónimo de cáncer intraepitelial, pudiendo desarrollar posteriormente un cáncer invasor mediante la infiltración del estroma subyacente, como se muestra en la **Figura 1.7**, tomada de Carrasco Colomer (2008) [2]. Estas etapas histológicas se asocian con cambios celulares característicos y definitorios de ellas, que pueden ser detectados mediante el estudio de células aisladas, obtenidas mediante muestreo cervical.

La técnica pionera en la detección precoz del cáncer cervical fue el estudio

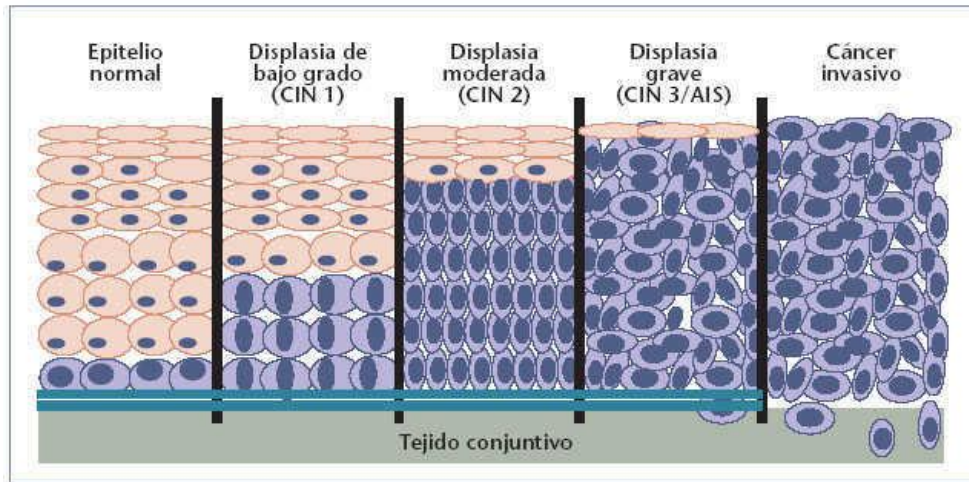


Figura 1.7: Esquema del desarrollo de las lesiones cervicales, de izquierda a derecha, desde epitelio normal hasta cáncer invasivo, pasando por CIN I, II y III. tomada de Carrasco Colomer (2008) [2].

de las muestras citológicas cervicales mediante la técnica descrita en 1928 por Papanicolaou (1941) [10]. Inicialmente, esta técnica fue desarrollada como un método sencillo de conservación y análisis para detectar cambios celulares que pudieran conducir al desarrollo del cáncer cervical. Tras su publicación, no fue bien aceptada por la comunidad científica, debido, en parte, a que no existía un conocimiento real de la historia natural de este cáncer y los cambios celulares iniciales asociados a él, pero gracias al mejor conocimiento de los factores participantes en el desarrollo de la enfermedad, esta técnica se ha revelado como la mejor prueba de screening poblacional disponible por el momento.

El análisis de la muestra citológica se podría resumir como un proceso subjetivo, en el que un patólogo experto examina cada muestra con atención, y las clasifica como normal, premaligna o maligna. Dado que se trata de un proceso de clasificación humana subjetiva, presenta, inherentes a él, un número de falsos

negativos y falsos positivos, dependiendo de la calidad del muestreo, errores o artefactos de procesamiento, fatiga u otras situaciones, además de presentar una alta variación interpersonal, con un índice kappa bajo, tal y como publican Ismail y cols. (1989) [85], Stoler y cols. (2001) [86] y más recientemente Bigras y cols. (2013) [87].

En la actualidad, la sensibilidad del screening de las citologías cervicovaginales estudiadas por patólogos expertos varía entre las publicaciones consultadas, pero oscila entre el 68,8 % y el 93,8 %, con una media del 77,4 %, y una especificidad que varía desde el 70,6 % hasta el 95,6 %, con una media del 81,3 %, con una fiabilidad en la detección de CIN II y grados superiores que oscila entre el 74,1 % y el 83,8 %, con una media del 79,4 %, como reflejan distintos autores como Arbyn y cols. (2004)[88], Bigras y cols. (2013) [87], Wright y cols. (2014) [89] y Sorbye y cols. (2017) [90]. Esta variación de la sensibilidad y especificidad se relaciona en parte con la naturaleza subjetiva del proceso de clasificación así como con los errores tanto técnicos como humanos que pueden surgir durante todo el proceso, desde la toma de la muestra, su procesamiento y su análisis al microscopio. La toma citológica clásica consistía en la toma de material exfoliativo exocervical mediante la utilización de una espátula, como la expuesta en la **Figura 1.8a**, realizando un muestreo de los fondos de saco laterocervicales y del exocérnix propiamente dicho, y posteriormente se tomaba una muestra mediante cepillado del canal endocervical, extendiendo cada una de las tres muestras en un cristal portaobjetos y fijando la muestra mediante fijación química utilizando para ello lacas especiales. Para disminuir los errores que pueden ocurrir durante este muestreo y la aparición de artefactos durante el proceso de fijación, se han desarrollado varios sistemas que simplifican este procedimiento, como la toma de la muestra para el estudio mediante un único cepillado de muestreo

simultáneo exo y endocervical y su conservación en medio líquido, según publican Ronco y cols. (2007) [91] y Arbyn y cols. (2008) [92]. Estas muestras se recogen utilizando un cepillo citológico especial denominado *Liquid-Based Cytology Specimen brush* (LBC), como se presenta en la **Figura 1.8b**, que es lavado en líquido preservativo, tal y como aparece en la **Figura 1.8c**. Este líquido se centrifuga y se filtra, situando las células sobre un portaobjetos de cristal formando una monocapa uniforme, que será teñida, y para reducir la concurrencia de posibles errores durante el procesamiento de la muestra, se puede utilizar un sistema automatizado y robotizado de preparación de las muestras, que puede aportar una reducción del tiempo necesario para fijar y teñir cada muestra, evitando posibles errores derivados del manejo humano de las mismas.

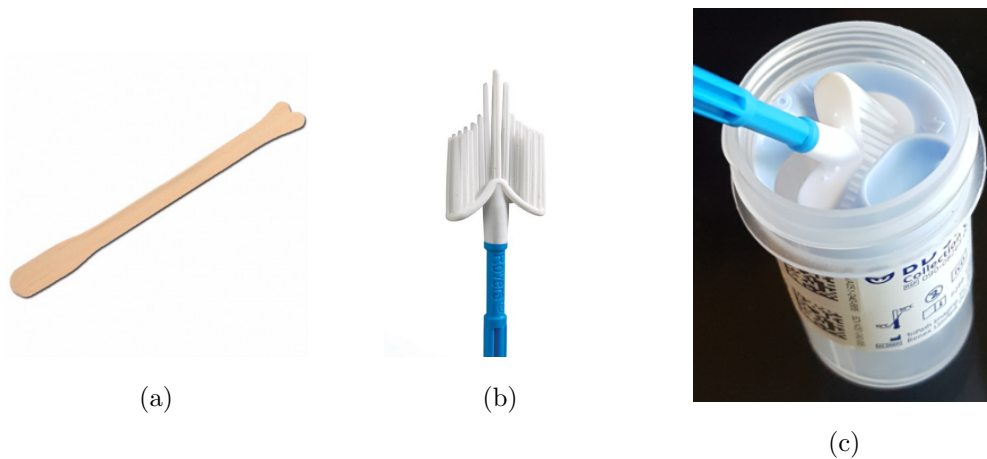


Figura 1.8: Material utilizado habitualmente para la toma de muestras cervicovaginales para realizar citologías (a) Espátula de Ayre, (b) cepillo citológico LBC, y (c) lavado del cepillo en líquido preservativo BD Surepath®.

Para la tinción de las muestras se utilizan distintos colorantes, según Montuenga y cols. (2009) [93] e instrucciones de uso para los reactivos ITW Reagents de Panreac Química S.L.U., AppliChem (versión 3: JMBAPR2009):

- Hematoxilina: Este colorante es ampliamente utilizado en Anatomía Patológica y en estudios histológicos dado que tiene afinidad por los núcleos celulares, permitiendo valorar su estructura. Habitualmente, se utiliza el pigmento denominado Hematoxilina de Harris.
- Orange G: Este colorante sintético es de carácter ácido, por lo que tiene afinidad por los compuestos de naturaleza básica como la prequeratina, que se tiñe de color rosado, o la queratina que adquiere un color naranja brillante.
- Eosina amarillenta: Este colorante tiene la propiedad de teñir el citoplasma de las células escamosas maduras, de las células ciliadas y de los hematíes de color rosa-anaranjado.
- Verde Luz SF amarillento: Con este colorante las células escamosas no superficiales se tiñen de color verde azulado, por lo que se muestra en las células inmaduras o parcialmente maduras.
- Pardo Bismark R: Este colorante tiñe la mucina, pero no el citoplasma celular, por lo que resalta células productoras de mucina, típicas del epitelio glandular endocervical y endometrial.
- Ácido fosfotúngstico: Esta sustancia no es un colorante, sino que se trata de un mordiente. La función de este ácido consiste en aumentar la afinidad de las células por el colorante, en este caso, consigue su efecto sobre las células escamosas inmaduras y parcialmente maduras, consiguiendo una coloración más marcada en las mismas.

Dado que la tinción utiliza varios colorantes, permite identificar distintos tipos de células, por lo que se trata de una tinción ideal para estudios de tipo citológico. Para su aplicación, se utilizan tres soluciones diferentes, para agrupar los colorantes, que se aplican de forma correlativa siguiendo las instrucciones que a continuación se detallan:

- Solución de Hematoxilina.
- Solución de Papanicolaou OG, que contiene el Orange G.
- Solución de Papanicolaou EA, que contiene el resto de colorantes.

El cristal portaobjetos debe ser sumergido durante un minuto, y de forma sucesiva, en alcohol 80 %, alcohol 70 %, alcohol 50 % y agua. Posteriormente, se sumerge en la solución de hematoxilina de Harris durante 5 minutos, y se retira el exceso de tinción mediante la inmersión del cristal en agua 6 veces durante un segundo cada una.

Posteriormente se sumerge el cristal en ácido clorhídrico en dilución 0,5 %, 8 veces durante un segundo. A continuación se lava el cristal con agua corriente durante 5 minutos, y se vuelve a pasar la muestra por alcoholes aumentando la graduación de forma sucesiva de 50 %, 70 %, 80 % y 96 % durante 30 segundos en cada uno de ellos.

Tras este paso por alcoholes, se realiza la tinción con Solución de Papanicolaou OG durante un minuto o minuto y medio. El exceso de colorante se elimina mediante dos baños en alcohol al 96 %, sumergiendo la preparación en dos ocasiones durante 3 o 4 segundos cada vez.

La siguiente tinción con la solución de Papanicolaou EA se realiza mediante la inmersión en la misma durante un minuto y medio o dos, y posteriormente se lava la preparación sumergiéndola dos veces en tres recipientes distintos con etanol al 96 % durante 3 o 4 segundos en cada ocasión. Tras retirar el exceso de colorante de esta forma, se lava la preparación con etanol absoluto durante 30 segundos.

Para finalizar, se sumerge la preparación durante 4 minutos en un baño 1:1 de xileno, mezcla de isómeros y etanol absoluto, tras lo que se aclara con xileno, mezcla de isómeros, mediante inmersión de la preparación durante 3 minutos en baño.

1.3.5. Análisis automático de citologías cervicovaginales

Tras la tinción, las muestras aún deben ser examinadas por patólogos o citotécnicos humanos expertos, pues los sistemas de análisis automatizado aplicados a la clasificación de citologías cervicovaginales, por el momento, no han demostrado ofrecer mejores resultados que el análisis por un experto para la detección de lesiones de grado CIN II o mayor, independientemente del sistema automático utilizado o el tipo de muestra examinada (convencional o líquida), según Alves y cols. (2004) [94]. Los análisis de coste-efectividad muestran que los sistemas automatizados son más caros que el examen manual, cuando son aplicados a programas de screening poblacional como se demuestra en distintas revisiones como las publicadas por el *Medical Services Advisory C.* (2009) [95] y el *Scottish Cervical Cytology Review Group* (2009) [96].

La única ventaja publicada de la aplicación del análisis automatizado, según Kitchener y cols. (2011) [97] y Boost (2009) [98], es que su eficiencia es mayor que la del análisis manual, en términos de número de muestras analizadas por unidad de tiempo.

Los sistemas de análisis automático para la clasificación de citologías cervicovaginales han sido diseñados para ser aplicados en citologías procesadas en monocapa, según Boost (2009) [98]. Estas muestras se toman utilizando un cepillo citológico especial, que es lavado en líquido preservativo. Este líquido se filtra y las células se sitúan sobre un portaobjetos de cristal formando una monocapa uniforme, que será teñida. Tras este procesamiento, se escanea el portaobjetos completo, y el sistema determina la densidad óptica de los núcleos celulares en 120 campos visuales, mostrando de 15 a 20 campos con mayor atipia (determinada por la densidad óptica nuclear) para ser revisados. Estos sistemas

reducen el tiempo de trabajo, simplificando el análisis de todo el portaobjetos a los campos visuales más representativos, lo que también permite evitar la revisión de las muestras normales, según Queiro Verdes y cols. (2013) [99].

Los sistemas pioneros en este campo fueron aprobados por la Administración Americana de Alimentos y Medicamentos, la *Food and Drugs administration* (FDA) en la década de los años 90, y desde entonces, muchos otros sistemas han intentado mejorar los resultados, como reflejan Koss y cols. (1994) [100] y Patten y cols. (1997) [101], pero la mayoría de ellos no están disponibles hoy en día por los motivos anteriormente expuestos.

En España, disponemos de dos sistemas comercializados: el BD FocalPoint GS Imaging System® (BD Diagnostics-TriPath, Burlington, NC. USA), y el ThinPrep Imaging System® (Hologic, Bedford, MA, USA).

El BD FocalPoint Slide Profiler® escanea el portaobjetos y le asigna un valor numérico en quintiles, dependiendo de la intensidad de la atipia celular detectada. El resultado viene dado en cuatro categorías:

1. Revisar: Quintil del 1 al 5;
2. No se requiere revisión: la muestra, entonces se encuentra dentro de los límites de la normalidad;
3. Revisar proceso: Problema técnico;
4. Revisar para control de calidad: Revisar todas las imágenes de la pantalla.

El ThinPrep Imaging System® está diseñado para ser aplicado sobre citologías tomadas en medio líquido, y teñidas con su método de tinción comercializa-

do Hologic Imager®), según según Queiro Verdes y cols. (2013) [99]. El sistema escanea el portaobjetos completo y selecciona 22 campos visuales para ser analizados, pero este sistema requiere un entrenamiento previo. Si este sistema encuentra cualquier anomalía en alguno de estos campos, entonces, toda la muestra debe ser revisada. Este clasificador no asigna ninguna categoría diagnóstica a la muestra, pero permite al patólogo evitar la revisión de muestras catalogadas como normales.

Actualmente, ninguno de estos sistemas está recomendado por el Programa Europeo contra el Cáncer (*Europe Against Cancer Programme*) para su aplicación en programas de screening poblacional, debido a la falta de evidencia requerida para la elaboración de guías clínicas, según publica Arbyn y cols. (2010) [102].

1.4. Inicios de la Inteligencia Artificial

En las últimas décadas del siglo XX y primeras del XXI, la tecnología ha evolucionado para dar respuesta a problemas derivados de la necesidad de gestionar grandes volúmenes de información, derivando así en el desarrollo de sistemas denominados genéricamente como Inteligencia Artificial (IA), y concretamente, como técnicas de Aprendizaje Automático, o como también es conocido en inglés, *Machine Learning* (ML). Estos sistemas, inicialmente consistían en algoritmos de estudio probabilístico cuyo objetivo era predecir, con mayor o menor grado de acierto, la posibilidad que un fenómeno tenía o no de ocurrir, en base a unos datos previos y en función de un entrenamiento concreto que ponderaba la relevancia de cada uno de los datos y su influencia en la ocurrencia del fenómeno a estudiar. Por ejemplo, se han desarrollado modelos para realizar tasaciones teóricas de los precios de viviendas con resultados muy cercanos a la

realidad, tomando como valores de referencia ponderados el barrio donde se encuentra la vivienda y los precios de otras viviendas cercanas, la superficie útil de la vivienda, orientación, año de construcción, reformas realizadas, y número de habitaciones, entre otros. De esta forma, el algoritmo toma de cada ejemplo de muestra unas características que serán analizadas y se le entrenará en la asignación de una etiqueta o categoría concreta, que será el objetivo de clasificación propuesto. Este tipo de aprendizaje se denomina Aprendizaje Supervisado, en el que el algoritmo realiza ajustes de cálculo internos con cada ejemplo que le es presentado, ajustando así sus parámetros para obtener una etiqueta de salida adecuada. Cuando el número de ejemplos de muestra presentados sea el adecuado, el algoritmo será capaz de predecir o generalizar, con una probabilidad determinada, cuál será la etiqueta o categoría que corresponda a las siguientes muestras que se le presenten, según Ng (2011) [103].

Como podemos ver, la IA tiene entre sus objetivos conseguir que las máquinas sean capaces de aprender y actuar de forma similar a como lo hace el ser humano, mejorando su capacidad de aprendizaje a lo largo del tiempo y de forma autónoma mediante la incorporación de datos e información procedentes del mundo real.

Hoy día, las técnicas de ML se encuentran presentes en la mayoría de dispositivos electrónicos disponibles, siendo la base de tecnologías tan útiles como la conducción automática de vehículos, el reconocimiento de voz, filtros de búsqueda en internet, filtros de correo anti-spam o reconocimiento de imágenes. Estos ejemplos pueden agruparse generalmente dentro de dos tipos de clasificaciones: binaria o multiclase. En la clasificación binaria solamente es posible clasificar los datos en dos categorías, mientras que en la multiclase los datos deben ser clasificados en una categoría de entre varias posibles, según Ng (2011)

[103].

1.5. Aprendizaje Automático aplicado a la clasificación de tumores ováricos.

Respecto a la clasificación de tumores ováricos en benignos o malignos, debemos reseñar una reciente y valiosa contribución basada en técnicas de ML, publicada por Khazendar y cols. (2014) [104], (2015) [105] y (2016) [106]. En estos trabajos proponen un modelo de Support Vector Machine (SVM) basado en una fusión de niveles de decisión, desarrollado a partir del estudio de los descriptores obtenidos a partir de las imágenes ecográficas de una base de datos de 187 casos de pacientes reales. La fusión de niveles de decisión puede compararse con lo que ocurre cuando se consulta un caso clínico a un grupo de expertos, en este caso concreto, a dos expertos distintos: cada uno analiza la situación de forma independiente y exponen su conclusión, de forma que si ambos coinciden en el mismo diagnóstico, esa decisión será aceptada como válida, pero si difieren en su conclusión, surgirá un problema de clasificación para el que puede haber distintas vías de solución.

Para desarrollar esta herramienta, el equipo de Khazendar analizó dos características descriptivas, histogramas en escalas de grises e histogramas de patrones binarios locales, que se extraen mediante el procesamiento de las imágenes originales con la Transformada de Fourier (en inglés *Fourier Transform* - FT). Su estrategia para la clasificación de las características descriptivas se fundamenta en la fusión de niveles de decisión, basándose en dos situaciones principales:

- La primera situación consiste en que ambos análisis clasifican el tumor en el mismo grupo diagnóstico (benigno o maligno), por lo que entonces

se acepta como correctamente clasificado, y el nivel de confianza de esta clasificación depende de cada uno de los niveles de confianza de ambos sistemas.

- La segunda situación consiste en que los sistemas clasifican el tumor en diferentes grupos diagnósticos, es decir, uno lo clasifica como benigno y el otro como maligno, por lo que nos encontramos ante una clasificación inconcluyente, de forma que la imagen no puede ser clasificada por esta herramienta, salvo que una de las características sea clasificada con alto nivel de confianza y la otra con bajo nivel de confianza, por lo que el tumor podría ser clasificado en el grupo diagnóstico de mayor nivel de confianza, pero con un nivel de confianza global bajo.

De esta forma, encontraron que el 18,3 % de sus imágenes no podían ser clasificadas por este sistema, por lo que la exactitud media de su sistema de clasificación utilizando este método de fusión de niveles de decisión fue del 77 %.

