



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



ESCUELA TÉCNICA
SUPERIOR INGENIERÍA
INDUSTRIAL VALENCIA

TRABAJO FIN DE MASTER EN INGENIERÍA BIOMÉDICA

DISEÑO Y DESARROLLO DE UN SISTEMA AUTOMÁTICO BASADO EN ALGORITMOS DE "DEEP LEARNING" PARA IDENTIFICAR DISTINTOS GRADOS DE TUMOR " BUDDING" EN CÁNCER DE VEJIGA

AUTORA: Anna Esteve Domínguez

TUTORA: Valeriana, Naranjo Ornedo

COTUTORE: José Gabriel García Pardo
David Ramos Soler

Curso Académico: 2019-20

Agradecimientos

“ A Valery y a Gabri por depositar su confianza en mí, a querer realizar este proyecto juntos

A todos los profesores, ya sea a pequeñas o grandes dosis de docencia, que han hecho que a día de hoy sea capaz de ejercer como ingeniera biomédica.

A mi familia, por la incalculable paciencia que tienen conmigo y sobre todo con mi cabezonería de ser ingeniera biomédica.

A mi luz, mi flor blanca, mi todo, por ser la motivación de todo y el motivo la cual me entrego tanto para ser la mejor ingeniera biomédica del mundo.

Y por último y no menos importante, mi ejemplo a seguir mi prima Amparo Esther, por creer siempre en mí aún teniendo todo patas arriba. Por dar todo de ti para motivarme a continuar luchando. Sin ti probablemente no habría llegado tan lejos.

A todos muchas gracias.”

Resumen

El patrón de crecimiento tumor *budding* (TB) es una característica histológica que, por medio de una transición epitelio mesénquima en el frente tumoral, le confiere una alta agresividad y la gran capacidad de invadir tejidos y de provocar metástasis, dando lugar a un mal pronóstico del cáncer. Este patrón ha demostrado ser importante en los modelos tumorales más incidentes.

En el presente trabajo, se pretenden aplicar las técnicas de *deep learning* para diseñar y desarrollar algoritmos que sean capaces de detectar automáticamente los principales patrones de crecimiento en el cáncer de vejiga a partir una imagen histológica. Dichos patrones están ligados al grado de malignidad del tumor, por lo tanto, el principal propósito del proyecto es proporcionar a los expertos una herramienta que ayude a obtener un diagnóstico seguro mediante la clasificación automática del tumor.

En cuanto a la metodología empleada para la tesis, se desarrollan varias funciones para llevar a cabo el preprocesado de la base de datos con el fin de conseguir una base de datos acondicionada para los modelos. Posteriormente, se aplican algoritmos, tanto supervisados como no supervisados, que sean capaces de clasificar la presencia del cáncer y, además, diagnosticar la agresividad.

Finalmente, con la totalidad de los resultados conseguidos, se evalúa su funcionalidad mediante análisis cualitativos y cuantitativos de los modelos, y se realiza una comparación entre los diferentes resultados y modelos implementados.

Anna Esteve Domínguez

anesdo@etsii.upv.es

Palabras Clave: *Tumor budding*, cáncer de vejiga, imágenes histológicas, clasificación supervisada, Redes convolucionales neuronales, clasificación no supervisada, *autoencoder*, patrones de crecimiento.

Resum

El patró de creixement tumor *budding* (TB) és una característica histològica que, per mitjà d'una transició epiteli mesènquima en el front tumoral, li confereix una alta agressivitat i la gran capacitat d'envair teixits i de provocar metàstasis, donant lloc a un mal pronòstic del càncer. Aquest patró ha demostrat ser important en els models tumorals més incidents.

En el present treball, es pretenen aplicar les tècniques de *deep learning* per a dissenyar i desenvolupar algorismes que siguin capaços de detectar automàticament els principals patrons de creixement en el càncer de bufeta a partir una imatge histològica. Aquests patrons estan lligats al grau de malignitat del tumor, per tant, el principal propòsit del projecte és proporcionar als experts una eina que ajude a obtenir un diagnòstic segur mitjançant la classificació automàtica del tumor.

Quant a la metodologia emprada per a la tesi, es desenvolupen diverses funcions per a dur a terme el preprocessat de la base de dades amb la finalitat d'aconseguir una base de dades condicionada per als models. Posteriorment, s'apliquen algorismes, tant supervisats com no supervisats, que siguin capaços de classificar la presència del càncer i, a més, diagnosticar l'agressivitat.

Finalment, amb la totalitat dels resultats aconseguits, s'avalua la seua funcionalitat mitjançant anàlisis qualitatives i quantitatives dels models, i es realitza una comparació entre els diferents resultats i models implementats.

Anna Esteve

anesdo@etsii.upv.es

Paraules clau: Tumor *budding*, càncer de bufeta, imatges histològiques, classificació supervisada, Xarxes convolucionals neuronals, classificació no supervisada, *autoencoder*, patrons de creixement.

Abstract

The growth pattern of tumor budding (TB) is a histologic characteristic that, by means of a mesenchyme epithelium transition in the tumor front, confers a higher aggressiveness and the great ability to invade tissues and cause metastasis, leading to a poor prognosis of cancer. This pattern has been shown to be important in the most incident tumor models.

In this paper, it is intended to apply deep learning techniques to design and develop algorithms that are able to automatically detect the main growth patterns in bladder cancer from a histological image. These patterns are linked to the degree of malignancy of the tumor, so the main purpose of the project is to provide experts with a tool to help obtain a safe diagnosis by automatically classifying the tumor.

Regarding the methodology used for the thesis, several functions are developed to carry out the pre-processing of the database in order to obtain a database conditioned for the models. Subsequently, both supervised and unsupervised algorithms are applied that are able to classify the presence of cancer and, in addition, diagnose aggression.

Finally, with all the results obtained, its functionality is evaluated by qualitative and quantitative analysis of the models, and a comparison is made between the different results and models implemented.

Anna Esteve
anesdo@etsii.upv.es

Keywords: Tumor budding, bladder cancer, histological imaging, supervised classification, neural convolutional networks, unsupervised classification, autoencoder, growth patterns.

Índice general

I Memoria	1
1 Introducción	3
1.1 Motivación y descripción del problema	4
1.2 Objetivos	18
1.3 Guía de la memoria	19
2 Material	21
2.1 Base de datos.	22
2.2 Software y hardware	23
3 Metodología	25
3.1 Preparación de la base de datos de imágenes	27
3.2 Groundtruth para la clasificación	29
3.3 Partición de los datos	33
3.4 Clasificación supervisada	34
3.5 Clasificación no supervisada.	43
4 Resultados y discusión	49
4.1 Experimentos	50
4.2 Métricas de evaluación	51
4.3 Resultados del modelo supervisado.	54
4.4 Resultados del modelo no supervisado.	59
4.5 Estado del arte	62
5 Conclusiones y líneas futuras	65
5.1 Conclusiones.	65
5.2 Líneas futuras	66

Bibliografía	67
II Presupuesto	73
1 Presupuesto	3
1.1 Presupuestos parciales	3
1.2 Presupuestos total	6

Índice de figuras

1.1. Estimación de los tumores más frecuentemente diagnosticados en el mundo. Estimación para el año 2018, ambos sexos. Fuente: [5]	5
1.2. Estadios del cáncer dependientes del crecimiento. Obtenida en [7]	6
1.3. Incidencia a nivel mundial del cáncer de vejiga en la población masculina. Fuente: [4] . .	7
1.4. Incidencia a nivel mundial del cáncer de vejiga en la población femenina. Fuente: [4] . .	7
1.5. Incidencia en el año 2018 a nivel mundial del cáncer de vejiga en hombres. Fuente: [4] .	8
1.6. Incidencia en el año 2018 a nivel mundial del cáncer de vejiga en mujeres. Fuente: [4] . .	8
1.7. Estimación del número de casos de incidencia desde el 2018 a 2040 a nivel nacional, para el cancer de vejiga, población masculina y todas las edades. Fuente: [4]	9
1.8. Estimación del número de casos de incidencia desde el 2018 a 2040 a nivel nacional, para el cancer de vejiga, población femenina y todas las edades. Fuente: [4]	9
1.9. Estimación del número de casos de mortalidad desde el 2018 a 2040 a nivel nacional, para el cancer de vejiga, población masculina y todas las edades. Fuente: [4]	10
1.10. Estimación del número de casos de mortalidad desde el 2018 a 2040 a nivel mundial, para el cancer de vejiga, población femenina y todas las edades. Fuente: [4]	10
1.11. Estimación del número de casos de mortalidad desde el 2018 a 2040 a nivel mundial, para el cancer de vejiga, población masculina y todas las edades. Fuente: [4]	11
1.12. Estimación del número de casos de mortalidad desde el 2018 a 2040 a nivel nacional, para el cancer de vejiga, población femenina y todas las edades. Fuente: [4]	11
1.13. Clasificación del estadio TNM de la Unión Internacional Contra el Cáncer (UICC) para el cáncer de vejiga. Fuente: [13]	13
1.14. Esquema de preparación de las imágenes histológicas para su análisis bajo el microscopio, a partir de una biopsia. Elaboración propia.	15
1.15. Zona de frente tumoral invasor a pequeños aumentos (4×). (a) Tinción HE. (b) Tinción inmunohistquímica CK AE1/3. Fuente: [8].	15