Estos trabajos fueron desarrollados mediante técnicas clásicas de ML, concretamente, las utilizadas para el reconocimiento de imágenes, necesitan de un preprocesamiento inicial de las mismas para poder presentar al modelo datos estandarizados en la misma escala. Utilizando para su entrenamiento y evaluación la presentación de las características seleccionadas mediante distintos métodos de extracción de características, en lugar de las imágenes originales propiamente dichas. Esta necesidad de procesamiento previo supone un gasto computacional importante, con la consiguiente inversión de tiempo para cada imagen, además de la dificultad de segmentar los distintos objetos de la imagen. Esta deficiencia, se ve solucionada por un nuevo paradigma que ha surgido en los últimos años: una rama del ML capaz de clasificar las imágenes sin procesamiento previo, denominado Aprendizaje Profundo, o en inglés *Deep Learning*

(DL), según Le Cun y cols. (2015) [107].

1.6. *Deep Learning* y Redes Neuronales Convolucionales

Las técnicas de DL se encuentran presentes en la actualidad en numerosas aplicaciones, aportando resultados muy interesantes para ayudar a resolver problemas cotidianos como plantean Le Cun y cols. (2015) [107] y Jurtz y cols. (2017) [108]. Además, las Redes Neuronales Convolucionales, en inglés *Convolutional Neural Networks* (CNN) son una rama fundamental de las aplicaciones de DL para tratar problemas con imágenes, y están recibiendo una gran atención en los últimos años debido a sus innovadoras aplicaciones según Bernal y cols. (2018) [109]. Específicamente, las CNN son redes profundas que ofrecen grandes resultados para la clasificación de imágenes, como presenta Ahmad y cols. (2018) [110].

Los sistemas automáticos de clasificación han ido evolucionando, hasta desarrollar algoritmos cuyo proceso de entrenamiento se asemeja al aprendizaje intuitivo humano, pudiendo llegar a clasificar directamente imágenes sin necesidad de procesamiento previo, siendo ésta una gran ventaja y la primera diferencia con las técnicas clásicas de ML.

El aprendizaje intuitivo humano se basa en la observación de fenómenos y la extracción de las características comunes existentes entre ellos, por ejemplo, cuando se empieza a enseñar a un niño a diferenciar entre un perro y un gato, no se le explica sus diferencias biológicas ni filogenéticas, sino que se le muestra simplemente la imagen de un perro, y la de un gato, y el niño, observa diferencias y semejanzas entre ambas, y su capacidad para reconocer una u otra categoría irá en función de la cantidad de perros y gatos de diferentes razas, ta-

maños y pelajes que haya visto. Durante este proceso de aprendizaje, el cerebro del niño desarrolla sinapsis entre ciertas neuronas y elimina sinapsis de otras, en función de que la clasificación que obtiene sea acertada o no, seleccionando así los circuitos neuronales que mejores resultados clasificatorios obtienen, en función de una característica cerebral que se denomina plasticidad neuronal, capacidad que habitualmente va disminuyendo a la par que envejecemos y que se asemeja al procedimiento de “poda” de neuronas que se comentará más adelante. Obviamente, un perro pekinés es diferente a un gran danés, así como un gato de Angora lo es de un gato egipcio, pero el niño será capaz de reconocer cualquier raza de perro como un perro y cualquier raza de gato como un gato si ha visto suficientes ejemplos de cada categoría y ha podido seleccionar un circuito neuronal eficaz. Incluso sería capaz de reconocer correctamente una raza de perro o gato que nunca hubiera visto antes, pues tendría en cuenta sus características definitorias. Es más, cuando, siendo adultos, pensamos en la categoría perro y la categoría gato, no nos viene a la mente una raza en concreto, sino una idea genérica capaz de englobar a todas las razas que conocemos. Este es el funcionamiento básico del método conocido como el DL, entre cuyos algoritmos destacan para la clasificación y reconocimiento de imágenes las Redes Neuronales Convolucionales.

Como se ha dicho anteriormente, las técnicas de DL no necesitan preprocesado previo de las imágenes, sino que pueden trabajar directamente con las imágenes en bruto. Sin embargo, requieren una base de datos con un volumen importante para poder realizar el entrenamiento del sistema.

El DL es un área del Aprendizaje Automático que utiliza una estructura neuronal artificial jerarquizada siguiendo un diseño similar al cerebro humano, con

nodos neuronales interconectados entre sí de forma comparable a una tela de araña. La primera capa neuronal recibe los datos brutos directamente de la entrada, y los procesa, extrayendo de ellos sus características principales, que son transferidas a la siguiente capa neuronal. Las sucesivas capas neuronales realizan el mismo procedimiento, hasta que las características analizadas alcanzan la capa final, que es donde se realiza la predicción, asignando la etiqueta de salida. El resultado obtenido se compara con el resultado conocido, de forma que el sistema, realizando un análisis inverso, puede aprender cuáles son los factores y características que conducen a un resultado adecuado y puede eliminar las conexiones y las neuronas que lo lleven a una clasificación errónea, manteniendo las que aportan resultados aceptables, realizando así el proceso denominado como poda de neuronas, según Le CUN (2013) [111].

Los métodos clásicos de ML tienen un punto óptimo de entrenamiento, a partir del que comienzan a dar resultados peores si se continúa entrenando el modelo, en relación con un fenómeno denominado sobreentrenamiento. Este fenómeno no ocurre, sin embargo, con los modelos de DL, que continúan ofreciendo buenos resultados aunque se aumente el número de datos disponibles sin llegar a saturarse. Otra diferencia significativa es que el tiempo necesario para realizar el entrenamiento en los modelos de DL es mucho mayor al necesario para entrenar los de ML, debido a la presencia de muchos más parámetros para considerar en los DL, pero, esta aparente desventaja se invierte en el tiempo necesario para realizar las pruebas, siendo mucho menor en los modelos de DL que en los de ML.

Las células neuronales, en el ser vivo, requieren de la participación de un estímulo para que sean activadas e inicien su mecanismo de transmisión sinápti-

ca. De forma análoga, las neuronas artificiales de la red requieren de la utilización de una función de activación que las ponga en funcionamiento, pudiendo utilizarse para ello funciones sigmoideas o lineales que asignen un valor de salida a un determinado valor de entrada. Como veremos, en nuestro trabajo utilizamos una función de activación denominada Unidades Lineales Rectificadas, en inglés *Rectified Linear Units* (ReLU), que asigna el valor de salida 0 a todas las entradas con valor inicial negativo, y para las entradas con valor inicial positivo devuelve ese mismo valor como salida, según Agarap (2018) [112].

Los sistemas de Ayuda al Diagnóstico Médico asistidos por Computador, en inglés *Computer Assisted Diagnostic systems* (CAD), están adquiriendo mayor interés en los últimos años entre la comunidad científica, debido al gran desarrollo experimentado por los sistemas basados en métodos de ML, como las ANNs. Estas herramientas pueden ofrecer ventajas sobre el análisis del humano experto, en relación con las limitaciones inherentes al examinador humano, como por ejemplo el efecto deletéreo del cansancio físico y psíquico que conlleva una tarea prolongada en el tiempo, que puede afectar negativamente a los resultados. Por este motivo, en estos momentos hay diversos grupos de investigación desarrollando y mejorando sistemas para aplicaciones médicas multidisciplinarias como los trabajos publicados por García-Laencina y cols. (2014) [113] y (2015) [114]. Otra ventaja que aportan los sistemas CAD es que pueden ser aplicados en el mismo momento de la valoración inicial del caso clínico, no requiriendo por ello la realización de nuevas visitas de valoración para el mismo problema, pudiendo así aportar al clínico una información *in situ* que podría equivaler a una valoración del caso por un Comité de expertos, sirviendo así de ayuda *instantánea*. Esta posibilidad de disponer de una ayuda diagnóstica equivalente a una segunda opinión experta es especialmente valiosa para profesionales que

trabajan en pequeños centros clínicos donde no haya otro especialista a quien consultar, o para aquellos especialistas expertos que no tengan la posibilidad de consultar a otro con experiencia similar en caso de duda diagnóstica, así como en centros con alta carga asistencial en los que duplicar visitas para conseguir un diagnóstico preciso pueda suponer un problema de organización, y también podría ser interesante su utilidad como herramienta de entrenamiento para especialistas en formación.

Las CNN están obteniendo recientemente un gran éxito en la clasificación de imágenes a gran escala en todas las áreas de aplicación en que se utilizan como refleja Goodfellow y cols. (2016) [115], y en lo que a clasificación de imágenes médicas respecta, están creando grandes expectativas pues ofrecen resultados muy interesantes. Como ejemplos recientes destacables podemos citar a Godinez y cols. (2017) [116], donde se aplica una CNN multiescala para la identificación de fenotipos celulares basándose en imágenes de alto contenido celular. Los resultados obtenidos por esta CNN fueron más precisos que los obtenidos mediante otros sistemas conocidos. Moeskops y cols. (2016) [117] presenta un método para la segmentación automática en una serie de Resonancias Magnéticas Nucleares (RMN) de tejido cerebral utilizando una CNN, tomando una RMN como entrada. Los resultados, después de evaluar el método con cinco bases de datos, muestran que consigue segmentaciones muy precisas. En un trabajo publicado en 2017, Sun y cols. (2017) [118] presenta una red completamente convolucional multicanal para realizar la segmentación de tumores hepáticos a partir de imágenes obtenidas mediante Tomografía Computerizada con contraste, en multifase mejorada. Esta CNN se va dividiendo a la vez que se obtienen los datos mejorados en cada fase, y ofrece diferente información de las características patológicas. De esta forma, se entrena una CNN para cada fase

de la imagen de Tomografía Computerizada y es entonces cuando se fusionan las características de alta capa. Esta propuesta es evaluada con dos bases de datos, ofreciendo resultados precisos e interesantes.

También se ha aplicado estos sistemas para la clasificación de imágenes microscópicas, como se plantea Arévalo y cols. (2015) [119] desarrollando un marco de trabajo para la detección de carcinoma basocelular, basado en el aprendizaje no supervisado de descriptores. Para el análisis de imágenes histopatológicas, este marco de trabajo consta de tres etapas:

- (1) aprendizaje de representaciones locales;
- (2) aprendizaje global de representaciones, y
- (3) una interpretación visual.

Es en la segunda etapa donde los autores utilizan una CNN para obtener una visión global de las imágenes.

Esteva y cols. (2017) [120] presentan una CNN desarrollada para detectar melanomas al examinar imágenes dermatoscópicas de lunares. Este modelo de CNN consigue una exactitud similar a la que obtienen dermatólogos expertos. La clasificación de imágenes mediante datos de espectrometría de masa utilizando CNN se detalla en el trabajo publicado por Behrmann y cols. (2017) [121]. Los autores de este trabajo obtienen resultados muy interesantes proponiendo una arquitectura adaptada a su problema, basada en CNN, definiendo así una estrategia para interpretar el análisis de los modelos.

Otra aplicación de CNN para clasificar cáncer de mama se presenta por Mothlagh y cols. (2018) [122]. En este estudio, se desarrolla un modelo de CNN

muy preciso y fiable con el objetivo de clasificar tipos y subtipos de cáncer de mama utilizando imágenes histopatológicas teñidas con hematoxilina-eosina y preparaciones inmunohistoquímicas.

Zhang y cols. (2017) [123] se plantean un modelo de CNN para clasificar células del cérvix uterino. En este caso, los autores realizan un procesamiento previo de las imágenes limpiando y eliminando cualquier artefacto que pudiera dificultar el proceso de clasificación en células normales y atípicas. También realizaron en este trabajo distintas sesiones de preentrenamiento de las redes neuronales con distintas imágenes previamente procesadas. Para evaluar su trabajo, utilizan una base de datos pública y obtienen resultados interesantes mediante la extracción de características de las imágenes. Es de destacar que en este trabajo, todas las células utilizadas tanto en el entrenamiento como en la evaluación se presentan como células aisladas y libres de artefactos y solapamientos, no siendo la situación habitual en la práctica clínica real cuando se examina una citología cervicovaginal, puesto que, como veremos, estas muestras habitualmente se presentan junto a una gran diversidad de estirpes celulares, así como agentes bacterianos saprófitos o patógenos, agregaciones celulares, solapamientos y otros artefactos propios del procesamiento de la muestra.

Respecto al estudio de la clasificación automática de citologías cervicovaginales, debemos resaltar que en la actualidad no hay ningún sistema de clasificación automática que, aplicado a las citologías cervicovaginales, tenga suficiente fiabilidad como para ser recomendado en programas de screening poblacional, según distintas revisiones como las realizadas por Alves y cols. (2014) [94], por el *Medical Services Advisory C.* (2009) [95], por el *Scottish Cervical Cytology Review Group* (2009) [96] o por Queiro Verdes y cols. (2013) [99], posicionamiento adop-

tado por el Ministerio de Sanidad de España basándose en esta última revisión, al igual que hace también el *Europe Against Cancer Programme*. La detección precoz del cáncer de cérvix, y sus lesiones precursoras, es importante para mejorar la supervivencia y reducir la morbilidad asociada a los tratamientos requeridos en estadios avanzados en los que se presenta clínicamente un cáncer de cérvix invasivo.

HIPÓTESIS

2. HIPÓTESIS

Tal y como se deduce de la tendencia actual de la sociedad, el desarrollo tecnológico y la literatura publicada, hay evidencia que sugiere que las nuevas tecnologías concretamente, la Inteligencia Artificial pueden tener cabida entre las herramientas diagnósticas habituales que un profesional de la Medicina puede utilizar en su día a día, tal y como describe García-Laencina y cols. (2014) [113] y (2015) [114].

Respecto a las posibles aplicaciones en el campo de la Ginecología, se ha demostrado la aplicabilidad de algunas herramientas diagnósticas basadas en Aprendizaje Automático, concretamente en el área de diagnóstico por imagen en ecografías de tumores ováricos, como los trabajos de Timmerman y cols. (1999) [56] o los de Khazendar y cols. (2014) [104], (2015) [105] y (2016) [106], con resultados bastante aceptables. En cuanto a la aplicación sobre imágenes microscópicas de técnicas de Aprendizaje Automático, se ha desarrollado estudios en distintos campos que ofrecen resultados interesantes aplicados a la Dermatología sobre todo, como los propuestos por Esteva y cols. (2017) [120] para clasificación de lunares, aunque en el campo de Anatomía Patológica en relación con Ginecología se puede reseñar el método propuesto por Zhang y cols. (2017) [123] para clasificar citologías cervicovaginales mediante Deep Learning,

quien propone un sistema de clasificación aplicado sobre imágenes microscópicas previamente procesadas sobre un fondo limpio, de forma que obtiene buenos resultados, pero su aplicabilidad a muestras reales es dudosa.

De esta forma, parece evidente que las técnicas de Aprendizaje Automático pueden ser aplicables para el desarrollo de herramientas de ayuda al diagnóstico médico asistidas por ordenador, con resultados equiparables a la opinión de un experto.

En este punto se fundamenta la hipótesis de este trabajo de investigación: Los resultados de la clasificación de imágenes ecográficas de tumores ováricos mediante técnicas clásicas de aprendizaje automático, y la clasificación de imágenes microscópicas de citologías cervicovaginales mediante Redes Neuronales Convolucionales pueden ser equiparables a los resultados obtenidos por un observador experimentado.

OBJETIVOS

3. OBJETIVOS

3.1. Objetivo principal

- Aplicar técnicas de Aprendizaje Automático para la clasificación diagnóstica de imágenes médicas en el ámbito de la Ginecología, concretamente de imágenes ecográficas de tumores de ovario e imágenes microscópicas de citologías cervicovaginales teñidas con la técnica de Papanicolaou.

3.2. Objetivos secundarios

- Aplicar técnicas de Aprendizaje Automático en la clasificación de descriptores geométricos obtenidos de imágenes ecográficas de tumores ováricos mediante *Fast Fourier Transform* y compararlas con los resultados obtenidos en otros trabajos con la misma base de datos.
- Crear una base de datos mediante un método de fusión de imágenes para mezclar y solapar las células originales y multiplicar el número de imágenes disponibles para clasificación.
- Aplicar técnicas de Deep Learning en la clasificación de las imágenes microscópicas de la base de datos obtenida mediante fusión de imágenes y comparar los resultados con los obtenidos en la clasificación de la base de datos original.

MATERIAL Y MÉTODOS

4. MATERIAL Y MÉTODOS

Este capítulo se organiza en tres secciones. En la primera sección se describe cada una de las tres bases de datos utilizadas para realizar esta Tesis. En la segunda sección se detalla la configuración de los experimentos de clasificación realizados y en la tercera sección se describe el Hardware y el entorno Software utilizados para realizar los experimentos de clasificación.

4.1. Bases de datos utilizadas

4.1.1. Base de datos de ecografías de tumores ováricos

Para aplicar los métodos clásicos de ML a la clasificación de imágenes ecográficas de tumores de ovario, utilizamos una base de datos de descriptores matemáticos cedida por la Universidad de Buckingham, Reino Unido, extraída a su vez de otra base de datos de imágenes ecográficas, propiedad de la Universidad Católica de Leuven, Bélgica, y que originalmente fue utilizada por el grupo IOTA en diversos trabajos de clasificación de tumores ováricos, por lo que se trata de una colección de imágenes con cierta relevancia y representatividad, que permitiría la comparación de los resultados obtenidos con otros estudios realizados por estos otros autores.

La colección de imágenes fue originalmente recopilada por el grupo IOTA,

y está compuesta por 385 imágenes tomadas en modo B, sin señal Doppler. El trabajo original se encuadró dentro de los estudios de clasificación de tumores de ovario del grupo IOTA, y fue aprobado por el Comité de Ética Asistencial de los Hospitales Universitarios de Leuven (Bélgica), y por cada Comité Ético local de cada centro de reclutamiento que participó en los estudios (Malmö y Lund, Suecia; Londres, Reino Unido; Roma, Udine, Bolonia, Milán, Monza, Napoles y Cagliari, Italia; Lublin, Polonia; Praga, República Checa; Pekin, China; Ontario, Canadá; Genk, Bélgica) como se describe en la publicación de Timmermann y cols. (2010) [57]. Cada paciente que participó en el estudio fue debidamente informada del propósito del estudio inicial y que las imágenes obtenidas así como sus datos clínicos anónimos serían incluidos en un archivo para ser utilizados en futuras investigaciones, y cada una de ellas reflejó su consentimiento a participar en los estudios de forma verbal y por escrito, tras la entrega del correspondiente consentimiento informado.

Todas las imágenes ecográficas fueron tomadas previamente a la realización de la anexectomía y estudio histopatológico del ovario de la tumoración estudiada, independientemente de que la intervención requerida incluyese otras medidas quirúrgicas tales como histerectomía, anexectomía contralateral, cirugía de citorreducción u otras. El periodo comprendido entre la adquisición de la imagen y la intervención quirúrgica, en ningún caso superó los 120 días, por lo que el estudio anatomopatológico se realizó con un lapso de tiempo aceptable como para tomar su diagnóstico como el correspondiente al diagnóstico real de la imagen ecográfica. La colección de imágenes está compuesta por tumores ováricos de distinta naturaleza, tanto benignos como malignos, incluyendo tumores estromales, epiteliales, embrionarios y metastásicos, siendo representativa de todas las estirpes y presentaciones de tumores ováricos existentes.

A modo de ejemplo, la **Figura 4.1** muestra dos imágenes ecográficas de tumores extraídas de la base de datos original, una benigna (**Figura 4.1a**) y otra maligna (**Figura 4.1b**).

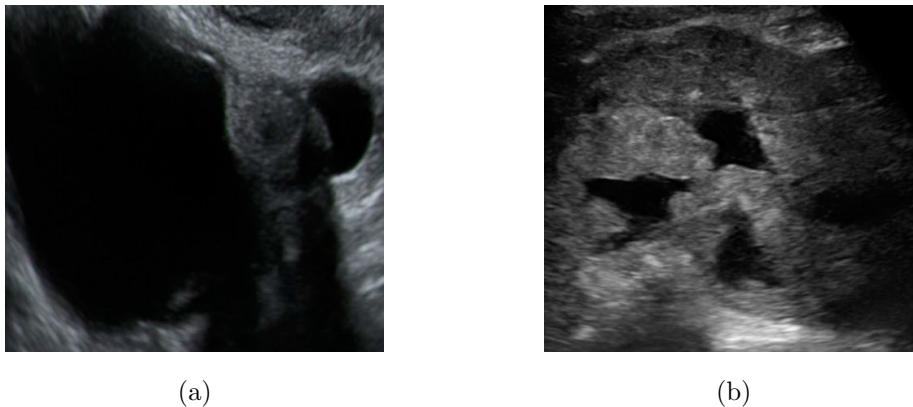


Figura 4.1: Ejemplos de imágenes ecográficas procedentes de la base de datos original creada por el Servicio de Ginecología de la Universidad KU de Leuven (Bélgica), utilizada para la extracción de descriptores: (a) Tumoración ovárica benigna y (b) Tumoración ovárica maligna.

De estas imágenes, el grupo de investigación de la Universidad de Buckingham, Reino Unido, seleccionó 187 imágenes representativas del total, de las que extrajeron las características descriptivas para su análisis mediante Transformada de Fourier, según Khazendar y cols. (2014) [104]. Estas características descriptivas nos fueron cedidas para realizar nuestro trabajo en un archivo completamente anonimizado, en lugar de las imágenes originales, por lo que nuestro estudio, dado que no trabajamos con datos ni imágenes de pacientes, no requirió autorización por Comité de Ética Asistencial. El protocolo del estudio de Khazendar fue aprobado por el Comité de Ética de la Universidad de Bucking-

ham en mayo de 2012, tal y como se refleja en su trabajo, defendido como Tesis Doctoral en 2016 [106].

De las 187 imágenes, 112 imágenes corresponden a tumoraciones benignas y 75 a malignas, de las que se extrajeron dos tipos de vectores para su clasificación, según Khazendar y cols. (2014) [104], (2015) [105] y (2016) [106]: Características o descriptores de Histogramas de Intensidad y Características o descriptores de Patrón Binario Local. Ambos métodos de extracción de imágenes se han realizado siguiendo cuatro procedimientos: Extracción de los descriptores desde la imagen original (*Original image*), Extracción de los descriptores tras preprocesamiento de la imagen con técnicas de mejora de imagen (*Enhanced image*), Extracción de los descriptores de la Región de Interés determinada de forma manual (*Segmented ROI*), y Extracción de los descriptores de la Región de Interés tras técnicas de mejora de la imagen (*Enhanced ROI*).

En el campo del procesamiento de imágenes, la Transformada de Fourier (FT) es una herramienta matemática utilizada para descomponer la onda de frecuencias de una imagen de entrada en sus componentes de seno y coseno, según González y Woods (1992) [124]. La salida obtenida se representa en el dominio de frecuencias (o dominio de Fourier), mientras que la entrada original se encuentra en el dominio espacial. La Transformada de Fourier Rápida (*Fast Fourier Transform*, FFT) es un algoritmo muy eficiente que permite calcular la FT Discreta (*Discrete Fourier Transform*, DFT), pues sólo trabajaremos con imágenes digitales, y su inversa, obteniendo una nueva imagen en el dominio espacial. Debemos remarcar que la FT realiza cálculos muy complejos para descomponer una imagen de entrada, obteniendo una imagen para la parte relativa a magnitudes de frecuencias y otra imagen distinta en relación con la fase de las ondas,

siendo la imagen referente a las magnitudes la que nos ocupa en este trabajo pues es la que contiene mayor información sobre la geometría del dominio espacial.

Como hemos comentado previamente, esta base de datos ha sido utilizada en los trabajos de la Universidad de Buckingham, publicados por Khazendar (2014) [104], (2015) [105] y (2016) [106], donde destacaron que los mejores resultados de clasificación se consiguen al aplicar los algoritmos sobre los descriptores obtenidos mediante *Enhanced ROI*. Siguiendo en esta línea, en nuestro trabajo realizamos la computación sobre la FT y esta información la utilizamos para entrenar y evaluar distintos modelos de clasificación basados en ML para evaluar su influencia en el resultado final del sistema CAD. Los detalles principales de la fase de extracción de características se muestra resumidamente en el siguiente algoritmo:

1. Descomponer una imagen en su dominio de frecuencia mediante la aplicación de FFT. Computación de su espectro de potencia.
2. Binarizar la imagen del espectro de potencia de FFT utilizando un umbral entrenado.
3. Calcular la elipse que mejor encaje en el centro de la imagen del espectro binario.
4. Extraer los diámetros mayor y menor y el área de la elipse en términos de número de píxeles obteniendo un vector de características 3D (*major, minor y area*).

4.1.2. Bases para la extracción de características de FFT:

LA DFT es una herramienta de análisis de señales que analiza y descompone funciones en el dominio tiempo/espacio, tales como las imágenes, en sus distintos componentes de ondas de frecuencias, de la misma forma en que un prisma descompone la luz del sol en sus distintos componentes de frecuencias obteniendo así un arcoiris de distintos colores. La DFT de una imagen f de tamaño MN para cualquier par de frecuencias (u,v) es un número complejo que depende de todos los valores de los píxeles $f(x,y)$ calculado por la fórmula:

$$F(u, v) = \left[\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \cos\left(2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)\right) - \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \sin\left(2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)\right) \right] \quad (4.1)$$

Como la transformación produce números complejos, la salida de la DFT no puede ser mostrada en una única imagen. Sin embargo, la presentación polar de $F(u,v)$ nos ofrece una forma más útil de obtener información de las características de la imagen respecto al espectro de F , definido como el módulo de F , según la siguiente fórmula:

$$\|F(u, v)\| = \sqrt{(\operatorname{Re}(F(u, v)))^2 + (\operatorname{Im}(F(u, v)))^2} \quad (4.2)$$

y su fase:

$$\angle(F(u, v)) = \arctan\left(\frac{\operatorname{Im}(F(u, v))}{\operatorname{Re}(F(u, v))}\right) \quad (4.3)$$

El espectro de Fourier, por sí mismo, ofrece información relativa a la fuerza o intensidad de las características de la imagen, especialmente en lo referente a la dirección de las discontinuidades dominantes en la imagen, como por ejemplo, los bordes u otras texturas geométricas. Estas discontinuidades se reflejan en los rayos más marcados que irradian desde la frecuencia central, localizada en

el punto $(0, 0)$, que representa la energía total de la imagen. Un especialista en Ginecología y Obstetricia sabe que una imagen ecográfica de un tumor ovárico maligno tiene tendencia a contener más detalles y estructuras complejas en comparación con las imágenes, habitualmente, mucho más simples de los tumores benignos. Las imágenes mostradas en la **Figura 4.2**, tomada de Martínez Más y cols. (2019) [3], confirman esta observación, cuando analizamos la silueta del espectro de bajas frecuencias de las características descriptivas obtenidas.

La **Figura 4.2** muestra que las imágenes de tumores malignos tienen tendencia a producir elipses más grandes y anchas que las obtenidas de tumores benignos, en relación con la presencia de mayor número de detalles y estructuras más complejas en las primeras. Cuanto mayor es el número de texturas en la imagen, mayor es el número de estructuras geométricas que contiene, y consecuentemente, habrá mayor concentración de energía en las regiones centrales del espectro. La binarización de la imagen FFT, utilizando un umbral de sensibilidad determinado, produce una mancha “elíptica” en el centro. Los parámetros que caracterizan la elipse que mejor representa la imagen obtenida, como por ejemplo, ejes mayor y menor, área, perímetro y orientación, reflejan la cantidad de energía concentrada en el espectro. Para nuestro análisis, se ha considerado el vector de triple característica denominado Características Geométricas de la transformada Rápida de Fourier, o *Fast Fourier Geometric Feature* (FFGF), formado por el eje mayor, eje menor y área de la elipse, para caracterizar las imágenes de entrada como benignas o malignas. Como se indica en las imágenes de ejemplo de la **Figura 4.2**, las imágenes de tumores benignos tienen menores discontinuidades geométricas, y por ello, la silueta elíptica obtenida de la imagen espectral tiende a aparecer más estrecha con mayor índice entre los ejes mayor/menor. Por otro lado, las imágenes procedentes de tumores malignos tienen muchas más discontinuidades geométricas, y por ello la silueta elíptica en el es-

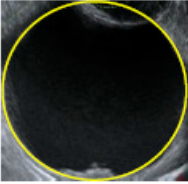
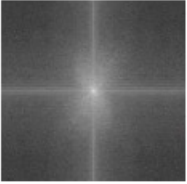
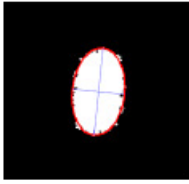
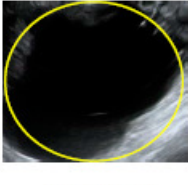
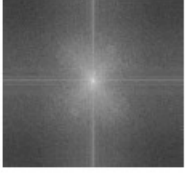
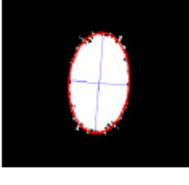
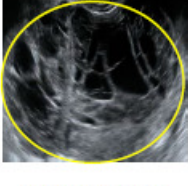
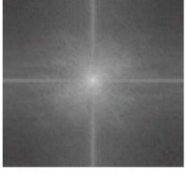
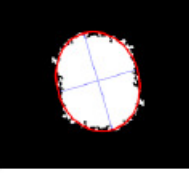
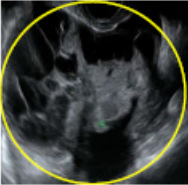
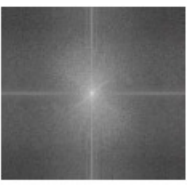
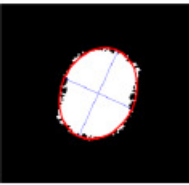
Enhanced ROI	FFT Spectrum	Binary spectrum with best fitted ellipse
 <p>(a) Benign</p>		
 <p>(b) Benign</p>		
 <p>(c) Malignant</p>		
 <p>(d) Malignant</p>		

Figura 4.2: Comparación del espectro de las imágenes ecográficas de tumores ováricos benignos y malignos, tomada de Martínez Más y cols (2019). [3].

pectro tiende a ser más ancha y con menor índice entre los ejes mayor/menor. Además, el tamaño de la silueta elíptica también es un buen indicador de la cantidad de discontinuidades geométricas. Por ello, este conjunto de indicadores podría servir para discriminar entre tumores benignos y tumores malignos.

4.1.3. Base de datos de imágenes microscópicas de células escamosas cervicales.

Recogida de datos

La base de datos que manejamos en este trabajo ha sido creada a partir de un grupo de 10 pacientes a quienes se iba a realizar una intervención denominada conización cervical ambulatoria por diferentes motivos clínicos, tales como presencia de lesión L-SIL persistente, H-SIL o carcinoma in situ, por lo que cada citología tiene su diagnóstico histopatológico obtenido mediante biopsia como control diagnóstico, siendo la biopsia el Gold-Standard. La conización consiste en realizar una exéresis de la porción cervical externa mediante distintas técnicas, con lo que se reseca la zona de transición cervical afectada por la lesión, mediante distintos métodos quirúrgicos, en nuestro caso, se realizó mediante el uso de asa de diatermia con infiltración previa de anestesia local. Las imágenes proceden del Servicio de Anatomía Patológica del Complejo Hospitalario Universitario Santa Lucía-Santa María del Rosell (Cartagena, Murcia), y forman parte del archivo de imágenes clínicas e histopatológicas de dicho Servicio. Todas y cada una de las pacientes cuyas imágenes fueron incluidas en este archivo expresaron su consentimiento informado tanto verbalmente como por escrito, autorizando expresamente a la utilización de su información clínica así como de las imágenes y tejidos obtenidos de su proceso asistencial con fines docentes y de investigación.

El protocolo fue aprobado por el Comité de Ética de la Universidad Católica de Murcia en enero de 2020, con el código CE012005.

La recogida de las muestras para las citologías se realizó mediante cepillado cervical, utilizando el sistema para citologías en medio líquido BD SurePath® liquid-based Pap test®, según se detalla en el capítulo de Introducción, y fueron proce-

sadas utilizando el sistema automatizado y robotizado PrepStain® . Para escanear las imágenes se empleó el iScan Coreo® , un escáner de campo claro para portaobjetos. Cada muestra, con todas las células, se dividió de forma automática en imágenes de tamaño 200 x 200 píxeles, donde cada imagen contendría una célula escamosa en alta resolución, que sería clasificada por dos patólogos, siguiendo el sistema Bethesda 2014 [77]:

- Células escamosas normales: Células epiteliales normales, sin ningún tipo de atipia o significación patológica (**Figura 4.3 a**).
- ASC-US: Anomalías celulares más marcadas que los cambios reactivos pero no tan marcadas como para ser clasificables como lesión intraepitelial (**Figura 4.3 b**).
- L-SIL: Anomalías celulares suaves o moderadas, que reflejan una infección concurrente por VPH (por ejemplo, halo perinuclear o koilocitosis) (**Figura 4.3 c**).
- H-SIL: Atipia celular muy marcada, frecuentemente asociada a infección persistente de VPH y alto riesgo de progresión (por ejemplo, displasia severa, anisocitosis, cromatina granulada, y/o nucleolos grandes) (**Figura 4.3 d**).

Clínicamente, cuando se diagnostica una citología cervicovaginal como ASC-US, L-SIL, o un H-SIL, se requiere revisión de la paciente por parte del ginecólogo, y su significación clínica puede conllevar la realización de otras pruebas para confirmación del diagnóstico. Por esta razón, en nuestro estudio agrupamos estas categorías como “Requiere revisión” y las citologías normales como “No requiere revisión” para plantear una clasificación binaria.

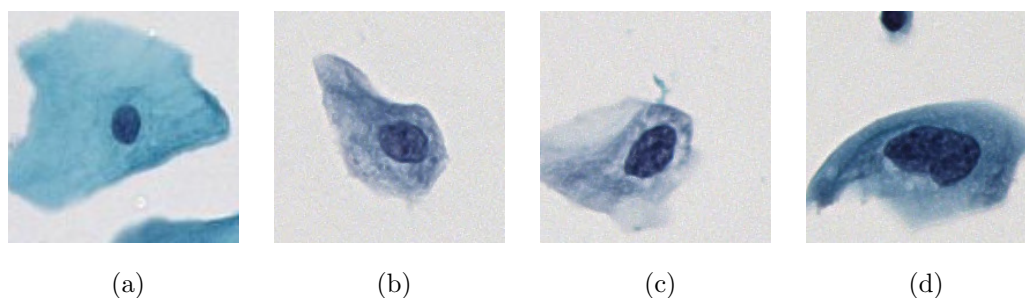


Figura 4.3: Ejemplos procedentes de la base de datos utilizada, obtenida tras la fragmentación de imágenes procedentes del archivo de imágenes del Servicio de Anatomía Patológica el Complejo Hospitalario de Cartagena Santa Lucía-Santa María del Rosell (Cartagena, Murcia): (a) Célula escamosa normal (b) ASC-US (c) L-SIL y (d) H-SIL.

En nuestro caso, para disminuir la posible influencia de los sesgos en el muestreo y procesamiento, las citologías fueron tomadas mediante cepillado exo y endocervical simultáneo con el cepillo LBC de BD®, realizando cinco giros del cepillo alrededor del orificio cervical externo en el sentido de las agujas del reloj (siguiendo el protocolo de toma de muestras de la marca comercial), y se utilizó el sistema de conservación para citologías en medio líquido BD SurePath® liquid-based Pap test®, como se muestra en la **Figura 1.8c**. Para el procesamiento de las muestras se empleó el sistema automatizado y robotizado PrepStain®, de forma que se disminuye la influencia del factor humano en la aparición de errores de procesamiento, pues se realiza tanto la centrifugación y filtrado como la fijación de una única lámina de células sobre el cristal portaobjetos de forma automatizada.

Posteriormente, se realizó el escaneo de las muestras citológicas, que fueron automáticamente fragmentadas en imágenes celulares simples mediante un programa de software y clasificadas una a una en sus categorías diagnósticas

por dos patólogos, trabajando de forma independiente, para posteriormente utilizarlas en el entrenamiento y test de la CNN.

Debe resaltarse que las células de la **Figura 4.3**, son células individuales con un fondo limpio, esto es, sin ninguna otra célula superpuesta o junto a ellas ni con artefactos de tinción. Con el objetivo de obtener células aisladas en estas condiciones, debe realizarse un preprocesamiento de las muestras. Realmente, la mayor parte de las células que se encuentran en una muestra citológica se encuentran plegadas, con aglomeraciones celulares, superposiciones con otras células, con manchas de tinción, artefactos de procesamiento, o borrosas. La borrosidad se debe al posicionamiento de las células en niveles diferentes de profundidad, aun tratándose de muestras en monocapa. Este efecto no es apreciable en las imágenes 2D, pero al escanear fijando un foco de ampliación, todas aquellas células que queden por delante o por detrás del foco se verán borrosas, efecto que en el examen manual con un microscopio se evita ajustando el foco en cada enfoque diferente del campo visual si así lo requiere, pero que en las muestras escaneadas no es posible evitar. La **Figura 4.4** muestra un ejemplo de células borrosas, superpuestas y plegadas, para cada una de las categorías de células escamosas que estudiamos, tal y como suelen encontrarse en una muestra citológica habitual.

Se debe tener en cuenta que para nuestro trabajo no se utiliza ningún método de preprocesamiento de las células, es decir, estas son escaneadas en alta resolución directamente de una muestra real, tal y como se ha expuesto. Por ello, se debe enfatizar que entre las imágenes clasificadas por los patólogos y empleadas para los experimentos (test, validación y entrenamiento) es posible encontrar imágenes con células superpuestas, dobladas o con cualquier otro elemento

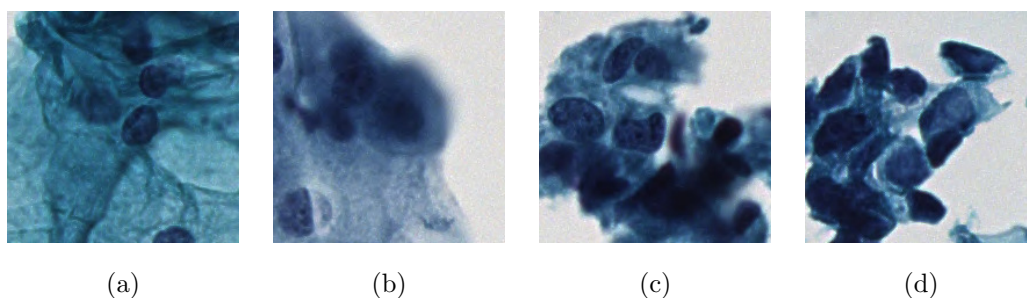


Figura 4.4: Distintos ejemplos donde se muestra la complejidad para la clasificación de las células debido a superposicionamientos, plegamientos, borrosidad por defectos de enfoque y artefactos de tinción en (a) Célula normal, (b) ASC-US, (c) L-SIL y (d) H-SIL. Imágenes procedentes de la base de datos utilizada, obtenida tras la fragmentación de imágenes procedentes del archivo de imágenes del Servicio de Anatomía Patológica el Complejo Hospitalario de Cartagena Santa Lucía-Santa María del Rosell (Cartagena, Murcia)

en el fondo distinto a células escamosas, tal y como ocurre cuando se examina una citología cervicovaginal real.

Si únicamente se clasificase células como las expuestas en la **Figura 4.3**, entonces, la clasificación sería mucho más fácil y la exactitud obtenida sería muy alta, sin embargo se perdería la perspectiva de la realidad de las células que se encuentran en una muestra de una paciente y se clasificaría una base de datos no representativa de la realidad mediante un procedimiento cuyos resultados difícilmente serían extrapolables o aplicables a la clínica. Por tanto, el planteamiento de este estudio es conseguir una clasificación satisfactoria considerando todas las posibles situaciones que pueden aparecer en una muestra, y comparar sus resultados con las tasas de detección de patólogos expertos cuando examinan muestras de pacientes reales.

4.1.4. Base de datos de nueva creación mediante la aplicación de técnicas de fusión de imágenes sobre la base de datos original de células escamosas

Método de fusión de imágenes

En una muestra de una citología cervicovaginal real suele aparecer una multitud de estirpes celulares, distintas a las células escamosas que son objeto de esta clasificación, como la mostrada en la **Figura 4.5a**, tales como células glandulares (**Figura 4.5b**), leucocitos polimorfonucleares (**Figura 4.5c**), eritrocitos, flora cocobacilar saprófita, flora patógena como hongos, bacterias o protozoos, espermatozoides, así como aglomerados celulares (**Figura 4.5d**), y artefactos de procesamiento o problemas de enfoque, con visualización de células borrosas (**Figura 4.5e**). Una vez obtenida una muestra suficiente de células escamosas individuales clasificadas, y dado que el sistema DL requiere un alto número de imágenes para su entrenamiento, se propone un método de fusión de imágenes, donde se superponen varias imágenes previamente clasificadas, modificando su posición mediante rotaciones, solapamientos y agregaciones celulares entre sí, multiplicando así el número de imágenes incluidas en cada grupo diagnóstico.