1.16. Representación de los tres patrones de crecimiento propuestos por Jimenez et al. [14]. I) Patrón nodular; II) Patrón trabecular; III) Patrón infiltrativo. Fuente: [8]	16
1.17. Ejemplos de patrones de crecimiento de cáncer de vejiga con imágenes histológicas reales. (a) Sano, sin presencia de cáncer. (b) Patrón nodular. (c) Patrón trabecular.(d) Patrón infiltrativo. Elaboración propia.	17
1.18. Zona de frente tumoral invasor a pequeños aumentos (4×) y zona destacada por la presencia del tumor <i>budding</i> a 40×. Fuente: [8].	17
2.1. A) Escáner Leyca Biosystems Aperio CS2 [19]. B) Software <i>Aperio ImageScope</i> . Elaboración propia. Dispositivo y programa empleado para la adquisición de imágenes.	22
3.1. Diagrama de flujo correspondiente a la metodología de este proyecto.	26
3.2. Preprocesado de la base de datos. Elaboración propia.	27
3.3. Esquema correspondiente a la anotación y extracción de las regiones de interés. Elaboración propia.	28
3.4. Pasos realizados para el descarte de los bloques con información menor al 35%. Elaboración propia.	29
3.5. Esquema que reproduce la zona de estudio del TB. Fuente: [8]	30
3.6. Frontera tumoral con diferentes patrones.	31
3.7. Distribución de clases en el <i>groundtruth</i>	32
3.8. Diagrama de flujo del procedimiento realizado para la preparación de los datos.	33
3.9. Dibujo explicativo de la metodología <i>kfold cross validation</i> con <i>K</i> valor igual a cinco. Elaboración propia.	33
3.10. Tabla que describe el número de imágenes por clase que hay en cada set elaborado por el método de <i>k fold cross validation</i> . Elaboración propia.	34
3.11. Similitud con las neuronas cerebrales. Fuente: [25]	35
3.12. Esquema de un ANN de N capas, una capa de entrada, N-2 capa oculta y una de salida. Elaboración propia.	36
3.13. (a) Representaciones de cómo busca el mínimo global de la función de pérdidas con diferentes valores de <i>learning rate</i> . (b) Cómo responde la función de pérdidas con la elección del valor del <i>learning rate</i> con cada época. Fuente: [26]	37
3.14. Tabla con los síntomas, remedios y efectos que provoca en problemas de regresión, clasificación y aprendizaje profundo cuando se entrenan pocas épocas <i>Underfitting</i> , número de épocas correcto y muchas épocas, <i>overfitting</i> . Fuente: [27]	38
3.15. Esquema del funcionamiento del optimizador para encontrar el error mínimo de la función de pérdidas. Fuente: [24]	39
3.16. Representación del funcionamiento de la convolución. Fuente: [29]	40
3.17. Tipos de muestreo. Elaboración propia	40

3.18. Arquitectura convolucional. Elaboración propia.	41
3.19. Familia de VGG. Fuente: [30]	42
3.20. Arquitectura de VGG19. Fuente: [33]	42
3.21. Esquema del algoritmo desarrollado para la clasificación no supervisada.	44
3.22. Arquitectura autoencoder diseñada para llevar a cabo la reconstrucción de las imágenes histológicas Xj.	45
3.23. Esquema del funcionamiento de la técnica <i>kmeans</i> . (a) Inicialización de los centroides. (b) Paso de asignación de los centroides iniciales a los datos. (c) Paso de la actualización de la posición de los centroides.	47
4.1. Esquema del funcionamiento de la técnica CAM. Elaboración propia	51
4.2. Representación de la matriz de confusión. Elaboración propia	52
4.3. Ejemplo de Curva ROC y área bajo la curva (AUC). Fuente : [37]	53
4.4. Matriz de confusión de los mejores modelos generados por los subconjuntos en la etapa de validación para el modelo supervisado. (a) <i>kfold</i> =1. (b) <i>kfold</i> =2. (c) <i>kfold</i> =3. (d) <i>kfold</i> =4. (e) <i>kfold</i> =5	55
4.5. Representación de las curvas de aprendizaje durante el entrenamiento para el modelo supervisado. (a) Valor de la exactitud y de la función de pérdidas en cada época. (a) <i>kfold</i> =1. (b) <i>kfold</i> =2. (c) <i>kfold</i> =3. (d) <i>kfold</i> =4. (e) <i>kfold</i> =5.	55
4.6. Resultados de los CAMs para los subconjunto dela etapa de validación. (a) Grado III. (b) Grado II. (d) Sano.	56
4.7. Matriz de confusión en la etapa de predicción para el modelo supervisado.	57
4.8. Resultados de los CAMs para los subconjunto de la etapa de validación. (a) Grado III. (b) Grado II. (d) Sano.	58
4.9. Matriz de confusión de los mejores modelos generados por los subconjuntos en la etapa de validación para el modelo no supervisado. (a) <i>kfold</i> =1. (b) <i>kfold</i> =2. (c) <i>kfold</i> =3. (d) <i>kfold</i> =4. (e) <i>kfold</i> =5	59
4.10. Representación de las curvas de aprendizaje durante el entrenamiento para el modelo no supervisado. (a) Valor de la exactitud y de la función de pérdidas en cada época. (a) <i>kfold</i> =1. (b) <i>kfold</i> =2. (c) <i>kfold</i> =3. (d) <i>kfold</i> =4. (e) <i>kfold</i> =5.	60
4.11. Matriz de confusión en la etapa de predicción para el modelo no supervisado.	61

Índice de cuadros

3.1. Tabla de número de imágenes por grado.	31
4.1. Promedio de las métricas de los modelos obtenidos por los subconjuntos en la etapa de validación para el modelo supervisado.	54
4.2. Métricas del modelo definitivo en la etapa de predicción para el modelo supervisado. . .	57
4.3. Promedio de las métricas de los modelos obtenidos por los subconjuntos en la etapa de validación para el modelo no supervisado.	59
4.4. Métricas del modelo definitivo en la etapa de predicción para el modelo no supervisado. .	60
4.5. Métricas del modelo definitivo en la etapa de predicción para el modelo no supervisado. .	61
1.1. Descomposición de los costes de mano de obra.	4
1.2. Descomposición de los costes de los materiales.	4
1.3. Descomposición de los costes del <i>hardware</i>	5
1.4. Descomposición de los costes del <i>software</i>	5
1.5. Coste total de la maquinaria	5
1.6. Presupuesto total del proyecto.	6

Listado de abreviaturas principales

OMS	Organización Mundial de la Salud
ADN	Ácido desoxirribonucleico
IARC	Agencia Internacional para la Investigación del Cáncer
SEOM	Sociedad Española de Oncología Médica
AECC	Asociación Española Contra el Cáncer
CVMI	Crecimiento músculo invasivo
CVNMI	Crecimiento no músculo invasivo
MRI	Magnetic Resonance Imaging
DRE	Digital del Recto
TURBT	Transurethral Resection of Bladder Tumor
HE	Hematoxilina-Eosina
ROI	Región de interés
TB	Tumor <i>Budding</i> .
TFM	Trabajo final de máster
MLP	Multi Layer Perceptron
RNA	Red Neuronal Artificial
SGD	Descenso por Gradiente Estocástico
CNN	Redes Neuronales Convolucionales
VGG	Visual Geometry Group

CAM	Class Activation Mapping
R	Recall
NPV	Negative Predictive Value
AUC	Area Under Curve

Parte I

Memoria

Capítulo 1

Introducción

Índice de contenidos

1.1 Motivación y descripción del problema	4
1.1.1 Cáncer, cáncer de vejiga y tumor <i>budding</i>	4
1.1.2 Procedimiento de diagnóstico.	12
1.2 Objetivos	18
1.3 Guía de la memoria	19

1.1 Motivación y descripción del problema

1.1.1 Cáncer, cáncer de vejiga y tumor budding

Cáncer

El cáncer, según la Organización Mundial de la Salud (OMS), es un proceso incontrolado de crecimiento y de diseminación de células. Esta alteración tiene la capacidad de aparecer prácticamente en cualquier lugar del cuerpo e invadir el tejido circundante o los vasos linfáticos y/o hemáticos provocando el crecimiento de un tumor en puntos distantes del organismo, lo que se conoce como metástasis [1]. Hoy en día se registran más de 200 tipos de cáncer donde se diferencian en 23 grupos según su localización. Cada uno de ellos tiene unas propiedades distintas que los categoriza como enfermedades independientes [2].

Existen dos tipos de factores de riesgo que han demostrado estar asociados con un mayor crecimiento tumoral. Por un lado, están aquellos que se controlan como los malos hábitos para la salud. Y por otro lado, están los factores inevitables como son los artefactos del ADN, es decir, las mutaciones hereditarias. Evitar la exposición a estos factores de riesgo, en la medida de lo posible, podría conducir a una disminución de la incidencia y, por consiguiente, de la mortalidad [3].

Es cierto que, una vez desarrollado el cáncer, si se detecta en fases tempranas, éste podría eliminarse empleando los tratamientos más comunes como la quimioterapia, la radioterapia y la cirugía. Sin embargo, si el diagnóstico del cáncer tiene lugar en una fase avanzada, su tratamiento se complica. Este tipo de casos se asocia a un aumento de la morbilidad y de la mortalidad, además de a una menor esperanza de vida. Para determinar en qué fase se encuentra y poder hallar un tratamiento personalizado, es necesario llevar a cabo un análisis detallado a nivel tisular.

Según las estadísticas recogidas por la Agencia Internacional para la Investigación del Cáncer (IARC), se prevé que el cáncer aumente a nivel mundial más de un 60%, en términos de incidencia y mortalidad en los próximos 20 años [4]. Esta es la razón por la que es necesaria la investigación y el desarrollo de sistemas de ayuda al diagnóstico para identificar el cáncer en sus fases iniciales, permitiendo así un diagnóstico precoz y, por ende, un tratamiento temprano que permita una disminución de la mortalidad. A continuación, se muestran en la Figura 1.1 unos gráficos sobre la incidencia en el 2018.

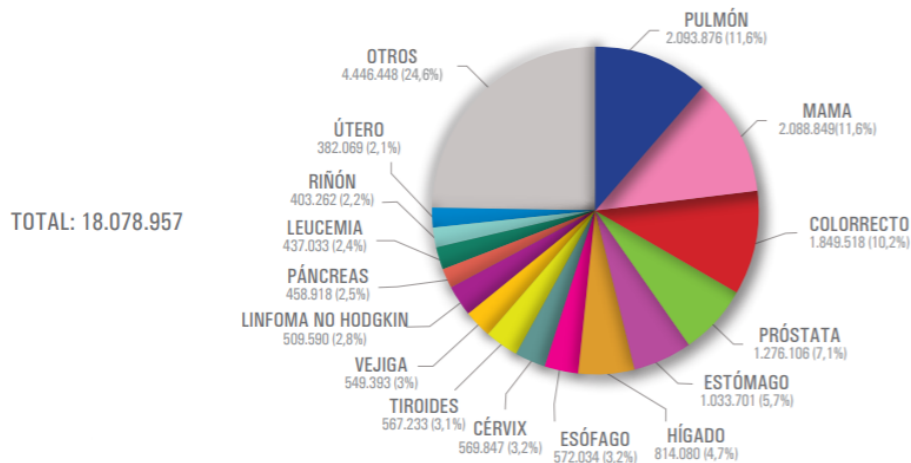


Figura 1.1: Estimación de los tumores más frecuentemente diagnosticados en el mundo. Estimación para el año 2018, ambos sexos. Fuente: [5]

Como se puede observar, pese a la gran variedad de tipos de cáncer, el cáncer de vejiga es uno de los cánceres más incidentes a nivel mundial, exactamente el undécimo carcinoma en frecuencia y se prevé un aumento en los próximos años. Además, es el segundo tumor más frecuente del tracto urinario tras el carcinoma de próstata.