Este procedimiento de fusión de imágenes se basa en la mezcla de distintas imágenes del mismo o distinto grupo diagnóstico, introduciendo rotaciones de las mismas y solapamientos entre ellas, con lo que se multiplica el número de imágenes de muestra disponibles. La categoría diagnóstica correspondiente a la imagen resultante de la fusión será la categoría de la célula con mayor grado de atipia, esto es, por ejemplo, al fusionar una célula normal con una L-SIL, la categoría de la imagen resultante será L-SIL, y al fusionar una célula L-SIL con una H-SIL, la categoría de la imagen resultante será H-SIL. Con este sistema, partiendo de una selección de 1.405 imágenes de la base original, se obtiene

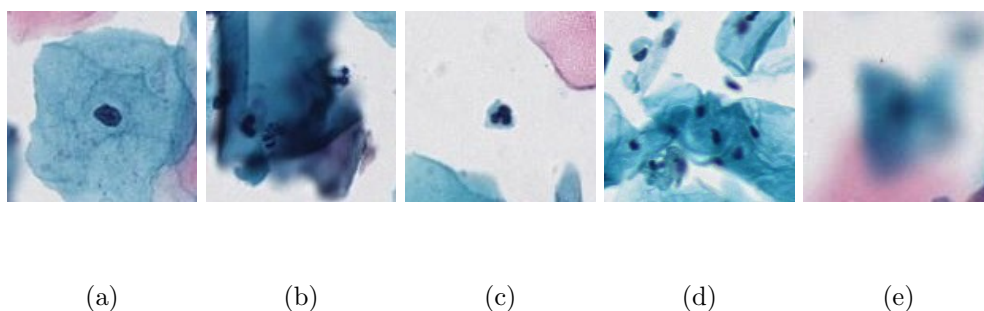


Figura 4.5: Ejemplos de distintas células procedentes de la base de datos del estudio, sin procesamiento de imagen: (a) Célula escamosa benigna, (b) células glandulares, (c) leucocitos polimorfonucleares, (d) células aglomeradas de distintas estirpes, o (e) célula escamosa borrosa, por defecto de enfoque. Imágenes procedentes de la base de datos utilizada, obtenida tras la fragmentación de imágenes procedentes del archivo de imágenes del Servicio de Anatomía Patológica el Complejo Hospitalario de Cartagena Santa Lucía-Santa María del Rosell (Cartagena, Murcia).

20.000 imágenes de nueva creación para cada categoría, es decir, un total de 80.000 imágenes distintas, con presentaciones similares a las que se encuentran en citologías reales. Con esta nueva base de datos, se realiza de nuevo el entrenamiento y evaluación del sistema de clasificación, y se compara sus resultados con los obtenidos en el experimento anterior de clasificación de la base de datos original.

Este planteamiento pretende crear un conjunto de imágenes de células que refleje la realidad de una muestra citológica cervicovaginal con tinción de Papanicolaou, que denominamos *Cell Merging Approach* (CMA). Nuestra finalidad es crear un modelo de CNN que automáticamente clasifique las células escamosas en categorías establecidas, sin necesidad de procesar previamente las células.

Se debe tener en cuenta que en una muestra real, habitualmente encontramos muchas más células normales que células atípicas. Por tanto, existe una dificultad añadida para la obtención de un gran número de células que serán catalogadas como “Requiere revisión”. Dada esta situación, la propuesta de este método tiene dos finalidades: La primera de ellas es aumentar el número total de células para el entrenamiento de la CNN para obtener una red neuronal más robusta con menor variabilidad en los resultados. La segunda es conseguir un volumen importante de células superpuestas, plegadas, borrosas y artefactadas para poder obtener una visión más real de cómo se encuentran estas células en la muestra original. Estos solapamientos, plegamientos y artefactos, como se ha comentado anteriormente, son muy frecuentes en muestras reales, y buscamos obtener resultados realistas y capaces de detectar células atípicas en su estado original, sin preprocesamiento ni sobre un fondo artificialmente limpio.

La **Figura 4.6** describe gráficamente el proceso de fusión de imágenes realizado, donde, como podemos ver, no sólo se rota la célula, sino también todo el fondo que la rodea. Por tanto, la nueva imagen obtenida contiene más información que las células iniciales, y refleja de una forma más fiable la realidad de una citología cervicovaginal tomada mediante el muestreo habitual del cérvix de una paciente.

El proceso del método propuesto se compone de los siguientes pasos:

1. Selección aleatoria de dos células;
2. Selección aleatoria del ángulo de rotación de cada célula, siendo 0° , 90° , 180° , 270° ;
3. Superponer ambas células y crear una única imagen.

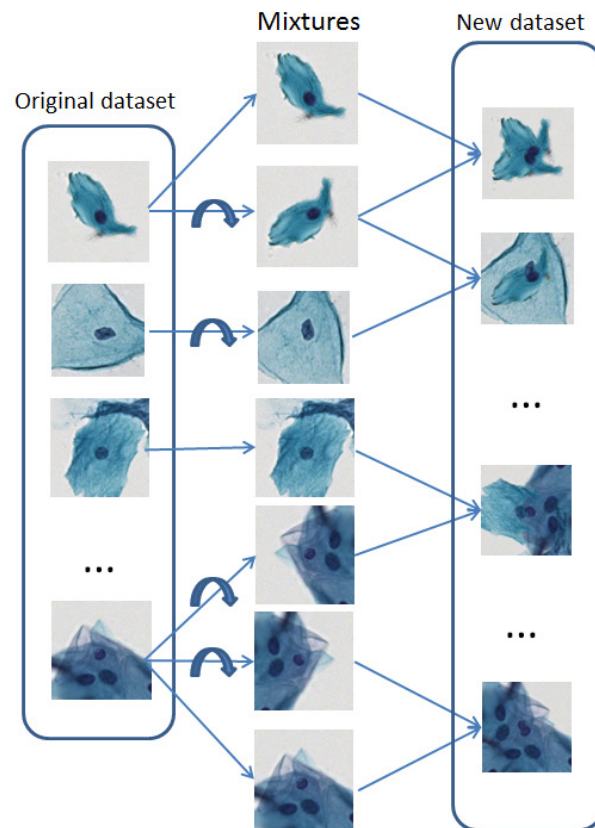


Figura 4.6: Descripción gráfica esquematizada del proceso de fusión de imágenes procedentes de la base de datos utilizada, obtenida tras la fragmentación de imágenes procedentes del archivo de imágenes del Servicio de Anatomía Patológica el Complejo Hospitalario de Cartagena Santa Lucía-Santa María del Rosell (Cartagena, Murcia). Figura tomada de Ortiz-Reina (2019) [4].

Con este procedimiento, crearemos 5.000 imágenes de cada categoría utilizando la CMA. De esta forma, obtendremos una muestra de 20.000 células de los 4 tipos, a las que añadiremos las rotaciones comentadas, obteniendo un total de 80.000 imágenes de muestra, siendo 20.000 de cada categoría, que serán utilizadas como entrada para el entrenamiento de la CNN.

4.2. Configuración de los experimentos.

4.2.1. Configuración de los experimentos para la clasificación de descriptores obtenidos de imágenes ecográficas de tumores de ovario.

En esta sección se describe las nociones principales de los cuatro métodos de clasificación basados en ML que se analizan en este trabajo: *K-Nearest Neighbors* (KNN), *Linear Discriminant* (LD), *Support Vector Machine* (SVM) y *Extreme Learning Machine* (ELM). Debe remarcarse que en este trabajo, las características geométricas FFT obtenidas serán utilizadas como entradas durante el entrenamiento y evaluación de los cuatro clasificadores. Por tanto, de acuerdo a nuestra notación, la base de datos \mathbf{X} está compuesta por $N=187$ vectores de entrada,

$$\mathbf{X} = \{\mathbf{x}_n\}_{n=1}^N \quad (4.4)$$

donde cada vector de entrada está compuesto por tres características geométricas FFT:

$$\mathbf{x}_n = [x_{n,1}, x_{n,2}, x_{n,3}]^T \quad (4.5)$$

Además, las dos posibles categorías de clasificación para un vector de entrada dado, C_1 (benigno) y C_2 (maligno) son etiquetadas respectivamente como +1 y -1. Entonces, el vector objetivo, o resultado deseado, es representado por

$$\mathbf{t} = \{t_n\}_{n=1}^N \quad (4.6)$$

donde t_n puede ser +1 (C_1 : benigno) o -1 (C_2 : maligno).

K-Nearest Neighbors (KNN)

El clasificador KNN asigna el vector de entrada al grupo de clasificación que más ejemplos de entrenamiento tiene entre el número K determinado de vecinos de dicho vector de entrada, según Alpaydin y cols. (2010) [125]. En la versión estándar de este clasificador no paramétrico, todos los vecinos tienen el mismo peso (o voto) y el grupo que tenga el mayor número de votantes entre los K vecinos será el elegido. Este clasificador es conocido como “algoritmo perezoso”, dado que no asigna ninguna categoría a la muestra, sino que memoriza el conjunto completo de datos y analiza la proximidad o similitud de la muestra a estudiar con el resto de ejemplos memorizados. En este método el valor de K es el parámetro principal que se debe seleccionar. Otro aspecto importante es la adecuada selección de la distancia métrica. En este trabajo se consideran dos métricas ampliamente utilizadas: *Distancia Euclidiana* y la *Métrica de bloques*. Ambos son casos especiales de la *métrica Minkowski*. Dados dos vectores de entrada diferentes \mathbf{x}_n y \mathbf{x}_m , su distancia, utilizando la métrica de Minkowski viene dada por

$$d_{n,m} = \sqrt[p]{\sum_{i=1}^d |x_{n,i} - x_{m,i}|^p} \quad (4.7)$$

siendo $d=3$ en nuestro trabajo. La Métrica de Bloques (*City block distance*) es una variación de la métrica de Minkowski, con $p=1$; y cuando $p=2$ se obtiene la distancia Euclidiana. Una desventaja del método KNN es que debe retener todos los datos de las muestras de entrenamiento para poder clasificar nuevos vectores en el futuro, con lo que requiere sistemas con una memoria relativamente grande.

Linear Discriminant (LD)

El método LD se fundamenta en encontrar una línea en el espacio de representación muestral que lo divida, dejando a cada lado de la línea cada una de las categorías a clasificar. Se trata de un modelo sencillo pero a la vez muy útil cuando se trata de clasificar dos categorías netamente separadas. El método obtiene la categoría de clasificación de \mathbf{x} utilizando una combinación lineal ponderada de sus vectores de entrada [125, 126]:

$$y(\mathbf{x}) = w_0 + w_1x_1 + w_2x_2 + \dots + w_dx_d \quad (4.8)$$

donde $\mathbf{w} = [w_1, w_2, \dots, w_d]$ es el peso del vector y w_0 es el sesgo. La magnitud del peso w_i muestra la importancia de x_i y su signo indica si su efecto es positivo o negativo. En un LD, \mathbf{x} es clasificado como C_1 si $y(\mathbf{x}) > 0$ y como C_2 si $y(\mathbf{x}) < 0$. El límite de decisión es aquel cuyos vectores de entrada dan un resultado $y(\mathbf{x}) = 0$ siendo un hiperplano dimensional $(d-1)$ en el espacio de entrada dimensional d .

A pesar de su simplicidad, el LD ha demostrado su utilidad en muchas aplicaciones en el mundo real [127]. De hecho, se ha comprobado que el discriminante óptimo es lineal cuando la clasificación de las muestras sigue una distribución Gaussiana, con una matriz de covarianza compartida [125, 126]. El método LD puede ser utilizado incluso cuando esta premisa no se sustenta y los parámetros ponderados pueden ser computados sin realizar suposiciones sobre la densidad de los grupos de clasificación [125].

Hay diversas técnicas para determinar los valores apropiados para los parámetros ponderados de un LD utilizando los datos de entrenamiento disponibles [126]. De forma particular, en nuestro trabajo se aplica la ampliamente utilizada

aproximación de los mínimos cuadrados (least squares approach), que minimiza la siguiente función de error:

$$E = \frac{1}{2} \sum_{n=1}^N (y_n - t_n)^2 = \frac{1}{2} \sum_{n=1}^N ((w_0 + \mathbf{w}_n^T \mathbf{x}_n) - t_n)^2 \quad (4.9)$$

y su solución de mínimos cuadrados habitual viene dada por

$$\hat{\mathbf{w}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{t} = \mathbf{X}^\dagger \mathbf{t} \quad (4.10)$$

donde \mathbf{X}^\dagger es la matriz inversa generalizada de Moore-Penrose de \mathbf{X} . Debe destacarse que la Descomposición en Valores Simples, o *Simple Value Decomposition* (SVD) de \mathbf{X} se utiliza para calcular la pseudoinversa y así asegurar la estabilidad numérica y cálculos computacionales más rápidos.

Support Vector Machine

Los *Support Vector Machines* (SVM) [125, 126] son un método de clasificación basado en Kernels, fundamentado en la teoría de aprendizaje estadístico. Un Kernel es una función que transforma los datos de entrada en un espacio multidimensional y que puede ser función lineal (dot product) y no lineal (como las gaussianas o polinómicas). Siguiendo las experiencias previas con nuestra base de datos [104, 105, 106], se ha elegido un kernel lineal para nuestro estudio. Después de la transformación de las entradas mediante la aplicación del kernel, el SVM determina el hiperplano marginal máximo que separa ambos grupos diagnósticos en el espacio multidimensional resultante. Su solución analítica viene dada por aproximaciones de optimización de convexidad. Han sido propuestos diversos métodos de entrenamiento para el SVM [126], y en este trabajo, utilizamos dos procedimientos bien conocidos: Mínimos cuadrados, o *Least Squares* (LS), y Optimización Mínima Secuencial, o *Sequential Minimal Optimization* (SMO).

Extreme Learning Machine

El *Extreme Learning Machine* (ELM) se basa en el concepto de que si la ponderación del Perceptrón Multicapa, o *Multi-Layer Perceptron* (MLP) de entrada es asignada con valores aleatorios, el MLP puede ser considerado como un sistema lineal y la ponderación de la salida puede ser fácilmente obtenida mediante la pseudoinversa de la matriz de neuronas ocultas \mathbf{H} para un conjunto de entrenamiento dado. Aunque los planteamientos iniciales de este sistema fueron previamente analizados en otros trabajos [128, 129], Huang fue quien lo formalizó [130, 131], demostrando que el ELM es un aproximador universal para un amplio rango de nodos computacionales aleatorios, y todos los parámetros nodales ocultos pueden ser generados aleatoriamente en función de una distribución continua de probabilidad sin ningún conocimiento previo. Por tanto, dado un conjunto de N vectores de entrada, un MLP puede aproximar N casos con error cero, $\sum_{i=1}^N \|\mathbf{y}_i - \mathbf{t}_i\| = 0$, siendo $\mathbf{y}_i \in R^m$ la red de salida para el vector de entrada $\mathbf{x}_i \in R^n$ con el vector objetivo $\mathbf{t}_i \in R^m$. Por tanto, existe $\beta_j \in R^m$, $\mathbf{w}_j \in R^n$ y $b_j \in R$ de forma que,

$$\mathbf{y}_i = \sum_{j=1}^M \beta_j f(\mathbf{w}_j \cdot \mathbf{x}_i + b_j) = \mathbf{t}_i, \quad i = 1, \dots, N. \quad (4.11)$$

donde $\beta_j = [\beta_{j1}, \beta_{j2}, \dots, \beta_{jm}]^T$ es el vector ponderado que conecta los j° nodos ocultos con los nodos de salida, $\mathbf{w}_j = [w_{j1}, w_{j2}, \dots, w_{jn}]^T$ es el vector ponderado que conecta los j° nodos ocultos con los nodos de entrada, y b_j es el sesgo del j° nodo oculto. Para una red con M nodos ocultos, las N ecuaciones previas pueden expresarse como

$$\mathbf{H}\mathbf{B} = \mathbf{T}, \quad (4.12)$$

donde

$$\mathbf{H}(\mathbf{w}_1, \dots, \mathbf{w}_M, b_1, \dots, b_M, \mathbf{x}_1, \dots, \mathbf{x}_N) = \quad (4.13)$$

$$= f(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) f(\mathbf{w}_M \cdot \mathbf{x}_1 + b_M) \ddots f(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) f(\mathbf{w}_M \cdot \mathbf{x}_N + b_M)_{N \times M} \quad (4.14)$$

$$\mathbf{B} = \beta_1^T; \beta_M^T \in R^{M \times m} \text{ and } \mathbf{T} = \mathbf{t}_1^T; \mathbf{t}_N^T \in R^{N \times m} \quad (4.15)$$

donde $\mathbf{H} \in R^{N \times M}$ es la matriz de capa de salida oculta del MLP, $\mathbf{B} \in R^{M \times m}$ es la matriz ponderada de salida, y $\mathbf{T} \in R^{N \times m}$ es la matriz objetivo de los N casos de entrenamiento. Por tanto, como \mathbf{w}_j y b_j con $j = 1, \dots, M$, son seleccionados de forma aleatoria, el entrenamiento del MLP viene dado por la solución del problema de mínimos cuadrados de (4.12), por ejemplo, la ponderación óptima de la capa de salida es $\hat{\mathbf{B}} = \mathbf{H}^\dagger \mathbf{T}$, donde \mathbf{H}^\dagger es la pseudoinversa de Moore-Penrose [132].

Por tanto, el ELM para el entrenamiento de MLP puede ser resumido de la siguiente manera:

Requisitos: Dado un conjunto de entrenamiento

$$\mathcal{D} = \{(\mathbf{x}_i, \mathbf{t}_i) | \mathbf{x}_i \in R^n, \mathbf{t}_i \in R^m, i = 1, \dots, N\},$$

una función de activación f y un número M de neuronas ocultas.

1. Asignar ponderaciones arbitrariamente a las entradas \mathbf{w}_j y sesgos b_j , $j = 1, \dots, M$.
2. Computar la matriz de capas ocultas de salida \mathbf{H} utilizando (4.14).
3. Calcular las ponderaciones de la matriz de salida $\mathbf{B} = \mathbf{H}^\dagger \mathbf{T}$, donde \mathbf{B} y \mathbf{T} son definidas en (4.15).

El ELM aporta un entrenamiento de MLP rápido y eficiente [133], pero requiere la fijación del número de neuronas ocultas. Para evitar la búsqueda exhaustiva del valor óptimo de M , se han propuesto distintos métodos de poda [134, 135, 136, 137, 138], de entre los que el más comúnmente utilizado es el conocido como *Optimally Pruned-ELM* (OP-ELM) [138]. Este método establece un número inicialmente muy alto de neuronas ocultas ($M \gg N$) y por medio de

la aplicación del algoritmo de Regresión del Mínimo Ángulo, o *Least Angle Regression* (LARS) [139], realiza una poda de neuronas en función de su relevancia para la resolución del problema (4.12). La poda de neuronas se realiza utilizando la validación cruzada en la que, para el número N total de muestras utilizadas en el estudio, se extrae una del conjunto y será utilizada como test y el resto $N-1$ es utilizado como conjunto de entrenamiento. Este procedimiento es repetido N veces, una por cada vector de entrada diferente. De esta forma, todos los casos son utilizados tanto para entrenamiento como para test durante las N repeticiones del procedimiento, por tanto, las medidas de evaluación de los resultados se calculan al final del procedimiento repetitivo. Este proceso es conocido como *Leave-One-Out Cross Validation* (LOO-CV), y permite seleccionar aquella combinación de neuronas, previamente escogida por el algoritmo LARS, que aporten el menor error LOO. El error LOO-CV se computa de forma eficiente utilizando la fórmula de Allen [138].

4.2.2. Evaluación de resultados y protocolo de entrenamiento y prueba para la base de datos de tumores ováricos.

Para comparar los resultados han sido utilizadas cuatro medidas [125]: Exactitud o *Accuracy* (ACC), Área bajo la Curva ROC o *Area Under Curve* (AUC), Sensibilidad (SEN) y Especificidad (SPE). La exactitud, sensibilidad y especificidad vienen definidas por las siguientes fórmulas:

$$ACC = \frac{VP + VN}{VP + FP + VN + FN} \quad (4.16)$$

$$Sensibilidad = \frac{VP}{VP + FN} \quad (4.17)$$

$$Especificidad = \frac{VN}{VN + FP} \quad (4.18)$$

donde VP, VN, FP y FN son los Verdaderos Positivos, Verdaderos Negativos, Falsos Positivos y Falsos Negativos respectivamente.

La Exactitud es la capacidad que la prueba diagnóstica tiene de predecir correctamente la presencia de la enfermedad cuando es positiva y la ausencia de la misma cuando es negativa.

La Sensibilidad es la capacidad que tiene una prueba diagnóstica de detectar los casos que tienen la enfermedad.

La Especificidad es la capacidad que tiene la prueba diagnóstica de detectar individuos sanos mediante el resultado negativo de la misma.

La Curva ROC, acrónimo de *Receiver Operating Characteristic*, es una representación gráfica de la sensibilidad de una prueba diagnóstica, en función de los falsos positivos, para distintos puntos de corte. Así, la curva es la representación de la capacidad de la prueba para discriminar entre pacientes sanos y enfermos, y el parámetro más utilizado para evaluar la capacidad diagnóstica mediante esta curva es el área que queda bajo la misma (AUC), que se puede interpretar como la probabilidad de que la prueba clasifique correctamente a dos individuos, uno sano y otro enfermo.

Tradicionalmente, la medida más ampliamente utilizada para evaluar la fiabilidad en problemas de clasificación es la ACC, sin embargo, esta medida ig-

nora las estimaciones de probabilidad de clasificación a favor de cada grupo de clasificación. En muchas áreas de investigación, y particularmente en aplicaciones biomédicas, el AUC supone una medición efectiva del resultado global de una prueba de clasificación. El AUC toma valores desde 0 hasta 1, donde 0 indica un modelo perfectamente inadecuado para el problema estudiado y 1 indica que se trata de un modelo perfectamente fiable para la clasificación del problema estudiado. De forma general, se considera un valor de 0,5 en el AUC como el mínimo exigible para una prueba de clasificación.

Para evaluar de forma fiel los resultados de los diferentes métodos de clasificación, en este trabajo se utiliza la validación (LOO-CV) [125], como ha sido comentado previamente, que evita sesgos indeseables derivados de la selección aleatoria de los conjuntos de entrenamiento y de test. Para el ELM, este procedimiento LOO-CV se repite 30 veces, y, debido a la determinación aleatoria de ponderaciones, los resultados se muestran en términos del mejor resultado y media y desviación estándar.

4.2.3. Configuración de la arquitectura de CNN para la clasificación de imágenes de células escamosas a partir de citologías cervicovaginales mediante Deep Learning.

La arquitectura de CNN implementada en este trabajo es la referencia BVLC del modelo CaffeNet, según Jia y cols. (2014) [140] que es una réplica de AlexNet con dos diferencias: La primera es que no requiere entrenamiento por lo que aumenta el número de datos disponible en la base, y la segunda es que el orden de las capas de pooling y normalización está fijado. La tasa de aprendizaje ha sido establecida en 0,001 y cada 1500 repeticiones disminuirá en un factor de 10.

La red neuronal utiliza una entrada de 227×227 , por lo que las imágenes son ajustadas a ese tamaño. Tiene 5 capas de convolución con kernel 11, 5, 3, 3 y 3 respectivamente, donde la primera, segunda y quinta capas tienen una capa de *max pooling* (todas con kernel 3), dos capas totalmente conectadas con 4096 neuronas y una última capa *softmax* con cuatro salidas para representar cada una de las categorías de clasificación multiclase (Normal, ASC-US, L-SIL y H-SIL). Utilizamos ReLU como función de activación. La **Figura 4.7** representa el esquema de funcionamiento de la CNN junto con los dos niveles de clasificación estudiados, el nivel 1 (multiclase) para asignar a cada célula su diagnóstico según la nomenclatura Bethesda 2014 y el nivel 2 (binaria) para clasificarlas en función de la necesidad de revisión o no de la muestra.

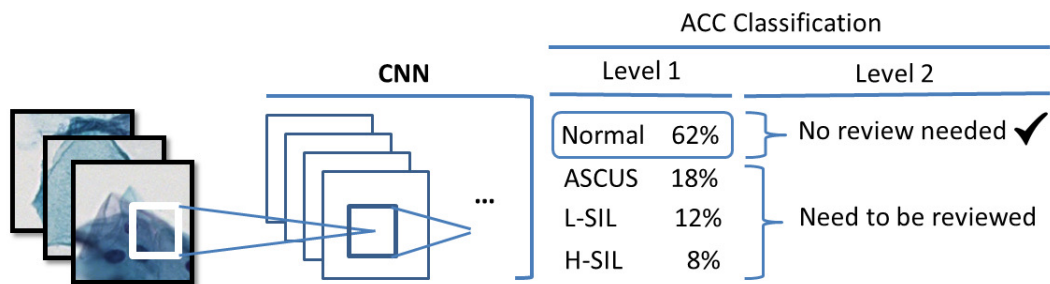


Figura 4.7: El método de clasificación realiza una medición de ACC en dos niveles: Una clasificación en 4 grupos diagnósticos (Normal, ASC-US, L-SIL y H-SIL) y otra clasificación respecto a si requiere o no una revisión por el patólogo. Figura tomada de Ortiz-Reina y cols. (2019) [4].

4.2.4. Configuración de los experimentos de clasificación de imágenes microscópicas de células escamosas obtenidas a partir de citologías cervicovaginales tras la creación de la base de datos por fusión de imágenes, mediante Deep Learning.

Los experimentos realizados para validar nuestras propuestas de clasificación de citologías cervicovaginales con tinción de Papanicolaou mediante DL y para crear nuevas muestras para validar y obtener un muestreo más real de los datos, han sido divididos en dos pasos, siguiendo la siguiente estructura común:

- La clasificación se realizará utilizando ambos tipos de categorización descritos anteriormente. Por un lado, las células se clasificarán de forma binaria considerando los grupos de "No requiere revisión" y "Requiere revisión". Por otro lado, se realizará una clasificación multiclase con el diagnóstico de cada célula, siendo categorizada como Normal, ASC-US, L-SIL y H-SIL.
- Cada imagen de entrada para la CNN será rotada tres veces en ángulos de 90°, de forma que obtenemos para cada célula, 4 imágenes diferentes mediante rotación.
- Las imágenes de entrada a la CNN se utilizan para entrenamiento y validación, pero para realizar los tests, se utilizarán 30 imágenes de cada categoría que fueron previamente seleccionadas de forma aleatoria, con las que se obtendrá la exactitud de la CNN.
- La fiabilidad de la CNN ha sido evaluada mediante la repetición de cada uno de los experimentos propuestos 5 veces. Así, por cada experimento, se muestra la media y desviación estándar, indicando así la estabilidad de la

técnica durante el proceso de validación de la CNN. Cuanto menor sea la SD, mayor estabilidad y fiabilidad tendrá la técnica.

A continuación se describe las diferencias entre ambos tipos de experimentos, realizados sobre la base de datos de células escamosas:

- El primer experimento se realiza utilizando las imágenes originales, clasificadas previamente por patólogos expertos, con 907 imágenes inicialmente catalogadas como Normales, 323 como ASC-US, 213 como L-SIL y 419 como H-SIL. Dada la facilidad para encontrar células normales entre las muestras habituales, pues suelen ser más numerosas que las células atípicas, para realizar los experimentos fueron seleccionadas aleatoriamente 450 imágenes Normales para poder equilibrar los grupos de clasificación. Este procedimiento deja una muestra de 1405 imágenes. Teniendo en cuenta las tres rotaciones en el espacio que fueron comentadas con anterioridad, se obtiene una muestra de 5620 imágenes de entrada. Este experimento es denominado “Clasificación de células originales”.
- En el segundo experimento, se valida la exactitud y fiabilidad de la CNN utilizando la base de datos obtenida mediante CMA. Este experimento es denominado “Clasificación de células fusionadas”.

4.3. Descripción del Hardware y Software utilizados

Los ensayos de clasificación de imágenes ecográficas de tumores ováricos han sido realizados en entorno MATLAB R2018a, siempre en el mismo ordenador.

El método propuesto para la clasificación de citologías cervicovaginales con tinción de Papanicolaou ha sido diseñado utilizando el marco de trabajo Caffe

según Jia y cols. (2014) [140], con su *interface* en *Python*. *Python* es un lenguaje de programación de alto nivel, ampliamente utilizado en el manejo de grandes volúmenes de datos y para la creación de algoritmos de DL. Este marco de trabajo se encuentra instalado en un sistema heterogéneo con configuración multicore+GPU. Este sistema consta de cuatro multiprocesadores Intel Xeon E5-2650 v4 funcionando a 2.2 GHz. Cada multiprocesador tiene 12 cores. La GPU es una NVIDIA GTX 1080 ti con 3584 CUDA-cores (28 Multiprocesadores Streaming y 128 cores en cada Multiprocesador funcionando con un reloj de estimulación de 162 GHz), ofreciendo una potencia de procesamiento bruto de hasta 10.8 TFLOPS. Además el sistema se completa con los 12 GB de una memoria GDDR5.

RESULTADOS

5. RESULTADOS

5.1. Resultados de la clasificación automática de imágenes ecográficas de tumores de ovario

La **Tabla 5.1** muestra los resultados obtenidos, en las cuatro medidas de valoración comentadas: Exactitud (ACC), Área bajo la Curva ROC (AUC), Sensibilidad (SEN) y Especificidad (SPE).

Tabla 5.1: Resultados Obtenidos (Exactitud (en %) -ACC-, Área bajo la curva ROC -AUC-, Sensibilidad (en %) -SEN- y Especificidad (en %) -SPE-) por los clasificadores KNN, LD, SVM y ELM utilizando Características Geométricas obtenidas por FFT. Evaluación realizada mediante procedimiento LOO-CV.

Método		ACC	AUC	SEN	SPE
KNN	distancia Euclidiana, K=1	50,27	0,4836	58	40
	distancia Euclidiana, K=10	52,94	0,4522	78	16
	distancia Euclidiana, K=15	56,68	0,4377	91	5
	distancia Euclidiana, K=30	55,08	0,3907	89	4
	distancia City block, K=1	53,48	0,5127	63	40
	distancia City block, K=10	57,20	0,5169	82	20
	distancia City block, K=15	58,29	0,4912	93	7
	distancia City block, K=30	58,29	0,4801	94	5
LD	método de Mínimos Cuadrados	85,56	0,8514	89	80
SVM (con kernel Lineal)	SMO training	87,70	0,8740	91	83
	LS training	86,10	0,8545	88	84
ELM (mejor resultado)	Linear Kernel	84,49	0,8551	94	71
	Kernel Sigmoide	87,17	0,8676	90	80
	Kernel Gaussiano	86,10	0,8620	92	79
	Kernel Lineal-Gaussiano	87,70	0,8740	92	80
	Kernel Sigmoide-Gaussiano	87,17	0,8692	93	79
	Kernel Lineal-Sigmoide-Gaussiano	87.17	0,8765	93	77

Los resultados obtenidos con el método KNN independientemente de la distancia métrica utilizada y el número de kernels fijado, obtiene unos resultados con una exactitud máxima alcanzada de 58,29 %, y un AUC muy cercano a 0,5 o incluso inferior. Este método ofrece una sensibilidad que puede alcanzar un 94 %, pero con una especificidad entre el 4 y el 40 %.

El método LD ofrece una Exactitud de 85,56 %, con un área bajo la curva de

0,85, una sensibilidad del 89 % y una especificidad de 80 %.

Con SVM, y mediante el procedimiento de Optimización Mínima Secuencial se obtiene una exactitud de 87,70 %, con un área bajo la curva de 0,87, sensibilidad de 91 % y especificidad de 83 %, que en la tabla se encuentran resaltados en negrita. Resultados similares se obtienen mediante ELM, alcanzando el 87,7 % de exactitud, un área bajo la curva ROC de 0,87, sensibilidad entre el 90 y 93 % y especificidad entre el 71 y el 80 %.

La **Tabla 5.2** muestra los resultados obtenidos por el clasificador ELM al realizar el procedimiento LOO-CV con la base de datos de Características Geométricas obtenidas mediante FFT, en términos de Exactitud y Área bajo la curva ROC (Media y Desviación Estándar), mostrando así mismo el número de neuronas ocultas fijadas.

Tabla 5.2: Resultados obtenidos (Exactitud -ACC-, Área bajo la curva ROC -AUC- y neuronas ocultas -HN-) por el clasificador ELM de distintos Kernels (Media y Desviación estándar) utilizando Características Geométricas obtenidas por FFT. Evaluación realizada mediante procedimiento LOO-CV.

Kernel	ACC (in %)	AUC	HN
Lineal	84,49±0,00	0,8551±0,0000	3,00±0,00
Sigmoide	82,16±1,87	0,8183±0,0200	15,21±0,48
Gaussiano	84,90±0,98	0,8486±0,0109	15,70±0,41
Lineal-Sigmoide	82,37±1,62	0,8199±0,0174	15,82±0,48
Lineal-Gaussiano	85,22±1,17	0,8513±0,0125	16,11±0,44
Sigmoide-Gaussiano	82,82±2,05	0,8260±0,0217	13,75±0,45
Lineal-Sigmoide-Gaussiano	82,30±2,04	0,8208±0,0214	13,40±0,37

Se puede observar en esta **Tabla 5.2** los resultados obtenidos mediante la aplicación de distintos kernel con un número variable de neuronas ocultas, oscilando los resultados entre una exactitud de 82,160 %±2,04 obtenido con kernel Sigmoide, hasta el 85,22 %±1,17 % obtenido con kernel Lineal-Gaussiano. El área bajo la curva ROC oscila desde el 0,8183±0,0200 obtenido con kernel Sigmoide al de 0,8551±0,000 obtenido con kernel Lineal, utilizando un número de neuronas ocultas que oscilan entre 3 y 16,11±0,44 dependiendo del tipo de kernel utilizado.

5.2. Resultados de la Clasificación de Células Escamosas Cervicales con tinción de Papanicolaou.

5.2.1. Resultados de la Clasificación de Células Originales.

En esta sección se presenta los resultados obtenidos por la CNN utilizando las imágenes originales como base de datos, sin aplicar la CMA. Esta base de datos contiene inicialmente 1405 muestras. La **Tabla 5.3** muestra la ACC media, SD, SPE y SEN de los 5 modelos obtenidos de la prueba de la CNN con la misma base de datos, clasificando las células de forma binaria como “No requiere revisión” y “Requiere revisión”.

Tabla 5.3: ACC media, SD, SEN y SPE de la clasificación de la base de datos imágenes originales, en “No requiere revisión” y “Requiere revisión”.

Base de datos	% ACC \pm SD	SEN	SPE
Original	79,5 % \pm 3,3 %	83	71

Al clasificar esta base de datos mediante la CNN, la exactitud obtenida es del 79,5 %, con una desviación estándar del 3,3 %, sensibilidad del 83 % y especificidad del 71 %.

Al analizar los resultados para valorar los errores de clasificación binaria, esto es, falsos positivos y falsos negativos, revisando en qué categoría se clasifica cada imagen que ha sido identificada de forma acertada, la **Figura 5.1** muestra cómo las categorías H-SIL y L-SIL obtienen una exactitud de un 92 % y un 82.7 % respectivamente, mientras que los resultados de clasificación de ASC-US y células Normales son de un 70 % y 73 % respectivamente.

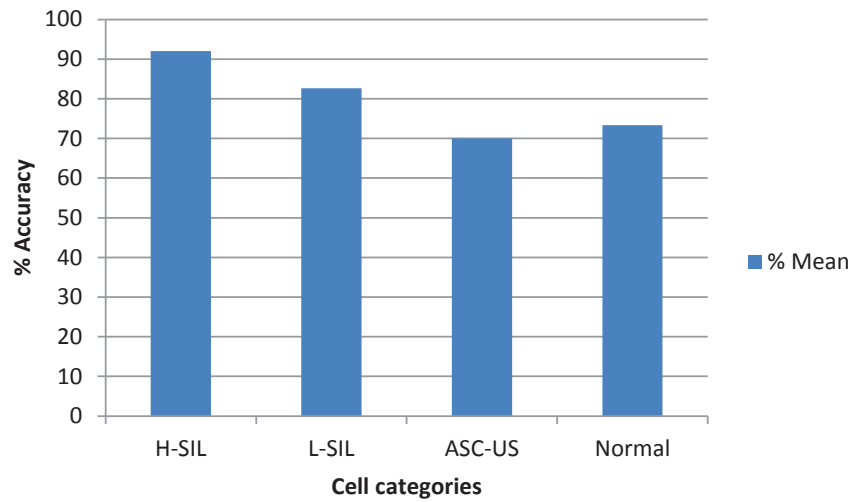


Figura 5.1: Los aciertos del modelo CNN en la clasificación binaria de “No requiere revisión” y “Requiere revisión” se muestran separados para cada categoría celular utilizando las imágenes originales como entrada. Figura tomada de Ortiz-Reina y cols. (2019) [4].

A continuación se presentan los resultados obtenidos por el modelo de CNN cuando clasifica en la modalidad multiclase, cada una de las categorías de células escamosas. La **Tabla 5.4** presenta los aciertos obtenidos de media por el modelo de CNN para cada categoría individual (Normal, L-SIL, H-SIL, ASC-US).

Tabla 5.4: Media y SD de ACC en % para clasificación multiclase en categorías individuales de las imágenes originales en H-SIL, L-SIL, ASC-US y Normal.

Tipos celulares	H-SIL	L-SIL	ASC-US	Normal
%ACC	67,33 %	10,67 %	30,67 %	73,33 %
%SD	5,48 %	2,79 %	2,79 %	3,33 %

La **Tabla 5.5** presenta la matriz de confusión del rendimiento del método de clasificación multiclase en categorías individuales aplicado sobre la base de datos de células escamosas original, resaltando en negrita los porcentajes de clasificación correctos en cada categoría.

Tabla 5.5: Matriz de confusión para la clasificación multiclase de categorías individuales (H-SIL, L-SIL, ASC-US y Normal) utilizando las imágenes originales como entrada, resaltando en negrita las clasificaciones correctas en cada categoría.

Tipos celulares	H-SIL	L-SIL	ASC-US	Normal
H-SIL	67,3 %	8,7 %	16,0 %	8,0 %
L-SIL	22,0 %	10,7 %	50,0 %	14,0 %
ASC-US	27,3 %	12,0 %	30,7 %	30,0 %
Normal	8,7 %	8,7 %	12,7 %	73,3 %

La clasificación células normales correctamente como tales, es del 73,3 %, con una clasificación errónea de este tipo celular como ASC-US en un 12,7 % y como L-SIL y H-SIL en un 8,7 % en ambas categorías. La detección correcta de los casos de ASC-US es de un 30,7 %, mientras que su clasificación errónea como células normales es del 30 %, como H-SIL un 27,3 %, y como L-SIL un 12 %. La detección correcta de L-SIL es de un 10,7 %, mientras clasifica erróneamente estas células como ASC-US en el 50 % de los casos, como normales en un 14 % de los casos y como H-SIL en un 22 % de los casos. La clasificación correcta de las células H-SIL es de un 67,3 %, y la clasificación errónea como L-SIL ocurre en el 8,7 % de los casos, como ASC-US en el 16 % y como células normales en el 8 % de los casos.

5.2.2. Resultados de la Clasificación de Células Fusionadas

Inicialmente, partimos de un modelo de CNN que clasifica las células de forma binaria en “No requiere revisión” y “Requiere revisión”. El número de células de entrada es de 80.000, utilizando el CMA. Los resultados, tras validar el modelo con las células de test, aparecen en la **Tabla 5.6**, mostrando la ACC media y su SD tras cinco repeticiones de la clasificación del modelo de CNN, especificidad y sensibilidad.

Tabla 5.6: Media y SD de la ACC, sensibilidad y especificidad para la muestra generada mediante CMA.

Base de datos	% ACC \pm SD	SEN	SPE
Fusión	88,8 % \pm 1,0 %	92 %	83 %

Como se muestra en la **Tabla 5.6**, la exactitud alcanzada en la clasificación de las imágenes generadas mediante el procedimiento de fusión de imágenes originales alcanza el 88,8 % con una desviación estándar del 1 %. La Sensibilidad obtenida es del 92 % y la Especificidad del 83 %.

Tras analizar los resultados de los modelos de CNN, entrenando el modelo con dos categorías de clasificación, realizaremos el análisis del comportamiento de los modelos de CNN teniendo en cuenta las cuatro categorías de células escamosas estudiadas. La **Tabla 5.7** presenta la ACC media y SD de los cinco modelos ejecutados para la clasificación multiclase las células como H-SIL, L-SIL, ASC-US y Normal.

Tabla 5.7: Media y SD de la ACC en la clasificación multiclase de la muestra generada mediante CMA en las categorías individuales H-SIL, L-SIL, ASC-US y Normal.