Según la Sociedad Española de Oncología Médica (SEOM), se prevé que los cánceres más frecuentemente diagnosticados para ambos sexos en España en 2020 serán los de colon y recto (44.231 nuevos casos), próstata (35.126), mama (32.953), pulmón (29.638) y vejiga urinaria (22.350). La población masculina en España en el 2020, los tipos de cáncer más incidentes serán el de próstata, colon y recto, pulmón y vejiga urinaria [5].

Por ello, este trabajo pretende ayudar a los expertos en términos de coste-eficacia, ya que para ellos supone una carga manual muy tediosa, además de subjetiva, pues cada patólogo puede diagnosticar de manera diferente una misma muestra. Por lo tanto, se pretende crear un sistema de apoyo para los patólogos que ayude a reducir esa carga de trabajo, por medio de algoritmos de inteligencia artificial basados en patología digital.

Cáncer de vejiga

El tumor de vejiga, o también conocido carcinoma urotelial, es una alteración en las células de la vejiga con una proliferación incontrolada que desarrolla el tumor. Según la Asociación Española Contra el Cáncer (AECC), el cáncer de vejiga es el quinto en incidencia entre los hombres en países desarrollados [6].

Inicio, crecimiento y propagación del cáncer de vejiga

La vejiga está compuesta de varias capas formadas por diferentes clases de células. El inicio más común del tumor suele ser en la primera capa del interior de la vejiga, es decir, el urotelio. A menudo, los cánceres de vejiga se describen basándose en el recorrido de crecimiento que toma el tumor desde la pared de la vejiga, pudiendo ser cáncer de vejiga músculo-invasivo o no invasivo. En la tabla mostrada en la Figura 1.2, se muestra que dependiendo del tipo de crecimiento, se determina el grado y el pronóstico del paciente.

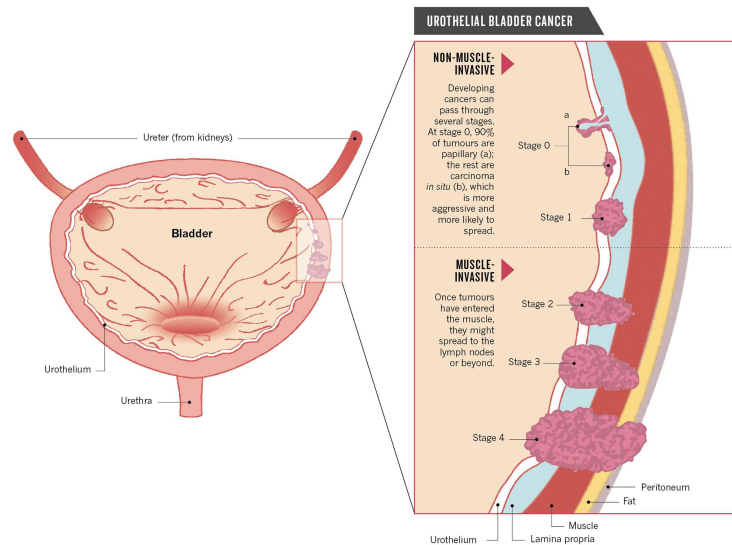


Figura 1.2: Estadios del cáncer dependientes del crecimiento. Obtenida en [7]

- **Crecimiento músculo invasivo (CVMI):** su principal característica es crecer hacia capas más profundas partiendo de la capa interna de la vejiga, llamada revestimiento. Esta propiedad es directamente proporcional del aumento de la facultad de propagación y, por consiguiente, dificulta el éxito al tratamiento, lo cual se traduce en un peor pronóstico del paciente. Cabe destacar que, durante los 2-3 primeros años tras la cistectomía (extracción de la vejiga y parte de la uretra), los pacientes suelen desarrollar metástasis a distancia, ya que cuando presentan síntomas hay una alta probabilidad de padecer metástasis.

Aproximadamente, el 25 % de los diagnósticos, se trata de un carcinoma vesical músculo-invasivo de estadio 2 a 4. Es importante notar que la tasa de mortalidad es elevada, incluso contando con los nuevos avances tecnológicos para el diagnóstico y tratamiento.

- **Crecimiento no invasivo (CVNMI):** Los tumores vesicales no invasivos son aquellos que se encuentran confinados en la mucosa o submucosa (pTa, pT1, o Cis) [8]. Como se observa en la Figura 1.2, estas no crecen hacia el exterior de la vejiga, sino que, crecen formando papilares, o células planas. Este tipo de crecimiento supone el 75 % de los diagnósticos de cáncer vesical.

En este trabajo se va a analizar el CVMI, por ser de peor pronóstico y por sus características de crecimiento y propagación hacia otras vísceras. La gran mayoría de las muertes del cáncer vesical corresponden a este tipo. Además, hoy en día, a pesar del gran desarrollo tecnológico, el CVMI sigue siendo una patología con mortalidad elevada [8].

En la literatura se conocen tres formas de propagación del CVMI primario:

- **Crecimiento local:** Se produce por invasión en profundidad desde las células uroteliales, traspasando la lámina propia e infiltrando la capa muscular. Puede afectar por proximidad a las vísceras como el útero o la vagina en las mujeres y la próstata en los varones.
- **Diseminación linfática:** Cuando el tumor afecta a la capa muscular de la vejiga, consigue afectar fácilmente a los ganglios linfáticos adyacentes y se puede diseminar a otras partes del cuerpo. En base a la experiencia del comportamiento del cáncer de vejiga músculo invasivo, es común que se propague a los ganglios linfáticos distantes, huesos, pulmones o hígado.

- **Diseminación hematológica:** Es común la diseminación a través de los vasos sanguíneos cuando llega a la capa muscular de la vejiga. Suele propagarse hacia el hígado, los pulmones y el hueso.

Incidencia

El cáncer de vejiga está directamente relacionado con el entorno laboral, como pueden ser las exposiciones ocupacionales (agentes químicos, pinturas, cuero y caucho), y también al hábito de fumar, radiación en la pelvis y el factor genético [9]. Es por eso que hay una diferencia considerable entre los países desarrollados y subdesarrollados. Hasta hace unos años, se podría decir que había una gran diferencia en la incidencia entre hombres y mujeres (Figura 1.3 y 1.4). Sin embargo, desde que la población femenina consiguió incorporarse dentro del mundo laboral y se incrementó el hábito de fumar, en la Figura 1.4 se puede observar que, debido a este desarrollo social, se prevé que el cancer de vejiga aumente considerablemente la incidencia del cáncer de vejiga en las mujeres.

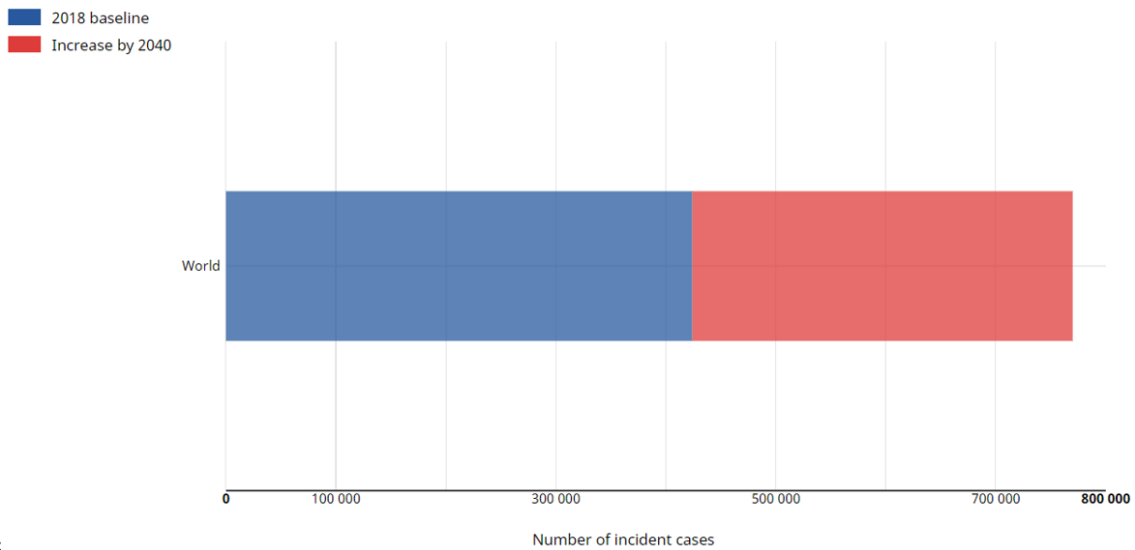


Figura 1.3: Incidencia a nivel mundial del cáncer de vejiga en la población masculina. Fuente: [4]

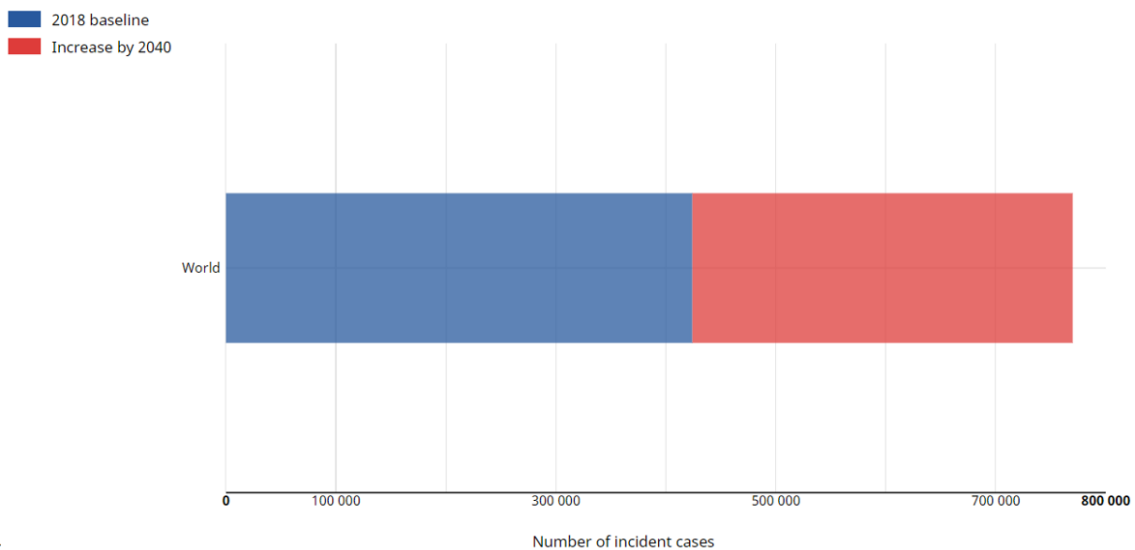


Figura 1.4: Incidencia a nivel mundial del cáncer de vejiga en la población femenina. Fuente: [4]

Como se ha detallado anteriormente, a nivel mundial el cáncer de vejiga supone el undécimo carcinoma respecto a todos los tipos de cáncer y el segundo más incidente del tracto urinario. Además, como en muchos otros tipos de cáncer, hay grandes diferencias entre los países según su nivel de industrialización (Figura 1.5 y 1.6). En Europa, aproximadamente se estima una incidencia de aproximadamente 118.000 casos al año [8].