	H-SIL	L-SIL	ASC-US	Normal
% ACC	68,00 %	16,67 %	52,00 %	86,00 %
% SD	1,83 %	3,33 %	3,80 %	3,65 %

Estos resultados de la **Tabla 5.7** indican el porcentaje de exactitud en los aciertos con la clasificación multiclase de cada categoría, pero si se analiza y añade a la tabla los falsos negativos, es decir, los errores de clasificación, de cada categoría por separado, obtenemos la **Tabla 5.8** de la matriz de confusión, que refleja que las células H-SIL son detectadas correctamente en un 68 %, clasificando erróneamente como L-SIL el 17,3 % y como ASC-US el 14 %, categorías incluidas dentro de la clasificación binaria como “Requiere revisión” mientras que clasifica erróneamente como Normales el 0,7 %. Con las células L-SIL ocurre una situación similar, aunque con menor exactitud, detectando correctamente el

16,7 %, con clasificación errónea como H-SIL el 18,7 % y como ASC-US el 50,7 %, que igualmente quedan incluidos en la categoría binaria de “Requiere revisión”, mientras que clasifica erróneamente como Normales el 10,7 % de estas células. Las imágenes catalogadas como ASC-US correctamente son el 52,2 %, clasificando erróneamente éstas como H-SIL en un 11,3 % y como L-SIL el 20,7 %, estando incluidas también en la categoría “Requiere revisión”, con un 16 % clasificado erróneamente como Normal. Respecto a las células Normales, éstas son correctamente detectadas en el 86 % de los casos, clasificando erróneamente como H-SIL el 6,7 %, y como L-SIL y ASC-US el 5,3 %.

Tabla 5.8: Los aciertos del modelo de CNN para la clasificación binaria en “No requiere revisión” y “Requiere revisión” se muestran separados para cada categoría celular individual utilizando como entrada la muestra obtenida mediante CMA.

Tipo de célula	H-SIL	L-SIL	ASC-US	Normal
H-SIL	68,0 %	17,3 %	14,0 %	0,7 %
L-SIL	18,7 %	16,7 %	50,7 %	10,7 %
ASC-US	11,3 %	20,7 %	52,0 %	16,0 %
Normal	6,7 %	5,3 %	5,3 %	86,0 %

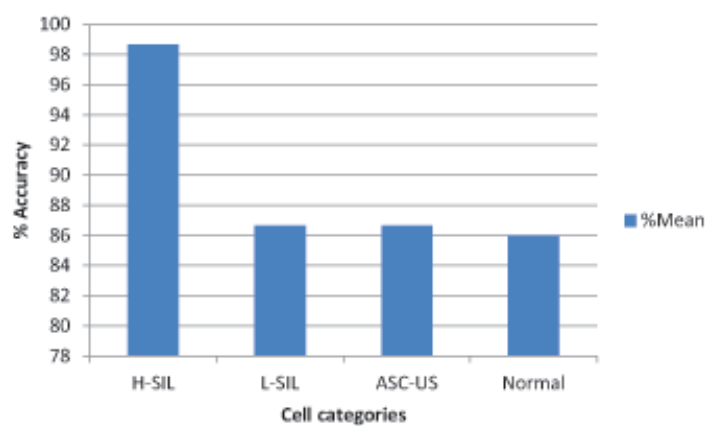


Figura 5.2: Presentación de los aciertos de las células clasificadas dentro de las categorías “Requiere revisión” y ‘No requiere revisión”, en relación al tipo de célula clasificada, utilizando la base de datos creada por CMA como entrada.

La **Figura5.2** presenta gráficamente la exactitud en la clasificación en la modalidad multiclase en cada categoría, destacando en la clasificación de H-SIL una exactitud del 98,7%, siendo superior al 85% para células normales, ASC-US y L-SIL.

DISCUSIÓN

6. DISCUSIÓN

6.1. Clasificación de tumores de ovario con técnicas clásicas de ML

En primer lugar, debemos tener en cuenta que para realizar nuestro trabajo utilizamos la misma base de datos que otros grupos de investigación, y que los resultados obtenidos previamente con los mismos descriptores (Histogramas de intensidad y Patrón Binario Local), se sitúan en una exactitud del 77%, según los trabajos de Khazendar y cols. (2014) [104], (2015) [105] y (2016) [106], en los que se utiliza un sistema de fusión de niveles de decisión.

Como primer comentario, se puede desprender de la observación de nuestros resultados que la clasificación mediante KNN ofrece un rendimiento muy pobre para analizar las imágenes ecográficas de tumores ováricos, independientemente de la distancia elegida (Euclidiana o City block), y del tamaño de aproximación local seleccionado, así como el número de vecinos cercanos. Aunque valores mayores de K y la distancia City block ofrecen una mejoría de los resultados, éstos siguen siendo muy poco fiables (menos del 59% de exactitud). Estos resultados pueden explicarse porque el FFGF actúa como un método muy efectivo de reducción de dimensiones, con una muy pequeña pérdida de información, por lo que la distancia Euclidiana para los vecinos cercanos no es un método de clasificación apropiado para este problema, debido a la relativa-

mente alta dimensión de los vectores implicados. Por ello, y de acuerdo a los resultados obtenidos, el KNN debería ser obviado en el diseño de la fase de clasificación de nuestro sistema CAD.

Con respecto a los otros tres clasificadores, LD, SVM y ELM, sus resultados en la clasificación son significativamente mejores que los obtenidos por el KNN. Comparándolos entre ellos, el ELM y el SVM ofrecen mejor ACC y AUC que el LD, con ventajas que son aún más claras cuando el discriminante basado en kernel del SVM se entrena utilizando el método de SMO, y el kernel Sigmoide-Gaussiano para el ELM.

Los métodos analizados para la clasificación de descriptores obtenidos a partir del procesamiento de imágenes ecográficas de tumoraciones ováricas pueden aportar una importante mejoría en los resultados, pues se ha llegado a alcanzar hasta un 87,7% de exactitud. En particular, y debido al hecho de que el SVM también ha sido utilizado en los trabajos anteriores de Khazendar y cols. (2014) [104], (2015) [105] y (2016) [106] obteniendo una exactitud del 77%, siguiendo el mismo procedimiento de LOO-CV, se puede afirmar que la mejoría observada en los resultados se debe a que la información resultante de los descriptores Geométricos obtenidos mediante FFT hacen que la clasificación sea más sencilla y exacta. El ELM con Kernel Lineal-Gaussiano obtiene resultados similares al SVM (en ACC y AUC).

Algunas herramientas que han sido desarrolladas para ayudar a la detección de tumoraciones ováricas malignas, utilizadas en la práctica clínica habitual, y que obtienen unos resultados de clasificación altos, como los modelos 1 y 2 de Regresión Lineal de IOTA (LR 1 y 2), no sólo incluyen el análisis de las imágenes,

sino que también tienen en cuenta datos clínicos de cada paciente, como describe Kaijser y cols. (2013) [141]. Estos modelos, como anteriormente fue expuesto, se basan en un algoritmo que realiza un cálculo del riesgo de que una imagen de un tumor ovárico sea maligna o benigna teniendo en cuenta distintos datos clínicos ponderados y el análisis que el examinador realiza de la imagen. Al utilizar estas herramientas, el clínico debe analizar la imagen, determinando si hay presencia de los marcadores a analizar, tales como ascitis, proyecciones papilares, sombras acústicas, paredes internas irregulares en los quistes, captación de señal Doppler en las proyecciones sólidas papilares, el diámetro máximo del componente sólido del tumor, además de tener en cuenta la edad de la paciente, presencia de dolor, uso de terapias hormonales, o antecedentes oncológicos. Como se ha comentado, se realiza un cálculo asignando una ponderación distinta a cada aspecto valorado, y se obtiene como resultado la probabilidad de que la lesión sea benigna o maligna. El modelo LR2 es una versión abreviada del LR1 en el que se simplifica el número de variables a estudiar para realizar el cálculo, quedándose únicamente con la edad, presencia de ascitis, captación Doppler en proyecciones papilares sólidas en la imagen tumoral, diámetro máximo del componente sólido, presencia de paredes internas irregulares en los quistes, y la presencia de sombras acústicas. Cuando el sistema ofrece una probabilidad muy alta de que se trate de una lesión benigna o maligna, se puede asumir que sirve de ayuda al clínico para su diagnóstico, pero esto no es así cuando ofrece probabilidades en rango intermedio para malignidad, siendo estos casos los que habitualmente suscitan más dudas diagnósticas en los casos clínicos, pues cuando una imagen ecográfica es claramente benigna o maligna, el clínico suficientemente experimentado es capaz de asignar un diagnóstico acertado sin necesidad de recurrir a sistemas de ayuda.

El modelo ADNEX, también desarrollado e implementado por el grupo IO-TA, para realizar su análisis, incluye junto a las variables ecográficas (diámetro máximo de la lesión, proporción del tejido sólido, presencia de más de 10 lóculos quísticos, número de proyecciones papilares, sombras acústicas y ascitis), la edad de la paciente, el nivel sérico de Ca125 y el tipo de centro médico donde se realiza el estudio (oncológico o no), siendo esta última una novedad respecto al resto de modelos de clasificación descritos, según Van Calster y cols. (2015) [142]. Igualmente, es destacable que este modelo no utiliza la presencia de flujo sanguíneo en las lesiones tumorales estudiadas, evidenciado por la presencia de captación Doppler en las ecografías, como clásicamente se ha hecho, y la introducción del marcador analítico Ca125. De esta forma, los equipos ecográficos necesarios para realizar la exploración de la paciente y poder aplicar el modelo ADNEX no requieren ser de alta gama y tener la funcionalidad de análisis Doppler, pero sí se hace necesario disponer de la determinación analítica de Ca125, lo que puede suponer un análisis de la imagen en diferido o requerir una segunda visita de la paciente a la consulta para su correcta categorización, con los inconvenientes que esta situación pueda conllevar. Este modelo tiene una capacidad muy alta para la detección de tumoraciones anexiales malignas, con un AUC de 0,94 según Van Calster (2014) [60], pero aún requiere de más estudios para su validación externa y comprobar así su aplicabilidad en distintos centros sanitarios y con diferentes poblaciones de estudio según refiere Sayasneh (2016) [61].

Tanto el modelo ADNEX como los modelos LR1 y LR2 tienen un índice de error, que en la práctica clínica se demuestra mayor en las imágenes que más dudas de clasificación pueden ocasionar al especialista durante el examen ecográfico de la paciente, pues son las imágenes que no se muestran como claramente

benignas ni claramente malignas.

La colección de imágenes que hemos utilizado en este trabajo está compuesta por tumores ováricos de diferente origen histológico, como tumores estromales, epiteliales, de origen embrionario y metastásicos, tanto benignos como malignos. La adecuada caracterización de estas lesiones es un reto diario para un ginecólogo, quien habitualmente puede tener dudas en cuanto a la benignidad o malignidad de una lesión que no responda claramente a los patrones que los definen como tales, encontrando una gran variabilidad interpersonal cuando se examina el mismo caso por distintos profesionales. Nuestro método de clasificación incluye imágenes de todos los tipos de tumores ováricos que se pueden encontrar en la práctica clínica, como hemos destacado, incluyendo algunas imágenes de difícil caracterización, que son aquellas para las que es más interesante disponer de un CAD que pueda apoyar la decisión del clínico.

Nuestros resultados son similares a los obtenidos por expertos clínicos (según publica Timmerman y cols. (1999) [143], Exactitud entre el 82 y el 92%), con una alta Sensibilidad (92 %) y Especificidad (80 %), con Exactitud de 87,70 % y AUC de 0,8765 cuando utilizamos ELM, teniendo en cuenta que no se considera ningún dato clínico de la paciente y únicamente se clasifica las imágenes ecográficas en Modo B, sin señal Doppler, que podrían mejorar los resultados de forma similar a como otros métodos hacen, modulando así la ponderación de la imagen en el proceso de clasificación. El objetivo inicial de este trabajo no consiste en mejorar los resultados de clasificación de ginecólogos experimentados, sino aportar una herramienta que pueda apoyar las decisiones del clínico en situaciones de cierta incertidumbre, sin pretender sustituir o priorizar los resultados clasificatorios del CAD frente a la exploración y análisis realizados por

el profesional.

Los algoritmos utilizados en este trabajo tienen un bajo peso computacional, lo que podría ser una ventaja para su eventual implementación en equipos médicos, por lo que podrían ser útiles para ayudar al clínico en casos de alta dificultad diagnóstica, así como también podrían ser utilizados como herramienta para entrenamiento en programas de especialización o en el periodo de formación MIR, pues aporta un resultado de clasificación equiparable al que realizaría un clínico experto.

Por el momento, no existe ningún sistema computerizado de ayuda al diagnóstico médico basado en Inteligencia Artificial que esté incorporado a los equipos de ecografía médica, o que sea ampliamente utilizado en el entrenamiento de residentes.

Con nuestro trabajo, se pretende remarcar que la Inteligencia Artificial puede tener un lugar importante en este campo, pudiendo ayudar al adecuado diagnóstico clínico de imágenes tumorales ováricas complejas, y también para ayudar a mejorar la formación de los especialistas en su periodo de residencia MIR.

6.2. Clasificación de Citologías mediante DL

Los sistemas de clasificación basados en DL aportan una importante ventaja frente a los métodos clásicos de clasificación, pues permiten trabajar sobre las imágenes directamente sin necesidad de procesamiento previo, a diferencia de dichos métodos, que no clasifican las imágenes sino sus descriptores matemáti-

cos extraídos mediante procesamiento. Esta ventaja se manifiesta en un primer momento en el menor coste de tiempo y gasto computacional previo a la clasificación.

Partimos de la situación actual en la que no hay ningún método automático de clasificación para citologías que esté avalado por el Ministerio de Sanidad de España para su aplicación en screening poblacional, basándose en la revisión publicada por la *Axencia de Avaliación de Tecnoloxías Sanitarias de Galicia*(2013) [99], que a su vez, analiza, entre otros, los estudios realizados por Alves y cols. (2004) [94], el análisis del *Medical Services Advisory Committee* (2009) [95] de Australia y el del *Feasibility Subgroup to Cervical Cytology Review Group* (2009) [96] de Escocia, en los que se fundamentó así mismo el informe del *Europe Against Cancer Programme* para no recomendar la utilización de estos sistemas para el cribado poblacional dada su baja fiabilidad respecto al estudio de las muestras por exminadores humanos expertos.

Con la aplicación de técnicas de IA, se supone que se evitaría la inversión de tiempo en el análisis de muestras normales, y se reduciría la tasa de falsos negativos en relación con errores de análisis en muestras de baja celularidad, o de baja concentración de atipias, o en aquellas en que las atipias se localicen de forma periférica en el cristal portaobjetos. De esta forma, estos sistemas podrían aumentar la detección de lesiones intraepiteliales en estos casos y almacenar imágenes de las células atípicas detectadas, según Kitchener y cols. (2011) [97] y Boost (2009) [98].

En este trabajo se presenta un modelo de análisis de imágenes microscópicas procedentes de una base de datos de creación propia, elaborada a partir de los

archivos de fotografías citológicas escaneadas por el Servicio de Anatomía Patológica del Hospital Universitario Santa María del Rosell-Santa Lucía, de Cartagena (Murcia, España). En un primer momento, tras fragmentar las imágenes de forma automática en células individuales, se procede a realizar una clasificación inicial binaria en dos categorías, células normales y células atípicas, y otra clasificación multiclase siguiendo las categorías diagnósticas del sistema Bethesda.

Para mejorar los resultados obtenidos por esta clasificación inicial y mejorar a su vez la aplicabilidad al análisis de muestras citológicas reales, se plantea la creación de un método de fusión de imágenes, creando solapamientos, plegamientos y cruces entre las distintas células de las imágenes manejadas anteriormente, con transparencia entre ellas, obteniendo un número mayor de imágenes, con distintas categorías diagnósticas entremezcladas que a priori pudiera parecer que disminuirían los resultados de clasificación dado que se aumenta la complejidad de las muestras a estudiar, pero finalmente, los resultados expuestos reflejan una mejoría de los mismos debido al aumento del número de imágenes de la muestra.

6.2.1. Clasificación de Células Originales

Los resultados obtenidos al realizar la clasificación binaria de las células de la base de datos original en las categorías “No requiere revisión” y “Requiere revisión” son bastante satisfactorios, pues se encuentran cercanos a la exactitud media publicada del 79,4% para patólogos expertos en distintos trabajos, como los de Arbyn y cols. (2004)[88], Bigras y cols. (2013) [87], Sorbye y cols. (2017)[90] y Wright y cols. (2014) [89]. Es importante destacar también que la SD

obtenida en nuestro trabajo no es muy alta, lo que indica una cierta estabilidad del modelo de CNN para clasificar células. Por ello, en esta primera parte del experimento se puede indicar que la bonanza de los resultados y la estabilidad del modelo son similares a los obtenidos por patólogos expertos.

Arbyn y cols. (2004) [88] publica una sensibilidad y especificidad del examen citológico cuando es informado como patológico en un 81,8% y 57,6% respectivamente, considerando como patológico la categoría que en nuestro trabajo se ha definido como “Requiere revisión”, es decir, el conjunto de células catalogadas como ASC-US, L-SIL y H-SIL. Sorbye y cols. (2017) [90] publica una sensibilidad media para esta categoría del 77,4%, y una especificidad del 81,3%. Wright y cols. (2014) [89] publica, tras analizar los resultados de cuatro laboratorios, una Sensibilidad que oscila para esta misma categoría diagnóstica entre el 42% y el 73%. Nuestro trabajo presenta una Sensibilidad del 83% y una especificidad del 71%, que comparativamente, no son peores que los publicados.

Si se analiza los resultados con mayor profundidad para valorar los falsos positivos y falsos negativos, se comprueba los aciertos de estos resultados considerando las categorías diagnósticas de las células individuales, es decir, que cada vez que el modelo de CNN obtiene una clasificación binaria acertada en la categoría “No requiere revisión” o “Requiere revisión”, se revisa en qué categoría individual estaba clasificada la imagen.

Los aciertos obtenidos de media por el modelo de CNN para cada categoría individual (Normal, L-SIL, H-SIL, ASC-US) demuestran cómo las categorías mejor clasificadas por el modelo son H-SIL y L-SIL, con un 92% y un 82,7% globales. Por otro lado, las células L-SIL obtienen un porcentaje bajo (10,67%), así como las ASC-US (30,67%). En la detección de células normales y ASC-US es donde ocurre la mayoría de errores, siendo este comportamiento del modelo

valorado por nuestro equipo como positivo, pues detecta de forma muy precisa las células con mayor atipia, que son las que no pueden dejarse escapar en el informe clínico de una citología, es decir, las células que deben ser detectadas con mayor exactitud, debido a las implicaciones clínicas que conllevan, son correctamente clasificadas en su mayoría, mientras que las células con menor atipia y menor grado de malignidad son clasificadas dentro de los grupos diagnósticos de bajo grado de malignidad o normalidad, lo que no reviste grandes consecuencias en el seguimiento de las pacientes de forma global.

La primera dificultad que aparece al enfrentarse a una muestra citológica cervical para entrenar un sistema automático de diagnóstico es la gran heterogeneidad de estirpes celulares que coexisten en la muestra, así como la variabilidad en la flora bacteriana saprófita habitual y de la patógena oportunista, junto a la presencia de artefactos procedentes de la tinción y manejo de la muestra, así como los posibles defectos de enfoque, añadidos a la agrupación y solapamiento existente entre las células. Esta dificultad, inicialmente puede solventarse fragmentando la imagen de la citología inicial en células individuales, y entrenar el sistema en el reconocimiento de células escamosas únicamente, sin reparar en el resto de células que pudiera encontrar. Esta táctica encuentra problemas de clasificación cuando se presenta la situación en la que aparece en la imagen un defecto de enfoque, un solapamiento de células o un artefacto de tinción, que pueda confundirlo y dar lugar a un error en la clasificación.

Por otro lado, cuando se realiza la clasificación multiclase en categorías diagnósticas, se detecta que, al igual que ocurre a un observador humano experimentado, ciertas categorías diagnósticas tienen características similares y en algunos casos son difíciles de diferenciar claramente, como ocurre entre ASC-US y L-SIL, por ejemplo, motivo por el que el estudio citológico tiene un índice kappa bajo,

como refleja Bigras y cols. (2013) [87], índice kappa de 0,59, que indica baja reproducibilidad.

La matriz de confusión de los experimentos realizados con la base de datos original, explica el motivo de la baja tasa de éxito para la clasificación individual de L-SIL, siendo la razón principal por la que en la mitad de los casos son clasificadas como ASC-US, de forma similar a como le puede ocurrir a un observador humano.

Aunque los resultados individuales no son muy satisfactorios para la clasificación multiclase, por la baja tasa de éxito de la clasificación de L-SIL (10.67%), si se analiza globalmente los resultados del modelo de CNN, los errores producidos en la clasificación de L-SIL como ASC-US, clínicamente no supondrían errores, dado que tanto L-SIL como ASC-US deben ser revisados por el patólogo, y en muchos casos, ante estos diagnósticos, las pacientes requieren de la realización de una biopsia para determinar con exactitud cada uno de ellos pues muchas veces la citología por sí sola no es capaz de caracterizar el diagnóstico, así como de la utilización de otros estudios, tales como la colposcopia o la determinación de la presencia del VPH. Recuérdese que la citología es una prueba de cribado, no de diagnóstico definitivo, y la detección de anomalías en la misma, conlleva un estudio en profundidad del caso detectado, lo que en la mayor parte de las ocasiones deriva en la realización de otras pruebas diagnósticas de confirmación.