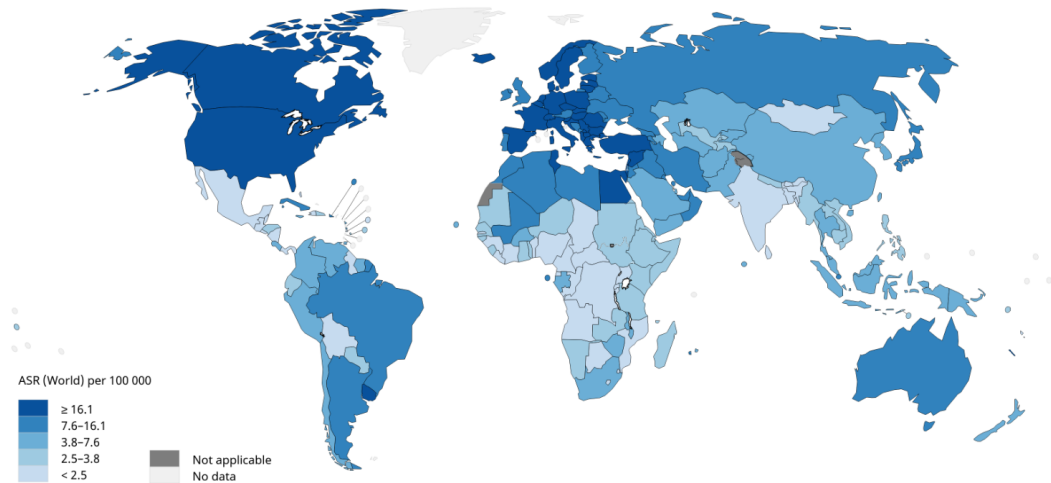


Figura 1.5: Incidencia en el año 2018 a nivel mundial del cáncer de vejiga en hombres. Fuente: [4]

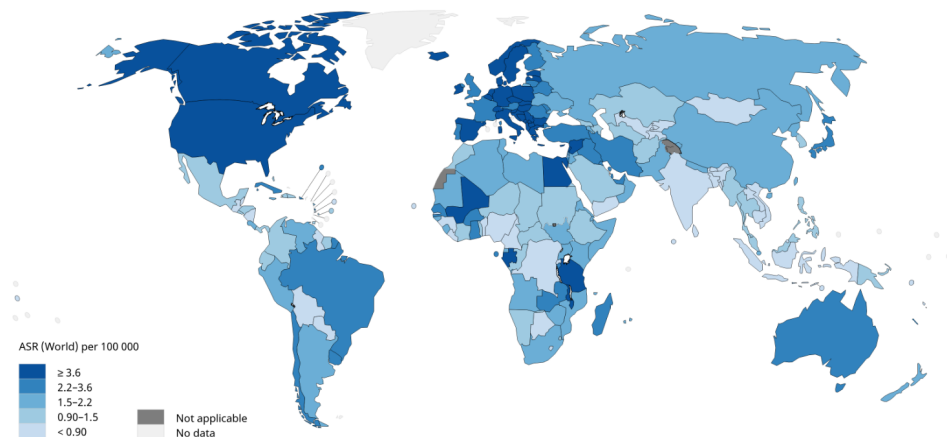


Figura 1.6: Incidencia en el año 2018 a nivel mundial del cáncer de vejiga en mujeres. Fuente: [4]

Según la OMS, España es uno de los países con más incidencia del cáncer de vejiga a nivel mundial, ya que se diagnostican alrededor de unos 19.000 casos cada año. En el año 2018 se registraron el 81 % de los tumores del sexo masculino (14.793 casos) y el 19% de los femeninos (3.475 casos). La mayoría de los casos se diagnostican entre los 65 y los 75 años, aunque se registran casos desde los 40-45 años [10]. En los últimos años, el cáncer de vejiga ha aumentado y se prevé que en el 2040 la incidencia en la población femenina aumentará un 30 % y en la población masculina un 60 % de su incidencia actual (véase las Figuras 1.7 y 1.8).

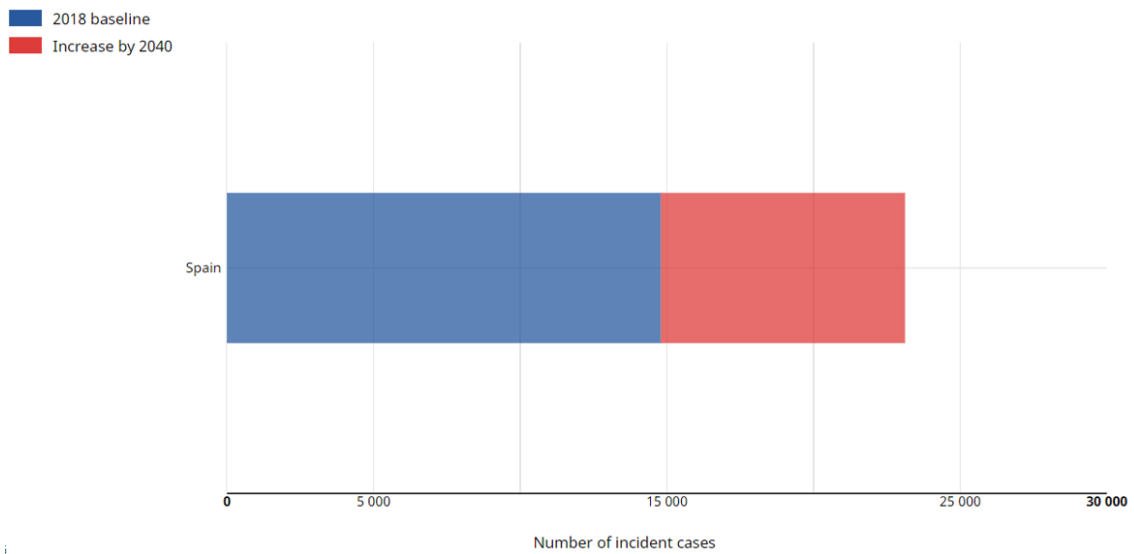


Figura 1.7: Estimación del número de casos de incidencia desde el 2018 a 2040 a nivel nacional, para el cancer de vejiga, población masculina y todas las edades. Fuente: [4]

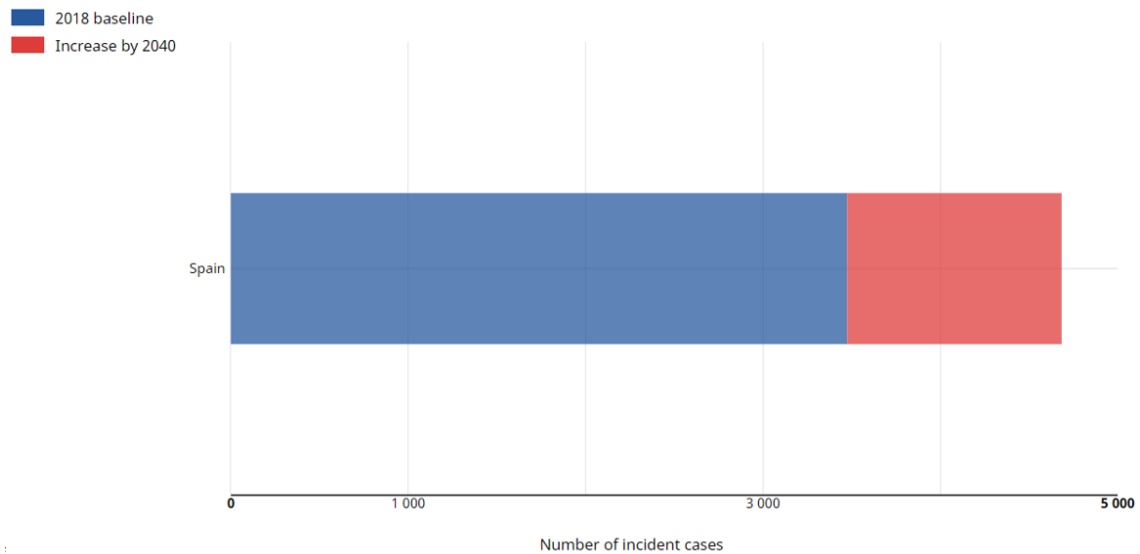


Figura 1.8: Estimación del número de casos de incidencia desde el 2018 a 2040 a nivel nacional, para el cancer de vejiga, población femenina y todas las edades. Fuente: [4]

Mortalidad

Pese a la gran incidencia que supone el cáncer de vejiga, con las tasas de mortalidad ocurre lo contrario gracias al avance de las nuevas tecnologías de diagnóstico precoz, al descenso del hábito de fumar, a la prevención de los riesgos laborales y a la investigación en busca de mejores tratamientos para el cáncer de vejiga. Como se muestra en la Figura 1.9 y 1.10, el número de casos de mortalidad en el 2018 fue alrededor de 150.000 muertes en la población masculina y entorno a 50.000 muertes en la población femenina. Cabe destacar que, se prevé un aumento en el 2040 del 94% en el número de los casos de mortalidad para toda la población a nivel mundial. De acuerdo a lo expuesto en las Figuras 1.11 y 1.12, en España, el cáncer de vejiga es la sexta causa de muerte por cáncer en los hombres, donde en el año

2018 se registraron muertes 4.576 hombres y 1.104 mujeres. Además, en el 2040 se prevé que la tasa de mortalidad del cáncer de vejiga aumentará un 63% para toda la población. Es importante destacar que, aproximadamente 10-15% de los pacientes con CVMI, ya presentan metástasis a distancia en el momento del diagnóstico [4] [8].

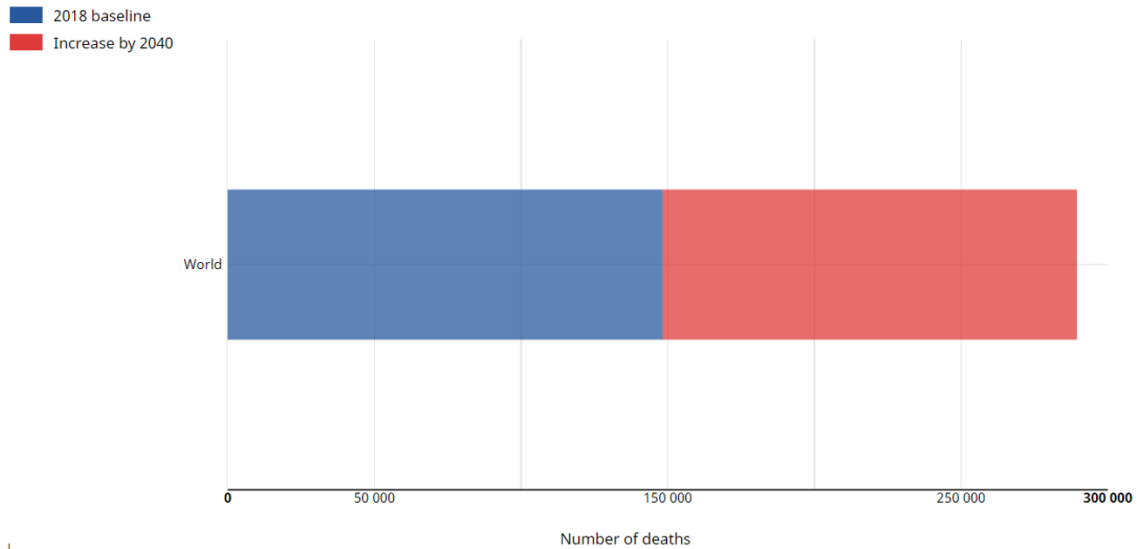


Figura 1.9: Estimación del número de casos de mortalidad desde el 2018 a 2040 a nivel nacional, para el cancer de vejiga, población masculina y todas las edades. Fuente: [4]

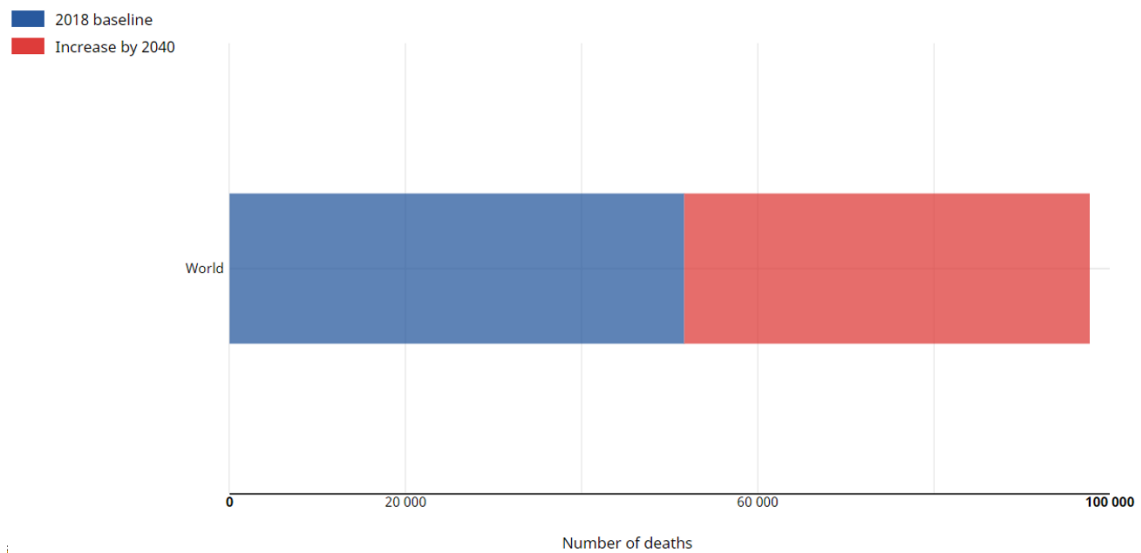


Figura 1.10: Estimación del número de casos de mortalidad desde el 2018 a 2040 a nivel mundial, para el cancer de vejiga, población femenina y todas las edades. Fuente: [4]

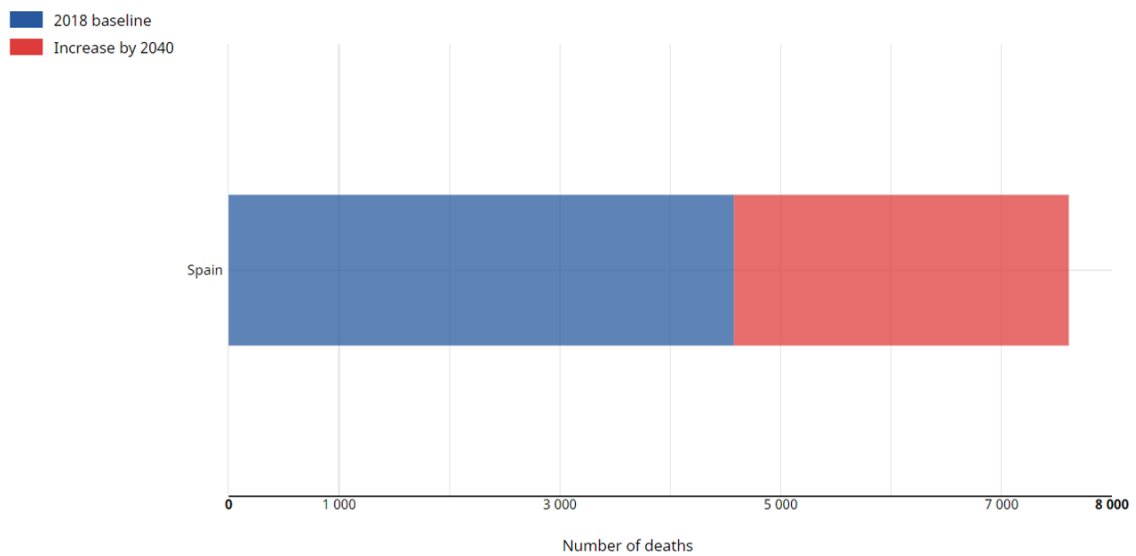


Figura 1.11: Estimación del número de casos de mortalidad desde el 2018 a 2040 a nivel mundial, para el cancer de vejiga, población masculina y todas las edades. Fuente: [4]

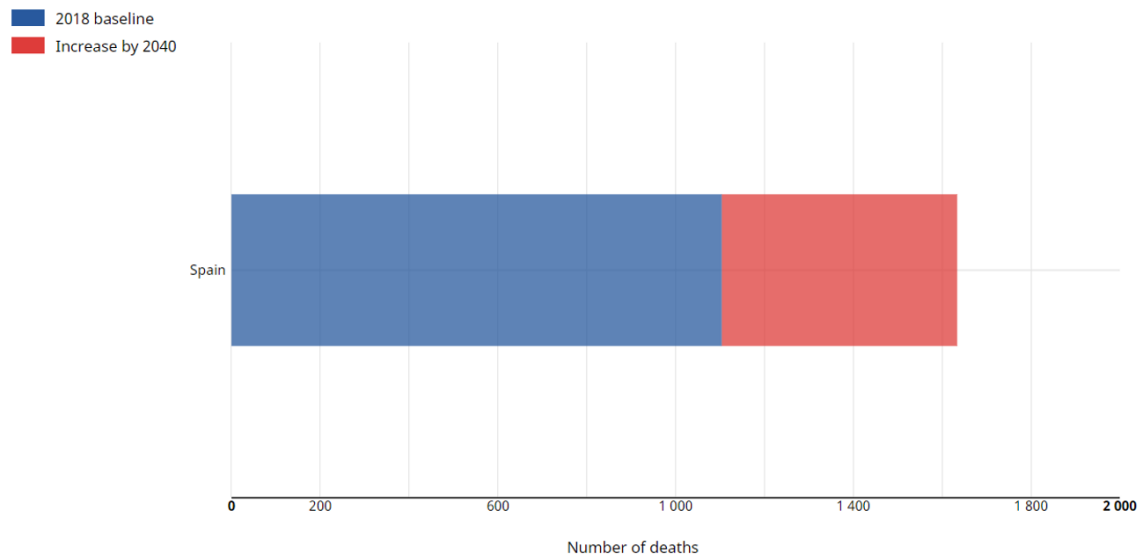


Figura 1.12: Estimación del número de casos de mortalidad desde el 2018 a 2040 a nivel nacional, para el cancer de vejiga, población femenina y todas las edades. Fuente: [4]

Visto las estadísticas tanto a nivel mundial como nacional, es necesaria la investigación sobre esta patología, ya que, en los próximos años, se prevé un incremento del número de muertes debido al aumento y envejecimiento de la población y al consumo del tabaco. En este contexto, se hace evidente la necesidad de la implantación de modelos de aprendizaje atendiendo a diferentes parámetros que nos permitan conseguir un diagnóstico más eficaz y seguro, así como desarrollar terapias individualizadas y más efectivas.

1.1.2 Procedimiento de diagnóstico

Generalmente, esta enfermedad es asintomática hasta que se desarrolla una fase avanzada. Una vez se muestra cualquier signo de sospecha de cáncer de vejiga, siendo el más habitual en su primer episodio la hematuria macroscópica, hay que realizar pruebas para el diagnóstico con el objetivo de determinar la extensión, es decir, el estadio del cáncer, y la malignidad del tumor.

Dentro de todos los test que existen de diagnóstico del cáncer de vejiga, en este trabajo se va a explicar el proceso de diagnóstico más común y que se aplica de manera rutinaria y fiable en la actualidad.

1. **Examen físico:** Se trata de un examen manual del recto, también conocido como examen digital del recto (DRE), si es varón, o en caso de ser mujer, se realiza un examen pélvico. Si el médico encuentra algo anormal se derivará a un urólogo para que se hagan más pruebas y tratamientos [8].
2. **Citología de orina:** Esta prueba determina si hay células cancerígenas o células precancerosas en la orina. Cabe destacar que este test necesita ser valorado con otras pruebas, debido a su baja fiabilidad por tener muchos falsos negativos. Es importante saber que, gracias a los grandes avances de los marcadores tumorales, se pueden identificar sustancias específicas que producen las células cancerosas de la vejiga [11].
3. **Ecografía vesico-prostática y renal:** Ambas técnicas son muy útiles para localizar el tumor sin irradiar al paciente, y tener una idea de del tipo de crecimiento (CVMI o CVNMI), ya que se puede observar si se ha propagado fuera de la vejiga. Además, se analizan los riñones [12].
4. **Urograma por MRI:** Es de gran utilidad para indicar si hay diseminación local y ganglionar y analizar la parte superior del sistema urinario, puesto que en algunas ocasiones no se puede localizar el tumor [12].
5. **Cistoscopia:** Si se sospecha de cáncer de vejiga, la cistoscopia es la prueba que recomiendan. Se trata de un tubo delgado, largo y flexible que en su extremo tiene una fuente de luz y una cámara de video [12]. Existen los siguientes tipos:

- **Resección transuretral de un tumor de vejiga (TURBT):** Esta técnica es la más usada para conocer las características del tumor. Consiste en extraer muestras del tejido con un aspecto anormal de la vejiga, es decir, del tumor y también del músculo que lo rodea para determinar si se trata de CVMI o CVNMI [12]. Este procedimiento también se conoce como resección transuretral (TUR).