En resumen, la clasificación multiclase a nivel individual de las células tiene inicialmente unos resultados globales satisfactorios, pues la mayoría de los errores ocurren dentro de las categorías a nivel global de “No requiere revisión”

y “Requiere revisión”, por lo que pueden ser considerados como “Errores sin significación clínica”, y la baja sensibilidad mostrada dentro de las categorías ASC-US y L-SIL puede estar en relación con la similitud existente entre estas células atípicas, que es la que justifica en parte el bajo índice kappa publicado para patólogos expertos.

6.2.2. Clasificación de Células Fusionadas

Mediante la fusión de imágenes se pretende crear un grupo de elementos nuevos a partir de la base de datos original mediante su giro, solapamiento y fusión con transparencia, tal y como suelen aparecer en muestras citológicas reales.

En este apartado se analiza la capacidad de predicción del modelo de CNN cuando se aumenta el número muestral, teniendo en cuenta que estas imágenes contienen células plegadas, borrosas, superpuestas, y manchas de procesamiento. También se valora si el modelo de CMA propuesto introduce errores en el CNN o si ayuda a aumentar la calidad de los aciertos de los resultados, reduciendo falsos negativos por el aumento de especificidad y sensibilidad.

Tras repetir en cinco ocasiones la clasificación del modelo de CNN, la ACC media calculada es 88.8 %, que se encuentra dentro del rango de detección que un patólogo experto puede conseguir, como se publica en diferentes trabajos, como los de Arbyn y cols. (2004) [88], Wright y cols. (2014)[89] y Sorbye y cols. (2017) [90], comentados anteriormente, por lo tanto, se puede afirmar que los modelos de CNN creados pueden obtener unos resultados similares a los obtenidos por patólogos expertos analizando citologías cervicovaginales con tinción

de Papanicolaou.

Si se analiza los resultados de la clasificación multiclase, se observa que, como ocurría en los experimentos anteriores, la tasa de aciertos en la clasificación de células L-SIL es menor que la obtenida con otras células. Sin embargo, en comparación con el experimento realizado con la base de datos original, se obtiene mejores resultados con la muestra obtenida mediante CMA (10,67%, frente al 16,67%), y en la detección de ASC-US se obtiene también una mejoría (30,67%, frente al 21,00%), de forma que los resultados son mejores cuando se incrementa el número de imágenes de la base de datos, aunque las imágenes sean de mayor complejidad por introducir artefactos y solapamientos en las mismas.

Analizando los errores en la matriz de confusión de este método, destaca el hecho de que sólo el 10,7% de células L-SIL se clasifica erróneamente como células Normales, mientras que el resto de los errores se encuentran entre ASC-US y H-SIL. Estos últimos errores en la clasificación de L-SIL, aunque no ayudan a obtener una clasificación exacta, no afectan al hecho de que las células L-SIL sean clasificadas dentro de categorías que deben ser igualmente revisadas por el patólogo, por lo que no se debe considerar estos errores como clínicamente relevantes, pues no alterarán la conducta del patólogo.

Junto con la exactitud, se debe analizar también los valores obtenidos de Desviación Estándar, Sensibilidad y Especificidad. La Desviación Estándar en este experimento es 3 veces menor a la obtenida en el análisis de las imágenes originales, lo que indica que los resultados que ofrecen los modelos de Redes Neuronales Convolucionales aplicados sobre la muestra creada mediante fusión de

imágenes son fiables y estables, por lo que se puede afirmar que los resultados no dependen de la aleatorización. Los valores de Sensibilidad y Especificidad son también mejores en este segundo experimento que en el realizado con las imágenes originales. Esta mejora sustancial de los resultados indica que el modelo de fusión de imágenes refleja una realidad más fiable de las muestras de entrada, lo que se traduce en un mejor comportamiento y ajuste del modelo de redes neuronales convolucionales para clasificar las células en las categorías de “No requiere revisión” y “Requiere revisión”.

Al analizar los aciertos obtenidos en la clasificación binaria de esta base de datos, se observa que la mayor tasa de aciertos se obtiene en la clasificación de células H-SIL, con una tasa de acierto del 98.7 %, siendo un valor altamente significativo. Para L-SIL, ASC-US y células Normales, la tasa de acierto es mayor al 85 % con un umbral de detección bastante satisfactorio, dado que es el mismo nivel de acierto obtenido por patólogos expertos, como se expone en distintos trabajos, como los publicados por Arbyn y cols. (2004) [88], Bigras y cols. (2013) [87], Wright y cols. (2014) [89] y Sorbye y cols. (2017) [90], comentados anteriormente.

Los resultados muestran una tasa de exactitud global del 88,8 %, con una Desviación Estándar del 1 % tras repetir el lanzamiento de la red neuronal convolucional en cinco ocasiones con la muestra obtenida mediante el modelo de fusión de imágenes.

En resumen, el método de fusión de imágenes propuesto genera una base de datos que ofrece unos resultados de clasificación más satisfactorios que cuando utilizamos la base de datos original. Esta mejoría en los resultados se refleja en la tasa de aciertos conseguida por el modelo, con una exactitud global de casi el

89 %, con una sensibilidad del 92 % y especificidad del 83 %, lo que indica que el modelo de CNN junto con el de CMA consigue realizar una clasificación a un nivel comparable a la realizada por un patólogo experto, según publica Sorbye y cols. (2017) [90] de un 77,4 % de especificidad y 81,3 % de especificidad. La estabilidad en la clasificación que realiza el sistema es importante, reflejada por la baja Desviación Estándar (1 %).

CONCLUSIONES

7. CONCLUSIONES

1. La Inteligencia Artificial puede ser tenida en cuenta en el diseño de sistemas de ayuda al diagnóstico médico asistidos por ordenador, en el campo de la clasificación de imágenes ecográficas de tumores de ovario, así como en la detección de atipias en células escamosas cervicales procedentes de citologías cervicovaginales teñidas mediante la técnica de Papanicolaou.
2. Los descriptores Geométricos obtenidos mediante *Fast Fourier Transform* aportan una información útil y relevante para realizar la clasificación de ecografías de tumores de ovario, resultando muy interesante la mejora de los resultados obtenidos en nuestro estudio en comparación con otros estudios anteriores realizados sobre la misma base de datos.
3. Se ha conseguido generar una base de datos mediante fusión de imágenes con transparencia a partir de la base de datos original de células escamosas cervicovaginales, obteniendo una nueva colección con 80.000 imágenes de nueva creación, con mayor complejidad.
4. Al aplicar las técnicas de *Deep Learning* sobre la base de datos de nueva creación mediante fusión de imágenes, se demuestra que aumentando el número de imágenes en la base de datos, con independencia de su complejidad, se mejora los resultados y la estabilidad de la clasificación.

LIMITACIONES DEL ESTUDIO

8. LIMITACIONES DEL ESTUDIO

1. La base de datos de imágenes ecográficas de tumores ováricos utilizada está compuesta por 187 imágenes seleccionadas por el equipo de la Universidad de Buckingham, Reino Unido, siendo un tamaño muestral suficiente para realizar nuestro estudio pero no para poder extrapolar los resultados a otras bases de datos.
2. La relativa facilidad para encontrar células normales en una citología cervicovaginal teñida con la técnica de Papanicolaou hace que este grupo de células sea mayor que cada una de las categorías con atipias celulares, por lo que se hace más difícil conseguir una muestra equilibrada para poder realizar el entrenamiento de la Red Neuronal Convolucional, siendo necesario eliminar de forma aleatoria algunas de las imágenes del grupo de células normales.
3. El método de fusión de imágenes desarrollado permite aumentar el número de imágenes para el entrenamiento y su complejidad, simulando los hallazgos que habitualmente se encuentran en el examen microscópico de una citología cervicovaginal, pero, no obstante, aunque los resultados de clasificación mejoran los obtenidos mediante el estudio de la base de datos de imágenes originales, este sistema de clasificación debería comprobarse con otras bases de datos para poder extrapolar sus resultados.

ANEXO

9. ANEXO

Durante la elaboración de este trabajo se ha publicado diferentes comunicaciones y artículos, cuyas referencias se exponen a continuación:

Comunicaciones:

Comunicación oral con el título “Clasificación de tumoraciones ováricas a partir de imágenes ecográficas mediante su análisis con técnicas de aprendizaje máquina”, presentada en las Jornadas Internacionales de Investigación y Doctorado: Doctorado Industrial de la Escuela Internacional de Doctorado de la UCAM, celebradas el 24 de junio de 2016.

Comunicación oral con el título “¿Tiene la Inteligencia Artificial cabida en el diagnóstico citopatológico?” presentada en el XXIX Congreso Nacional de la Sociedad Española de Anatomía Patológica, y XXIV Congreso Nacional de la Sociedad Española de Citología, celebrado en Granada el 24 de mayo de 2019.

Artículos:

Martínez-Más, J., Bueno-Crespo, A., Imbernón, B., Cecilia, J. M., Martínez-España, R., Remezal-Solano, M., ... y Martínez-Cendán, J. P. (2018, August). Deep Learning Approach for Classifying Papanicolau Cervical Smears. In Proceedings of the 47th International Conference on Parallel Processing Companion (pp. 1-5).

Martínez-Más, J., Bueno-Crespo, A., Khazendar, S., Remezal-Solano, M., Martínez-Cendán, J. P., Jassim, S., ... y Timmerman, D. (2019). Evaluation of machine learning methods with Fourier Transform features for classifying ovarian tumors based on ultrasound images. *PloS one*, 14(7).

Premios:

Premio a la mejor comunicación presentada por el trabajo titulado “Clasificación de tumoraciones ováricas a partir de imágenes ecográficas mediante su análisis con técnicas de aprendizaje máquina” en las Jornadas Internacionales de Investigación y Doctorado: Doctorado Industrial de la Escuela Internacional de Doctorado de la UCAM, celebradas el 24 de junio de 2016.

Premio de a Real Academia de Medicina de la Región de Murcia 2019, en la categoría de Pruebas Diagnósticas, al trabajo titulado “Comparación de modelos de Aprendizaje Automático para la clasificación diagnóstica de imágenes ecográficas de tumores ovárico”, 17 de enero de 2020.

Participación en el Plan Nacional de Investigación titulado “Desarrollo holístico de aplicaciones en sistemas heterogéneos”, con código TIN2016-788799-P, financiado con un importe de 83.853 euros en el marco del Programa Estatal de Fomento de la Investigación Científica y Técnica de Excelencia, Subprograma Estatal de Generación de Conocimiento, con una duración del proyecto de 36 meses, desde el 30 de diciembre de 2016 al 29 de diciembre de 2019.

BIBLIOGRAFÍA

Bibliografía

- [1] Netter FH. Atlas of human anatomy, Professional Edition E-Book: including NetterReference. com Access with full downloadable image Bank. Elsevier Health Sciences, 2014.
- [2] Colomer MC. Cáncer de cuello de útero: nuevas perspectivas en la prevención. *Offarm: farmacia y sociedad* 2008;27(5):120–123.
- [3] Martínez-Más J, Bueno-Crespo A, Khazendar S, Remezal-Solano M, Martínez-Cendán JP, Jassim S, et al. Evaluation of machine learning methods with Fourier Transform features for classifying ovarian tumors based on ultrasound images. *PloS one* 2019;14(7).
- [4] Ortiz-Reina S, Martínez-Más J, Bueno-Crespo A, Remezal-Solano M, Martínez-Cendán JP. ¿Tiene la Inteligencia Artificial cabida en el diagnóstico citopatológico? En XXIX Congreso Nacional de la Sociedad Española de Anatomía Patológica, y XXIV Congreso Nacional de la Sociedad Española de Citología.
- [5] Latarjet M, Liard AR. Anatomía humana. Ed. Médica Panamericana, 2004.
- [6] Sánchez LFP. Endocrinología clínica. Ediciones Díaz de Santos, 2013.
- [7] Ross MH, Wojciech P. Histología. 2013.

- [8] Sellors JW, Sankaranarayanan R. La colposcopia y el tratamiento de la neoplasia intraepitelial cervical: Manual para principiantes. Lyon, Francia: International Agency for Research on Cancer (IARC) 2003;:140.
- [9] Singer A, Monaghan JM, Quek SC. Lower genital tract precancer: colposcopy, pathology and treatment. John Wiley & Sons, 2008.
- [10] Papanicolaou GN, Traut HF. The diagnostic value of vaginal smears in carcinoma of the uterus. *American Journal of Obstetrics & Gynecology* 1941;42(2):193–206.
- [11] Moscicki AB, Schiffman M, Kjaer S, Villa LL. Updating the natural history of HPV and anogenital cancer. *Vaccine* 2006;24:S42–S51.
- [12] Darai E, Teboul J, Walker F, Benifla JL, Meneux E, Guglielmina JN, et al. Epithelial ovarian carcinoma of low malignant potential. *European Journal of Obstetrics & Gynecology and Reproductive Biology* 1996;66(2):141–145.
- [13] Eifel PJ, Gershenson DM, Kavanagh JJ, Silva EG. *Gynecologic cancer*, tomo 5. Springer Science & Business Media, 2006.
- [14] GLOBOCAN I. International Agency for Research on Cancer, GLOBOCAN 2018: Cancer Incidence, Mortality and Prevalence Worldwide in 2018 (Lyon), 2019.
- [15] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians* 2018;68(6):394–424.
- [16] Roldán F. Cáncer de ovario. *Boletín Oncológico*, 1998.

-
- [17] FROM ROD. Collaborative Group on Epidemiological Studies of Ovarian Cancer. DAVID A ZUBIK, ET AL, Petitioners, v SYLVIA BURWELL, ET AL, Respondents 2008;371:74.
- [18] Asociación Española Contra el Cáncer: Indidencia del cáncer de ovario.
- [19] Bray F, Loos AH, Tognazzo S, La Vecchia C. Ovarian cancer in Europe: cross-sectional trends in incidence and mortality in 28 countries, 1953–2000. *International journal of cancer* 2005;113(6):977–990.
- [20] Cabanes A, Vidal E, Pérez-Gómez B, Aragonés N, López-Abente G, Pollán M. Age-specific breast, uterine and ovarian cancer mortality trends in Spain: changes from 1980 to 2006. *Cancer epidemiology* 2009;33(3-4):169–175.
- [21] Red Española de Registros de Cáncer R. Estimaciones de la incidencia del cáncer en España, 2019. REDECAN 2019;.
- [22] Collaborative Group on Epidemiological Studies of Ovarian Cancer CG, et al. Ovarian cancer and oral contraceptives: collaborative reanalysis of data from 45 epidemiological studies including 23 257 women with ovarian cancer and 87 303 controls. *The Lancet* 2008;371(9609):303–314.
- [23] Kurman RJ, Shih IM. The Origin and pathogenesis of epithelial ovarian cancer—a proposed unifying theory. *The American journal of surgical pathology* 2010;34(3):433.
- [24] Bailey C, Ueland F, Land G, DePriest P, Gallion H, Kryscio R, et al. The malignant potential of small cystic ovarian tumors in women over 50 years of age. *Gynecologic oncology* 1998;69(1):3–7.

-
- [25] Cramer DW, Welch WR. Determinants of ovarian cancer risk. II. Inferences regarding pathogenesis. *Journal of the National Cancer Institute* 1983; 71(4):717–721.
- [26] Helzlsouer KJ, Alberg AJ, Gordon GB, Longcope C, Bush TL, Hoffman SC, et al. Serum gonadotropins and steroid hormones and the development of ovarian cancer. *Jama* 1995;274(24):1926–1930.
- [27] Parmley TH, Woodruff JD. The ovarian mesothelioma. *American Journal of Obstetrics & Gynecology* 1974;120(2):234–241.
- [28] Dubeau L. The cell of origin of ovarian epithelial tumors and the ovarian surface epithelium dogma: does the emperor have no clothes? *Gynecologic oncology* 1999;72(3):437–442.
- [29] Jacobs I, Kohler M, Wiseman R, Marks J, Whitaker R, Kerns B, et al. Clonal origin of epithelial ovarian carcinoma: analysis by loss of heterozygosity, p53 mutation, and X-chromosome inactivation. *JNCI: Journal of the National Cancer Institute* 1992;84(23):1793–1798.
- [30] Schorge JO, Muto MG, Welch WR, Bandera CA, Rubin SC, Bell DA, et al. Molecular evidence for multifocal papillary serous carcinoma of the peritoneum in patients with germline BRCA1 mutations. *JNCI: Journal of the National Cancer Institute* 1998;90(11):841–845.
- [31] Berchuck A, Kamel A, Whitaker R, Kerns B, Olt G, Kinney R, et al. Overexpression of HER-2/neu is associated with poor survival in advanced epithelial ovarian cancer. *Cancer research* 1990;50(13):4087–4091.
- [32] Wan M, Li WZ, Felix JC, Zhao Y, Dubeau L, Duggan BD. Telomerase activity in benign and malignant epithelial ovarian tumors. *Journal of the National Cancer Institute* 1997;89(6):437–441.

- [33] Rasila K, Burger R, Smith H, Lee F, Verschraegen C. Angiogenesis in gynecological oncology—mechanism of tumor progression and therapeutic targets. *International Journal of Gynecologic Cancer* 2005;15(5):710–726.
- [34] Stanford JL. Oral contraceptives and neoplasia of the ovary. *Contraception* 1991;43(6):543–556.
- [35] Whittemore AS, Harris R, Ltnyre J, Group COC. Characteristics Relating to ovarian Cancer Risk: Collaborative Analysis of 12 US case-Control Studies: IV. The Pathogenesis of Epithelial Ovarian Cancer. *American journal of epidemiology* 1992;136(10):1212–1220.
- [36] Beral V, Collaborators MWS, et al. Ovarian cancer and hormone replacement therapy in the Million Women Study. *The Lancet* 2007; 369(9574):1703–1710.
- [37] Níguez Sevilla I. La endometriosis atípica como lesión precursora del cáncer de ovario asociado a endometriosis un estudio prospectivo. Proyecto de investigación: 2018;.
- [38] Wiegand KC, Shah SP, Al-Agha OM, Zhao Y, Tse K, Zeng T, et al. ARID1A mutations in endometriosis-associated ovarian carcinomas. *New England Journal of Medicine* 2010;363(16):1532–1543.
- [39] Genkinger JM, Hunter DJ, Spiegelman D, Anderson KE, Arslan A, Beeson WL, et al. Dairy products and ovarian cancer: a pooled analysis of 12 cohort studies. *Cancer Epidemiology and Prevention Biomarkers* 2006; 15(2):364–372.
- [40] Genkinger JM, Hunter DJ, Spiegelman D, Anderson KE, Beeson WL, Buring JE, et al. A pooled analysis of 12 cohort studies of dietary fat, chole-

- terol and egg intake and ovarian cancer. *Cancer Causes & Control* 2006; 17(3):273–285.
- [41] Koushik A, Hunter DJ, Spiegelman D, Anderson KE, Buring JE, Freudenheim JL, et al. Intake of the major carotenoids and the risk of epithelial ovarian cancer in a pooled analysis of 10 cohort studies. *International journal of cancer* 2006;119(9):2148–2154.
- [42] Larsson SC, Giovannucci E, Wolk A. Dietary folate intake and incidence of ovarian cancer: the Swedish Mammography Cohort. *Journal of the National Cancer Institute* 2004;96(5):396–402.
- [43] Galeone C, Pelucchi C, Levi F, Negri E, Franceschi S, Talamini R, et al. Onion and garlic use and human cancer. *The American journal of clinical nutrition* 2006;84(5):1027–1032.
- [44] Meys E, Kaijser J, Kruitwagen R, Slangen B, Van Calster B, Aertgeerts B, et al. Subjective assessment versus ultrasound models to diagnose ovarian cancer: A systematic review and meta-analysis. *European Journal of Cancer* 2016;58:17–29.
- [45] Arenas JMB. *Fundamentos de ginecología*. Ed. Médica Panamericana, 2009.
- [46] Berek J. *Berek & Novak's Gynecology*. Lippincott Williams & Wilkins, 2015.
- [47] Menon U, Gentry-Maharaj A, Hallett R, Ryan A, Burnell M, Sharma A, et al. Sensitivity and specificity of multimodal and ultrasound screening for ovarian cancer, and stage distribution of detected cancers: results of the prevalence screen of the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS). *The lancet oncology* 2009;10(4):327–340.

- [48] Yazbek J, Ameye L, Testa A, Valentin L, Timmerman D, Holland T, et al. Confidence of expert ultrasound operators in making a diagnosis of adnexal tumor: effect on diagnostic accuracy and interobserver agreement. *Ultrasound in Obstetrics & Gynecology* 2010;35(1):89–93.
- [49] Kauff ND, Satagopan JM, Robson ME, Scheuer L, Hensley M, Hudis CA, et al. Risk-reducing salpingo-oophorectomy in women with a BRCA1 or BRCA2 mutation. *New England Journal of Medicine* 2002;346(21):1609–1615.
- [50] Rebbeck TR, Lynch HT, Neuhausen SL, Narod SA, Van't Veer L, Garber JE, et al. Prophylactic oophorectomy in carriers of BRCA1 or BRCA2 mutations. *New England Journal of Medicine* 2002;346(21):1616–1622.
- [51] Lynch HT, De la Chapelle A. Hereditary colorectal cancer. *New England Journal of Medicine* 2003;348(10):919–932.
- [52] Schmeler KM, Lynch HT, Chen Lm, Munsell MF, Soliman PT, Clark MB, et al. Prophylactic surgery to reduce the risk of gynecologic cancers in the Lynch syndrome. *New England Journal of Medicine* 2006;354(3):261–269.
- [53] Jacobs I, Oram D, Fairbanks J, Turner J, Frost C, Grudzinskas J. A risk of malignancy index incorporating CA 125, ultrasound and menopausal status for the accurate preoperative diagnosis of ovarian cancer. *BJOG: An International Journal of Obstetrics & Gynaecology* 1990;97(10):922–929.
- [54] Timmerman D, Valentin L, Bourne T, Collins W, Verrelst H. Vergote I; International Ovarian Tumor Analysis (IOTA) Group. Terms, definitions and measurements to describe the sonographic features of adnexal tumors: a consensus opinion from the International Ovarian Tumor Analysis (IOTA) Group. *Ultrasound Obstet Gynecol* 2000;16:500–505.

- [55] Timmerman D, Testa AC, Bourne T, Ameye L, Jurkovic D, Van Holsbeke C, et al. Simple ultrasound-based rules for the diagnosis of ovarian cancer. *Ultrasound in obstetrics & gynecology* 2008;31(6):681–690.
- [56] Timmerman D, Verrelst H, Bourne T, De Moor B, Collins W, Vergote I, et al. Artificial neural network models for the preoperative discrimination between malignant and benign adnexal masses. *Ultrasound in Obstetrics and Gynecology* 1999;13(1):17–25.
- [57] Timmerman D, Ameye L, Fischerova D, Epstein E, Melis GB, Guerriero S, et al. Simple ultrasound rules to distinguish between benign and malignant adnexal masses before surgery: prospective validation by IOTA group. *Bmj* 2010;341:c6839.
- [58] Kaijser J, Sayasneh A, Van Hoorde K, Ghaem-Maghami S, Bourne T, Timmerman D, et al. Presurgical diagnosis of adnexal tumours using mathematical models and scoring systems: a systematic review and meta-analysis. *Human reproduction update* 2014;20(3):449–462.
- [59] Huerta Sáenz I. Criterios sonográficos de evaluación de malignidad del tumor de ovario Metodología IOTA (International Ovarian Tumor Analysis Group). *Revista Peruana de Ginecología y Obstetricia* 2019;65(3):311–316.
- [60] Van Calster B, Van Hoorde K, Valentin L, Testa AC, Fischerova D, Van Holsbeke C, et al. Evaluating the risk of ovarian cancer before surgery using the ADNEX model to differentiate between benign, borderline, early and advanced stage invasive, and secondary metastatic tumours: prospective multicentre diagnostic study. *Bmj* 2014;349:g5920.

- [61] Sayasneh A, Ferrara L, De Cock B, Saso S, Al-Memar M, Johnson S, et al. Evaluating the risk of ovarian cancer before surgery using the ADNEX model: a multicentre external validation study. *British journal of cancer* 2016;115(5):542–548.
- [62] WHO. Early diagnosis and screening in Cervical Cancer, 2019.
- [63] Hakama M, Rasxnen-Virtaen U. Effect of a mass screening program on the risk of cervical cancer. *American Journal of Epidemiology* 1976; 103(5):512–517.
- [64] Saslow D, Solomon D, Lawson HW, Killackey M, Kulasingam SL, Cain J, et al. American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer. *CA: a cancer journal for clinicians* 2012;62(3):147–172.
- [65] García-Asenjo JAL, Carreño RG. Situación del cribado de cáncer de cérvix en España. *Libro Blanco* 2013;:167–172.
- [66] Sáez A, Gallego I, García-Andrade C, Carrizo R, Romero I, Cortés J, et al. Análisis inicial de eficiencia en la aplicación clínica de las nuevas recomendaciones para la prevención secundaria del cáncer de cuello de útero. *Progresos de Obstetricia y Ginecología* 2013;56(3):140–143.
- [67] International Agency Research on Cancer. Cervix cancer screening. IARC working group on the evaluation of cancer preventive strategies, 2004.
- [68] Pérez-Gómez B, Martínez C, Navarro C, Franch P, Galceran J, Marcos-Gragera R, et al. The moderate decrease in invasive cervical cancer incidence rates in Spain (1980–2004): limited success of opportunistic screening? *Annals of oncology* 2010;21(suppl_3):iii61–iii68.

- [69] Badía X, Tort M, Manganelli AG, Camps C, Díaz-Rubio E. The burden of cancer in Spain. *Clinical and Translational Oncology* 2019;21(6):729–734.
- [70] Kolstad P. Diagnosis and management of precancerous lesions of the cervix uteri. *International Journal of Gynecology & Obstetrics* 1970; 8(4P2):551–560.
- [71] Younge PA. The natural history of carcinoma in situ of the cervix uteri. *BJOG: An International Journal of Obstetrics & Gynaecology* 1965;72(1):9–12.
- [72] Muñoz AR. Correlación entre colposcopia-citología y patología en el diagnóstico precoz del Ca de cérvix. *Colombia Médica* 1972;3(4):153–156.
- [73] Luthra U, Prabhakar A, Seth P, Agarwal S, Murthy N, Bhatnagar P, et al. Natural history of precancerous and early cancerous lesions of the uterine cervix. *Acta Cytologica* 1987;31(3):226–234.
- [74] Holowaty P, Miller AB, Rohan T, To T. Natural history of dysplasia of the uterine cervix. *Journal of the National Cancer Institute* 1999;91(3):252–258.
- [75] Cuestas E. El Premio Nobel de fisiología o medicina 2008 ;.
- [76] Anttila A, Kotaniemi-Talonen L, Leinonen M, Hakama M, Laurila P, Tarkkanen J, et al. Rate of cervical cancer, severe intraepithelial neoplasia, and adenocarcinoma in situ in primary HPV DNA screening with cytology triage: randomised study within organised screening programme. *Bmj* 2010;340:c1804.
- [77] Nayar R, Wilbur DC. The Pap test and Bethesda 2014. *Acta Cytologica* 2015;59(2):121–132.

- [78] Pelea CL, González JF. *Citología ginecológica de Papanicolaou a Bethesda*. Editorial Complutense, 2003.
- [79] Solomon D, Davey D, Kurman R, Moriarty A, O'Connor D, Prey M, et al. The 2001 Bethesda System: terminology for reporting results of cervical cytology. *JAMA* 2002;287(16):2114–2119.
- [80] National Cancer Institute. *Treatment Summary for Health Professionals*, 2002.
- [81] Darragh TM, Colgan TJ, Cox JT, Heller DS, Henry MR, Luff RD, et al. The lower anogenital squamous terminology standardization project for HPV-associated lesions: background and consensus recommendations from the College of American Pathologists and the American Society for Colposcopy and Cervical Pathology. *Archives of pathology & laboratory medicine* 2012;136(10):1266–1297.
- [82] IARC. *Monographs on the evaluation of carcinogenic risks to humans, vol 90: Human Papillomaviruses*. Lyon: International Agency for Research in Cancer 2007;.
- [83] McKeage K, Lyseng-Williamson KA. 9-Valent human papillomavirus recombinant vaccine (Gardasil® 9): a guide to its use in the EU. *Drugs & Therapy Perspectives* 2016;32(10):414–421.
- [84] Hartwig S, Baldauf JJ, Dominiak-Felden G, Simondon F, Alemany L, de Sanjosé S, et al. Estimation of the epidemiological burden of HPV-related anogenital cancers, precancerous lesions, and genital warts in women and men in Europe: potential additional benefit of a nine-valent second generation HPV vaccine compared to first generation HPV vaccines. *Papillomavirus Research* 2015;1:90–100.

- [85] Ismail SM, Colclough AB, Dinnen JS, Eakins D, Evans D, Gradwell E, et al. Observer variation in histopathological diagnosis and grading of cervical intraepithelial neoplasia. *BMJ* 1989;298(6675):707–710.
- [86] Stoler MH, Schiffman M, et al. Interobserver reproducibility of cervical cytologic and histologic interpretations: realistic estimates from the ASCUS-LSIL Triage Study. *JAMA* 2001;285(11):1500–1505.
- [87] Bigras G, Wilson J, Russell L, Johnson G, Morel D, Saddik M. Interobserver concordance in the assessment of features used for the diagnosis of cervical atypical squamous cells and squamous intraepithelial lesions (ASC-US, ASC-H, LSIL and HSIL). *Cytopathology* 2013;24(1):44–51.
- [88] Arbyn M, Buntinx F, Ranst MV, Paraskevaidis E, Martin-Hirsch P, Dillner J. Virologic versus cytologic triage of women with equivocal Pap smears: a meta-analysis of the accuracy to detect high-grade intraepithelial neoplasia. *Journal of the National Cancer Institute* 2004;96(4):280–293.
- [89] Wright Jr TC, Stoler MH, Behrens CM, Sharma A, Sharma K, Apple R. Interlaboratory variation in the performance of liquid-based cytology: Insights from the ATHENA trial. *International Journal of Cancer* 2014;134(8):1835–1843.
- [90] Sørbye SW, Suhrke P, Revå BW, Berland J, Maurseth RJ, Al-Shibli K. Accuracy of cervical cytology: comparison of diagnoses of 100 Pap smears read by four pathologists at three hospitals in Norway. *BMC Clinical Pathology* 2017;17(1):18.
- [91] Ronco G, Cuzick J, Pierotti P, Cariaggi MP, Dalla Palma P, Naldoni C, et al. Accuracy of liquid based versus conventional cytology: overall results

- of new technologies for cervical cancer screening: randomised controlled trial. *BMJ* 2007;335(7609):28.
- [92] Arbyn M, Bergeron C, Klinkhamer P, Martin-Hirsch P, Siebers AG, Bulten J. Liquid compared with conventional cervical cytology: a systematic review and meta-analysis. *Obstetrics & Gynecology* 2008;111(1):167–177.
- [93] Montuenga Badía L, Ruiz E, Francisco J, Calvo González A. Técnicas en histología y biología celular. 611.018 611.018076 M6T4 MON. 2009.
- [94] Alves VAF, Bibbo M, Schmitt FCL, Milanezi F, Longatto Filho A. Comparison of manual and automated methods of liquid-based cytology. *Acta Cytologica* 2004;48(2):187–193.
- [95] Medical Services Advisory C. Automation assisted and liquid based cytology for cervical cancer screening. MSAC reference 2009;1122.
- [96] from the Feasibility Subgroup to Cervical Cytology Review Group R. Scottish Cervical Cytology Review Group Feasibility Subgroup: Cervical Cytology ThinPrep Imager (TIS) Feasibility Study, 2009.
- [97] Kitchener HC, Blanks R, Cubie H, Desai M, Dunn G, Legood R, et al. MAVARIC, A comparison of automation-assisted and manual cervical screening: a randomised controlled trial. *Clinical Governance: An International Journal* 2011;16(3).
- [98] Boost T. A comparison of screening times between the ThinPrep Imager and conventional cytology. *Diagnostic Cytopathology* 2009;37(9):661–664.
- [99] Verdes MTQ, Riobóo JP. Métodos automatizados de lectura de citología cervical uterina. Axencia de Avaliación de Tecnoloxías Sanitarias de Galicia, 2013.

- [100] Koss LG, Lin E, Schreiber K, Elgert P, Mango L. Evaluation of the PAPNET™ cytologic screening system for quality control of cervical smears. *American Journal of Clinical Pathology* 1994;101(2):220–229.
- [101] Patten Jr SF, Lee JS, Wilbur DC, Bonfiglio TA, Colgan TJ, Richart RM, et al. The AutoPap 300 QC System multicenter clinical trials for use in quality control rescreening of cervical smears: A prospective intended use study. *Cancer Cytopathology: Interdisciplinary International Journal of the American Cancer Society* 1997;81(6):337–342.
- [102] Arbyn M, Anttila A, Jordan J, Ronco G, Schenck U, Segnan N, et al. European guidelines for quality assurance in cervical cancer screening. Summary document. *Annals of Oncology* 2010;21(3):448–458.
- [103] Ng A. Advice for applying machine learning. En *Machine learning*.
- [104] Khazendar S, Al-Assam H, Du H, Jassim S, Sayasneh A, Bourne T, et al. Automated classification of static ultrasound images of ovarian tumours based on decision level fusion. En *Computer Science and Electronic Engineering Conference (CEEC), 2014 6th. IEEE*, 148–153.
- [105] Khazendar S, Sayasneh A, Al-Assam H, Du H, Kaijser J, Ferrara L, et al. Automated characterisation of ultrasound images of ovarian tumours: the diagnostic accuracy of a support vector machine and image processing with a local binary pattern operator. *Facts, views & vision in ObGyn* 2015; 7(1):7.
- [106] Khazendar S. *Computer-aided Diagnosis of Gynaecological Abnormality using B-mode Ultrasound Images*, February 2016.
- [107] LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521(7553):436.

-
- [108] Jurtz VI, Johansen AR, Nielsen M, Almagro Armenteros JJ, Nielsen H, Sønderby CK, et al. An introduction to Deep Learning on biological sequence data: examples and solutions. *Bioinformatics* 2017;33(22):3685–3690.
- [109] Bernal J, Kushibar K, Asfaw DS, Valverde S, Oliver A, Martí R, et al. Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: a review. *Artificial Intelligence in Medicine* 2018;;In Press.
- [110] Ahmad J, Muhammad K, Baik SW. Medical image retrieval with compact binary codes generated in frequency domain using highly reactive convolutional features. *Journal of Medical Systems* 2018;42(2):24.
- [111] LeCun Y, Ranzato M. Deep learning tutorial. En *Tutorials in International Conference on Machine Learning (ICML'13)*. Citeseer, 1–29.
- [112] Agarap AF. Deep learning using rectified linear units (relu). arXiv preprint arXiv:180308375 2018;.
- [113] García-Laencina P, Rodríguez-Bermúdez G, Roca-Dorda J. Feed-Forward Neural Network Architectures Based on Extreme Learning Machine for Parkinson's Disease Diagnosis. En *XIII Mediterranean Conference on Medical and Biological Engineering and Computing* 2013. Springer, 746–749.
- [114] García-Laencina PJ, Abreu PH, Abreu MH, Afonoso N. Missing data imputation on the 5-year survival prediction of breast cancer patients with unknown discrete values. *Computers in biology and medicine* 2015; 59:125–133.
- [115] Goodfellow I, Bengio Y, Courville A, Bengio Y. *Deep learning*, tomo 1. MIT press Cambridge, 2016.

-
- [116] Godinez WJ, Hossain I, Lazic SE, Davies JW, Zhang X. A multi-scale convolutional neural network for phenotyping high-content cellular images. *Bioinformatics* 2017;33(13):2010–2019.
- [117] Moeskops P, Viergever MA, Mendrik AM, de Vries LS, Benders MJ, Išgum I. Automatic segmentation of MR brain images with a convolutional neural network. *IEEE Transactions on Medical Imaging* 2016;35(5):1252–1261.
- [118] Sun C, Guo S, Zhang H, Li J, Chen M, Ma S, et al. Automatic segmentation of liver tumors from multiphase contrast-enhanced CT images based on FCNs. *Artificial Intelligence in Medicine* 2017;83:58–66.
- [119] Arevalo J, Cruz-Roa A, Arias V, Romero E, González FA. An unsupervised feature learning framework for basal cell carcinoma image analysis. *Artificial Intelligence in Medicine* 2015;64(2):131–145.
- [120] Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017;542(7639):115.
- [121] Behrmann J, Etmann C, Boskamp T, Casadonte R, Kriegsmann J, Maaß P. Deep Learning for tumor classification in imaging mass spectrometry. *Bioinformatics* 2017;34(7):1215–1223.
- [122] Motlagh NH, Jannesary M, Aboulkheyr H, Khosravi P, Elemento O, Tononchi M, et al. Breast Cancer Histopathological Image Classification: A Deep Learning Approach. *bioRxiv* 2018;:242818.
- [123] Zhang L, Lu L, Nogues I, Summers RM, Liu S, Yao J. DeepPap: Deep convolutional networks for cervical cell classification. *IEEE Journal of Biomedical and Health Informatics* 2017;21(6):1633–1643.

-
- [124] Gonzales RC, Woods RE. *Digital Image Processing*, Addison & Wesley Publishing Company. Reading, MA 1992;.
- [125] Alpaydm E. *Introduction to Machine Learning* 2nd edition Ed, 2010.
- [126] Bishop C. *Bishop Pattern Recognition and Machine Learning*, 2010.
- [127] Rodriguez-Bermudez G, Garcia-Laencina PJ, Roca-Dorda J. Efficient automatic selection and combination of eeg features in least squares classifiers for motor imagery brain–computer interfaces. *International journal of neural systems* 2013;23(04):1350015.
- [128] Pao YH, Park GH, Sobajic DJ. Learning and Generalization Characteristics of the Random Vector Functional-Link Net. *Neurocomputing* 1994; 6(2):163–180.
- [129] IgelNIK B, Pao YH. Stochastic Choice of Basis Functions in Adaptive Function Approximation and the Functional-Link Net. *IEEE Transactions on Neural Networks* 1997;8(2):452–454.
- [130] Guang-Bin GB, Chen L. Convex incremental extreme learning machine. *Neurocomputing* 2007;70(16):3056–3062.
- [131] Huang GB, Wang D, Lan Y. Extreme learning machines: a survey. *International Journal of Machine Learning and Cybernetics* 2011;2(2):107–122.
- [132] Serre D. *Matrices: Theory and Applications*. 2002.
- [133] Huang GB, Zhu QY, Siew C. Extreme learning machine: theory and applications. *Neurocomputing* 2006;70(1):489–501.
- [134] Rong HJ, Ong YS, Tan AH, Zhu Z. A fast pruned-extreme learning machine for classification problem. *Neurocomputing* 2008;72(1):359–366.

-
- [135] Miche Y, Bas P, Jutten C, Simula O, Lendasse A. A Methodology for Building Regression Models using Extreme Learning Machine: OP-ELM. En ESANN. 247–252.
- [136] Miche Y, Sorjamaa A, Lendasse A. OP-ELM: theory, experiments and a toolbox. En International Conference on Artificial Neural Networks. Springer, 145–154.
- [137] Mateo F, Lendasse A. A variable selection approach based on the delta test for extreme learning machine models. En Proceedings of the European symposium on time series prediction. 57–66.
- [138] Miche Y, Lendasse A. A faster model selection criterion for OP-ELM and OP-KNN: Hannan-Quinn criterion. En ESANN, tomo 9. 177–182.
- [139] Similä T, Tikka J. Multiresponse sparse regression with application to multidimensional scaling. En International Conference on Artificial Neural Networks. Springer, 97–102.
- [140] Jia Y, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, et al. Caffe: Convolutional architecture for fast feature embedding. En Proceedings of the 22nd ACM international conference on Multimedia. ACM, 675–678.
- [141] Kaijser J, Van Gorp T, Van Hoorde K, Van Holsbeke C, Sayasneh A, Vergote I, et al. A comparison between an ultrasound based prediction model (LR2) and the risk of ovarian malignancy algorithm (ROMA) to assess the risk of malignancy in women with an adnexal mass. *Gynecologic oncology* 2013;129(2):377–383.
- [142] Van Calster B, Van Hoorde K, Froyman W, Kaijser J, Wynants L, Landolfo C, et al. Practical guidance for applying the ADNEX model from the IO-

TA group to discriminate between different subtypes of adnexal tumors. Facts, views & vision in ObGyn 2015;7(1):32.

- [143] Timmerman D, Schwärzler P, Collins W, Claerhout F, Coenen M, Amant F, et al. Subjective assessment of adnexal masses with the use of ultrasonography: an analysis of interobserver variability and experience. *Ultrasound in Obstetrics and Gynecology: The Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology* 1999;13(1):11–16.