Uno de los métodos más usados para la identificación del cáncer de vejiga es la biopsia con aguja que, sin hacer una cirugía, con una aguja larga, hueca y delgada guiada por una ecografía se recoge una muestra pequeña.

- **La cistoscopia de fluorescencia:** Esta técnica es menos común que la cistoscopia rutinaria, pero pueden complementarse para ayudar al experto a encontrar con mayor precisión la localización de todas las células tumorales. Para ello, se coloca un medicamento sensible a la luz fluorescente que es atraído por las células cancerígenas. La respuesta es la emisión de fluorescencia de las células tumorales [12].

Biopsia

Una vez se ha evidenciado la presencia de cáncer vesicular por las pruebas que se han explicado anteriormente, el diagnóstico definitivo se determina con el análisis histológico del tumor. Por lo tanto, mediante una TURBT se extraen las áreas sospechosas y se envían al departamento de anatomía patológica para que los expertos consigan el resultado de la biopsia a nivel celular. El diagnóstico se determina con dos parámetros muy importantes a resolver:

- **Nivel de invasión del cáncer:** El estadio del cáncer vesical viene establecido por la invasión en las capas profundas de la vejiga. Se establece el estadio TNM en cada caso (Figura 1.13) [8].
- **Grado del cáncer:** Esta variable se decide en base a la apariencia y estructura que toman las células cancerosas. Más adelante, se explican los patrones de crecimiento en el frente tumoral que, como se ha estudiado en la parte clínica, son un factor pronóstico muy importante que determina el grado del tumor. Por lo general, en la literatura se pueden diferenciar en dos subgrupos:
 - *Bien diferenciados:* Son aquellos en los que las células cancerígenas son muy similares al tejido normal de la vejiga. También se suele llamar “cáncer de grado bajo de diferenciación” y se relaciona con un pronóstico favorable, ya que su capacidad de crecer y diseminarse es muy baja. [12].
 - *Pobrementemente diferenciados:* Corresponde a los cánceres de alto grado, y son aquellos que menos se parecen al tejido normal. Estos presentan una agresividad mucho mayor, con mucha facilidad de crecer en la pared de ella y propagarse fuera de la vejiga. Asimismo, el pronóstico no suele ser muy bueno, y este tipo de cáncer suele ser muy complicado de tratar [12].

T – Tumor primario	
Tx	No se puede evaluar el tumor primario
T0	Ausencia de datos de tumor primario
Ta	Carcinoma papilar no invasor
Tis	Carcinoma in situ: “tumor plano”
T1	El tumor invade el tejido conjuntivo subepitelial
T2	El tumor invade el músculo
	T2a El tumor invade el músculo superficial (mitad interna)
	T2b El tumor invade el músculo profundo (mitad externa)
T3	El tumor invade el tejido perivesical
	T3a Microscópicamente
	T3b Macroscópicamente (masa extravesical)
T4	El tumor invade cualquiera de las estructuras siguientes: próstata, útero, vagina, pared de la pelvis o pared abdominal
	T4a El tumor invade la próstata, el útero o la vagina
	T4b El tumor invade la pared de la pelvis o la pared abdominal
N – Ganglios linfáticos	
Nx	No se pueden evaluar los ganglios linfáticos regionales
N0	Ausencia de metástasis ganglionares regionales
N1	Metástasis en un solo ganglio linfático en la pelvis verdadera (hipogástricos, obturadores, ilíacos externos o presacros)
N2	Metástasis en varios ganglios linfáticos en la pelvis verdadera (hipogástricos, obturadores, ilíacos externos o presacros)
N3	Metástasis en uno o varios ganglios linfáticos ilíacos comunes
M – Metástasis a distancia	
M0	Ausencia de metástasis a distancia
M1	Metástasis a distancia
	M1a Ganglios linfáticos no regionales
	M1b Otras metástasis a distancia

Figura 1.13: Clasificación del estadio TNM de la Unión Internacional Contra el Cáncer (UICC) para el cáncer de vejiga. Fuente: [13]

La determinación de los parámetros T, N y M para su diagnóstico, se hace en base a las imágenes histológicas correspondientes a la zonas elegidas de la biopsia. El diagnóstico seguro va a depender de la estructura histológica que ha desarrollado el tumor, y por ello, se hace necesario el análisis a nivel celular.

Imagen histológica

Para poder obtener el resultado de la biopsia, es necesario que la muestra pase por un procedimiento con la finalidad de conseguir los portaobjetos junto al tejido. Seguidamente, una vez se consiguen las imágenes histológicas, los expertos deben analizar todas ellas para determinar el diagnóstico. Cabe destacar que, establecer un diagnóstico de cada biopsia se debe estudiar con detalle todas las características del tumor bajo el microscopio. El tiempo para efectuar esta tarea depende del tamaño del tumor, pudiendo relantizar el diagnóstico. El presente trabajo, pretende abordar el problema del tiempo y eficiencia del diagnóstico con un método de gradación del tumor automático que reduce considerablemente el tiempo.

El procedimiento general que se suele usar para preparar el portaobjeto es viene detallado en la Figura 1.14. En primer lugar, el patólogo toma las muestras que considera importantes del tejido procedente de la pieza de cistectomía, que después son fijadas en formalina tamponada al 10%, a fin de preservar los tejidos y evitar los procesos de muerte celular. Durante esta fase, el tejido es deshidratado mediante un procesador automático que extrae el agua de los mismos, y luego se elimina el alcohol de los tejidos provocando que se vuelvan transparentes (Figura 1.14, Fase 2). Seguidamente, se impregna la muestra con parafina líquida que, cuando se enfría, se solidifica y crea un bloque consistente (fase de la inclusión). Es necesario que el bloque sea resistente ya que, en la siguiente fase se realizan cortes finos de 3 micras (μm) en un equipo de precisión llamado micrótopo y se colocan en una placa llamada portaobjetos (Figura 1.14, Fases 4 y 5). A continuación, se realiza una tinción donde se añaden colorantes a los tejidos, como es en nuestro caso uno de los métodos más comunes la tinción Hematoxilina-Eosina (HE), para diferenciar sus estructuras celulares. Finalmente, una vez realizada toda la preparación, los portaobjetos quedan listos para que lo patólogos los analicen por medio de un microscopio [8].

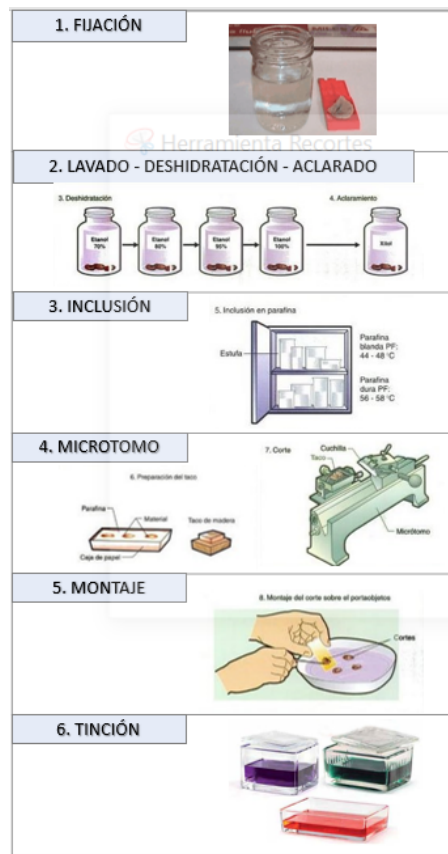


Figura 1.14: Esquema de preparación de las imágenes histológicas para su análisis bajo el microscopio, a partir de una biopsia. Elaboración propia.

Para el presente proyecto, se utilizaron las mismas imágenes histológicas que el estudio clínico [8]. Para la adquisición de las imágenes, se visualizaron primero en $\times 4$ - $\times 10$ aumentos, y tras ello, a $\times 40$ aumentos para seleccionar entre todas las laminillas de cada caso la que mostraba, una mayor densidad del patrón de agresividad más alto, es decir, la más representativa. Una vez escogidas todas ellas, se aplicó la técnica de inmunohistoquímica CK AE1/3 que, cuando se produce la unión antígeno-anticuerpo se hace visible las células cancerígenas produciendo coloración (cromógeno). Como se aprecia en la Figura 1.15, esta tinción facilita la localización del frente tumoral.

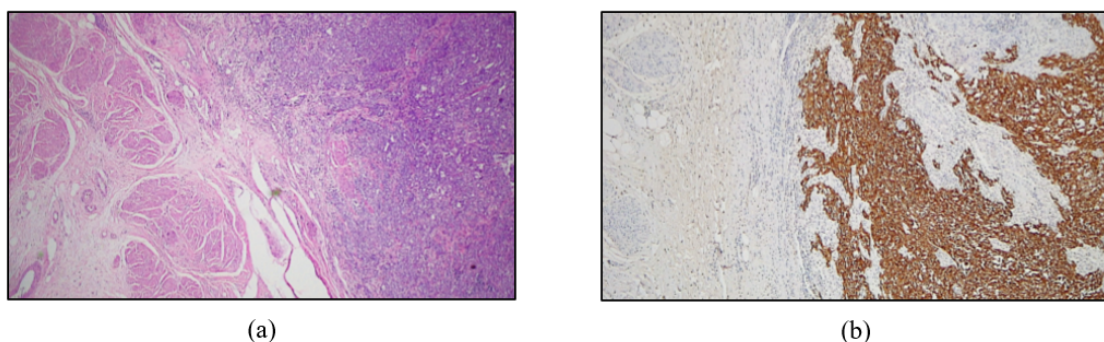


Figura 1.15: Zona de frente tumoral invasor a pequeños aumentos (4 \times). (a) Tinción HE. (b) Tinción inmunohistoquímica CK AE1/3. Fuente: [8].

Patrones morfológicos

En el año 2001, Jimenez et al. [14] describieron los patrones infiltrativos para el cáncer de vejiga. Dichos patrones morfológicos, a pesar de no estar incorporados en la práctica como factor predictor, determinan el grado de malignidad que tiene el tumor. En el cáncer de vejiga invasor no suele haber casos de NUPBPM o carcinoma de bajo grado. Todos los casos son carcinomas uroteliales de alto grado (grado II o III en la clasificación de la OMS de 1973) [15]. De acuerdo con lo expuesto en la Figura 1.17, se identifican los siguientes patrones [8]:

- **Patrón nodular:** Está compuesto mayoritariamente por nidos de células tumorales bien delineados con forma circular. Se asocia a un grado II del tumor.
- **Patrón trabecular:** El conjunto de células tumorales está dispuesto en bandas interconectadas entre sí. La forma de las bandas son alargadas, conocidas como trabéculas. Este patrón también se asocia al grado II del tumor.
- **Patrón infiltrativo:** Compuesto por estrechos cordones de células tumorales y un conjunto de menor a 4 células “*buds*” considerado grado III del tumor, también conocido como “tumor *budding*”.

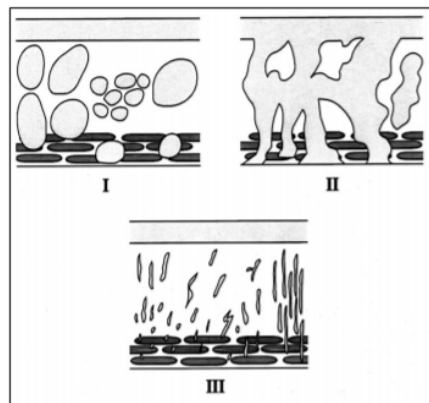


Figura 1.16: Representación de los tres patrones de crecimiento propuestos por Jimenez et al. [14]. I) Patrón nodular; II) Patrón trabecular; III) Patrón infiltrativo. Fuente: [8]

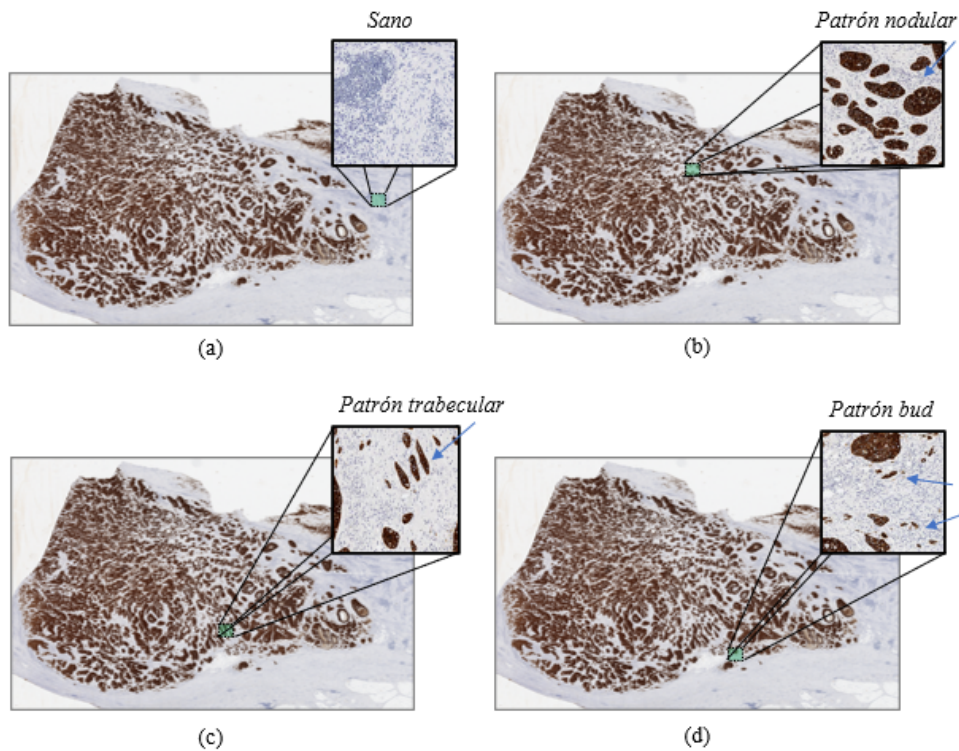


Figura 1.17: Ejemplos de patrones de crecimiento de cáncer de vejiga con imágenes histológicas reales. (a) Sano, sin presencia de cáncer. (b) Patrón nodular. (c) Patrón trabecular. (d) Patrón infiltrativo. Elaboración propia.

Tumor budding

Tumor *budding* (TB) es un concepto anatomopatológico que consiste en la presencia de células tumorales aisladas o en pequeños grupos de hasta 4 células tumorales (“*buds*”) (véase la Figura 1.18), que le confiere al tumor una alta agresividad y de peor pronóstico .

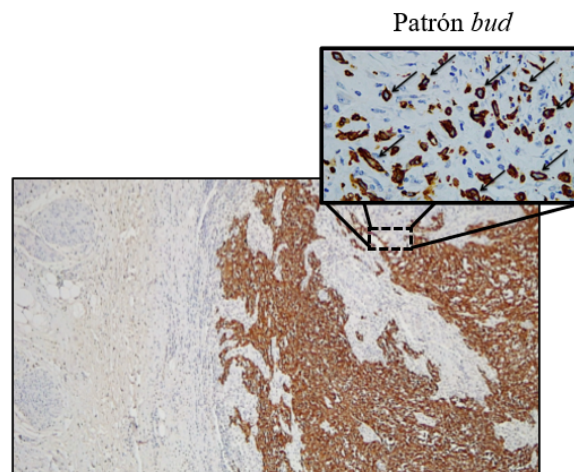


Figura 1.18: Zona de frente tumoral invasor a pequeños aumentos (4×) y zona destacada por la presencia del tumor *budding* a 40×. Fuente: [8].

Por primera vez, Imai [16] en el año 1954 definió que la presencia de TB se relacionaba con una alta agresividad tumoral. En el 1986, Jeremy R. Jass [14] identificó que, para el cáncer de colon, el pronós-

tico estaba determinado por el frente tumoral, es decir, que se trataba de un indicador muy importante para el pronóstico. Más adelante, otros expertos indicaron que, la presencia del patrón de crecimiento tipo infiltrativo, como es el TB, correspondía a una alta capacidad invasora y, por consiguiente, futura metástasis [8].

Cabe destacar que, hasta hace pocos años el análisis del posible factor pronóstico de la presencia del patrón TB en el frente tumoral para el cáncer de vejiga no había sido investigado. Actualmente, existen pocos estudios que demuestran tener carácter pronóstico para el cáncer de vejiga. Este fenómeno está íntimamente ligado con el inicio de la metástasis tumoral, ya que le confiere al tumor una alta capacidad de separarse de la masa tumoral principal y migrar a los vasos linfáticos y sanguíneos a otras partes del cuerpo.

Como se ha comentado anteriormente, se ha demostrado en varios estudios [17] [18] que, la presencia de los “*buds*” en pacientes con estadios tempranos, es un indicio de micrometástasis no detectables en la clínica. Si la situación de un paciente empeora hacia estadios más avanzados, este fenómeno tiene como resultado una reducción de la tasa de supervivencia a cinco años y un gran aumento en la recaída de la enfermedad una vez realizado el tratamiento [8]. De modo que, la presencia o no de este patrón de crecimiento en el frente tumoral aporta información de la gravedad de la patología y sobre todo del manejo terapéutico frente a estas características.

Hoy en día, sigue sin estar establecida en la práctica una técnica rutinaria, tanto clínica como computacional, relacionada con la detección del TB en el CVMI. Para este proyecto se va a desarrollar un modelo predictivo de ayuda al diagnóstico atendiendo a los patrones de crecimiento, tanto en el frente invasor como en el interior, que permita extraerlos, cuantificarlos y categorizarlos en los grados II y III.

1.2 Objetivos

El objetivo final de este proyecto es el diseño y desarrollo de un sistema de inteligencia artificial capaz gradar el cáncer de vejiga. En particular, se pretende llevar a cabo un modelo predictivo basado en algoritmos de *deep learning* que sirvan de ayuda al diagnóstico a fin de facilitar la tediosa tarea de los patólogos en términos de tiempo y eficacia. Además, con este modelo también se pretende reducir el nivel de subjetividad que existe entre diferentes patólogos cuando tienen que analizar una determinada muestra.

Se va a desarrollar un algoritmo con el que se consiga diferenciar los grados de agresividad atendiendo a la arquitectura morfológica celular tanto en el frente tumoral como en el interior del tumor, haciendo gran hincapié en los patrones explicados en el Capítulo 1.1.2.

Para lograr la consecución del objetivo principal de este trabajo final de máster se proponen una serie de objetivos más específicos que se listan a continuación:

1. Conocer el estado de arte sobre las técnicas de clasificación en imágenes histopatológicas para cáncer de vejiga y determinar el valor pronóstico que supone la presencia de tumor *budding* en la arquitectura tumoral.
2. Construir una base de datos de imágenes usando la tecnología *Whole Slide Image* (WSI) cuyas dimensiones sean las correctas para poder trabajar sobre ellas.

3. Estudiar en profundidad y aprender los patrones más comunes de los diferentes grados del cáncer de vejiga para determinar un *groundtruth* correspondiente a cada bloque de la imagen para entrenar y validar el modelo.
4. Realizar una exhaustiva partición de datos empleando técnicas de validación cruzada para reportar resultados fiables y objetivos, teniendo en cuenta el número de imágenes del conjunto y de cada etiqueta.
5. Aplicar técnicas supervisadas y no supervisadas de análisis de imagen para llevar a cabo la clasificación del cáncer de vejiga a partir de algoritmos de *deep learning*.
6. Obtener métricas de evaluación de los modelos de clasificación y una breve comparación entre ellos.
7. Identificar y especificar las mejoras que se puedan incluir en investigaciones futuras y determinar las limitaciones encontradas durante la realización del proyecto para ahondar en ellas en posteriores trabajos.

1.3 Guía de la memoria

A continuación, en el Capítulo 2 se detalla, con mayor profundidad, el material utilizado para este proyecto y en el Capítulo 3 la metodología usada para diseñar y desarrollar la clasificación de cada imagen y, así, identificar la presencia de cáncer o no o la ausencia de cáncer de vejiga, así como sus respectivos grados de malignidad II y III.

Posteriormente, en el Capítulo 4 se llevan a cabo las discusiones sobre los resultados obtenidos para analizar el rendimiento del modelo. Además, se describen y examinan tanto el conjunto de validación de la fase entrenamiento y validación, como la fase de predicción test. También, se muestran los resultados cualitativos y cuantitativos obtenidos durante la fase de predicción.

Para terminar, en el Capítulo 5 se exponen las conclusiones más importantes sobre el trabajo realizado, y se detallan las líneas futuras que podrían abordarse tras finalizar este trabajo final de máster.

Capítulo 2

Material

Índice de contenidos

2.1 Base de datos.	22
2.2 Software y hardware	23
2.2.1 Software.	23
2.2.2 Hardware	23

2.1 Base de datos

Gracias al Departamento de Anatomía Patológica del Hospital Politècnic Universitari la Fe, se ha podido adquirir la base de datos de las imágenes utilizadas para realizar el presente trabajo fin de máster se han obtenido mediante la digitalización de las muestras de tejido de vejiga extraídas en las biopsias e incluidas en los portaobjetos, tal y como se ha explicado en el Apartado 1.1.2. La digitalización del tejido histológico se ha llevado a cabo gracias al escáner Aperio CS2 (Leica Biosystems, San Diego, CA, USA) que crea portaobjetos digitales de alta calidad y con una capacidad de escaneo del 20x y 40x con una alta fiabilidad [19]. Dicho escáner está guiado por un software llamado Aperio ImageScope que permite el manejo del mismo y la adquisición de las imágenes en formato “.svs” (Figura 2.1).

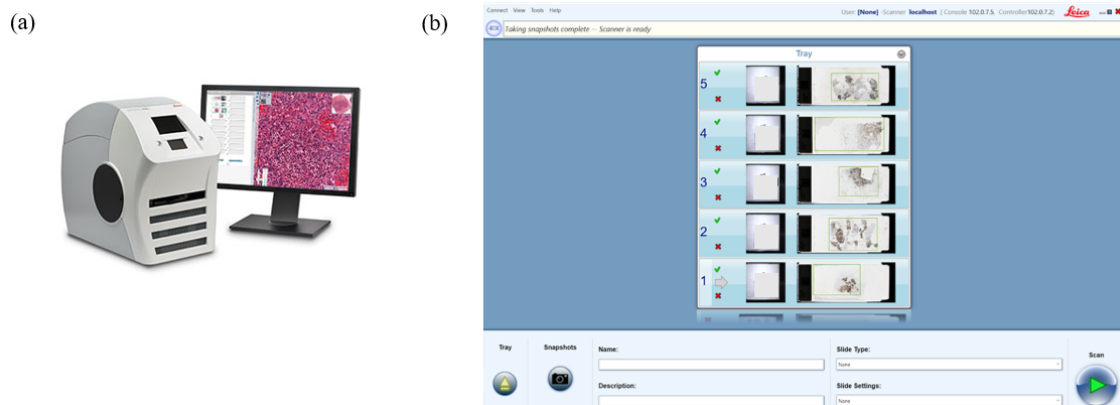


Figura 2.1: A) Escáner Leyca Biosystems Aperio CS2 [19]. B) Software *Aperio ImageScope*. Elaboración propia. Dispositivo y programa empleado para la adquisición de imágenes.

La finalidad de la privacidad de la base de datos es debido a la confidencialidad de los pacientes por la ley de protección de datos. Estos firman un consentimiento informado para garantizar la seguridad de los derechos de intimidad, la protección de los derechos y la seguridad.

La base de datos está compuesta por 136 cortes histológicos (biopsias), todos ellos con presencia de cáncer de vejiga. Una vez digitalizados las muestras de tejido tras la biopsia, se obtienen las imágenes histológicas escaneándolas con una resolución de 40x. Cada una de ellas es del orden de [30.000, 20.000, 3] píxeles por imagen RGB, lo cual conlleva unas dimensiones de memoria de almacenamiento de aproximadamente 1,5GB. Este gran tamaño de las imágenes presenta ciertas limitaciones a la hora de cargarlas en el software y de trabajar con ellas. Por esa razón, es necesario llevar a cabo un preprocesado de la base de datos antes de encarar el desarrollo de los modelos predictivos. A continuación, se adjunta un esquema donde se explica el proceso utilizado para la obtención de la base de datos final, que se explica más adelante en el Apartado 3.

2.2 Software y hardware

2.2.1 Software

Para el desarrollo de esta tesis final de máster, se ha hecho uso de dos *softwares* distintos. Por un lado, se ha utilizado el lenguaje MATLAB (C) para la extracción de las regiones de interés, la creación del *groundtruth* y la partición de datos por el método *kfold cross validation*, v. R2019b, de *The MathWorks, Inc. (Natick, Massachusetts, Estados Unidos)*.

Por otro lado, se ha usado el *software* Python 3.7, para la creación de los bloques sobre las imágenes completas y el desarrollo de los modelos de aprendizaje automático, mediante *IDE JetBrains Pycharm* 2019.1.3x64. Cabe añadir que, para el desarrollo del modelo, se ha hecho uso de los distintos *framework* (*keras*) con *TensorFlow* como *backend*.

2.2.2 Hardware

Para evaluar el coste computacional que conlleva la ejecución de los algoritmos, hay que tener en cuenta las características internas del *hardware* para así poder realizar futuras comparaciones entre otros equipos con características diferentes.

El desarrollo y la ejecución de este trabajo final de máster se ha realizado mediante un equipo DELL XPS 9370 con un sistema operativo Windows 10 Pro de 64 bits, un procesador Intel® Core™ i7-8550U (1,8 GHz, 4 núcleos), un controlador gráfico Intel HD Graphics 620, 16GB DDR3-SDRAM 1866 MHz de memoria instalada (RAM) y un disco duro de 512 GB SSDM.2.

Para la etapa de entrenamiento de los modelos, ha resultado necesario el uso de un servidor de computación, situado en el departamento de *CVB Lab*, con capacidad suficiente para llevar a cabo el proyecto. Respecto a las características del servicio, éste está compuesto por un procesador Intel i7 @4.20GHz, 32GB de RAM y tarjetas gráficas NVIDIA Titan V de 12 GB de RAM.

Capítulo 3

Metodología

Índice de contenidos

3.1 Preparación de la base de datos de imágenes	27
3.1.1 Anotación y extracción de la región de interés	27
3.2 Groundtruth para la clasificación	29
3.3 Partición de los datos	33
3.4 Clasificación supervisada	34
3.4.1 Fundamentos de las redes neuronales artificiales: Perceptrón multicapa (del inglés - MultiLayer Perceptron (MLP))	34
3.4.2 Arquitectura red neuronal convolucional: <i>Deep Learning</i>	39
3.5 Clasificación no supervisada	43
3.5.1 Clasificación no supervisada basada en autoencoders	43

Para la conseguir el principal objetivo del presente proyecto, tal y como se aprecia en el diagrama de flujo de la Figura 3.1, el trabajo parte de una base de datos de imágenes histopatológicas que posteriormente se le aplica un pre-procesado para la anotación y extracción de la región de interés, la creación de los bloques y el descarte de aquellos que no contienen información relevante. Una vez se tienen los bloques histopatológicos, se crea el *groundtruth* para cada uno de los bloques. Más adelante, se realiza una partición de datos para el entrenamiento de los modelos. En este trabajo final de máster, se clasifican los bloques desde una clasificación supervisada y otra no supervisada. Finalmente, se exponen los resultados y se comparan entre ellos.

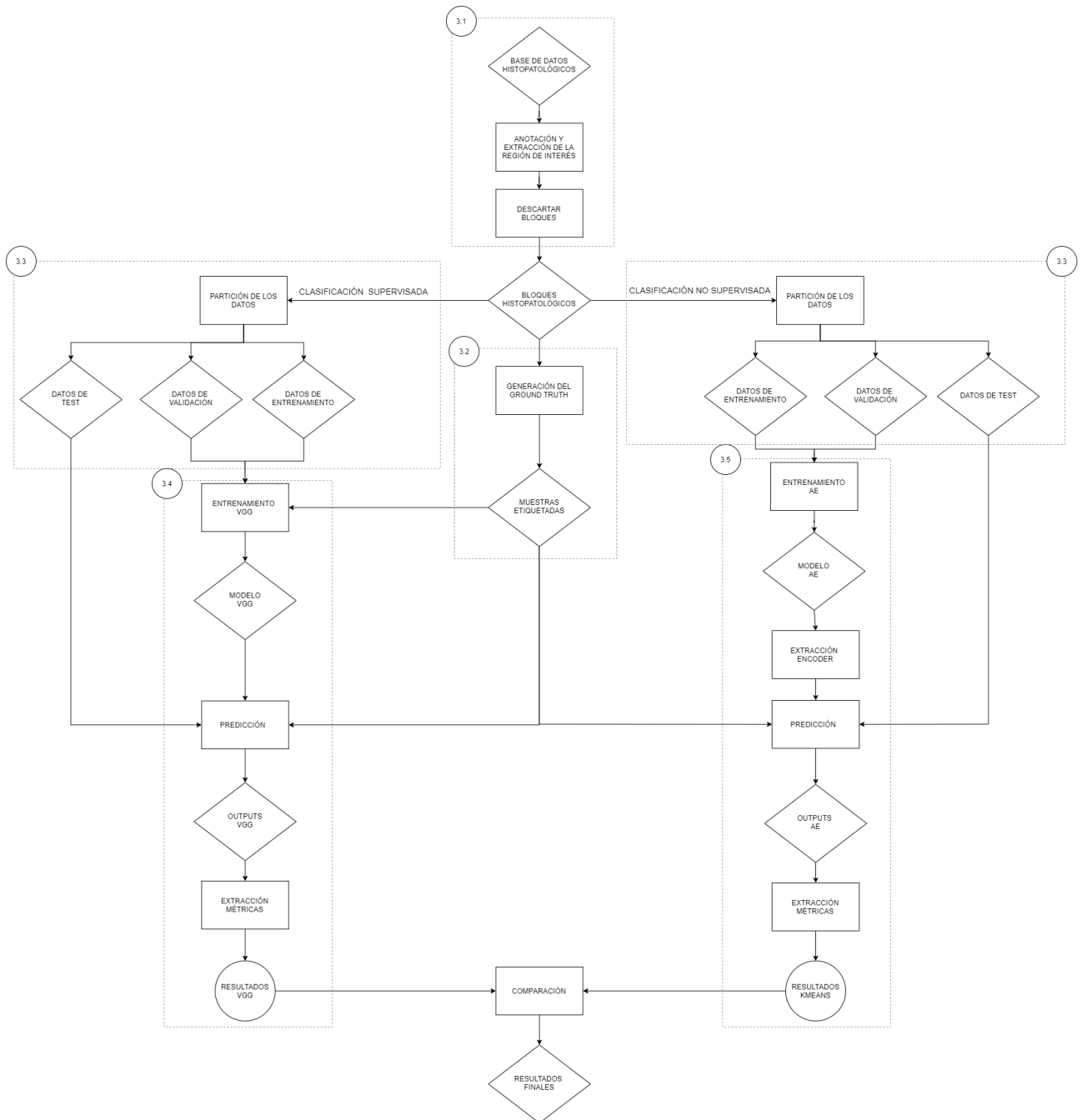


Figura 3.1: Diagrama de flujo correspondiente a la metodología de este proyecto.

3.1 Preparación de la base de datos de imágenes

La preparación de la base de datos de imágenes conlleva realizar un procedimiento muy importante ya que analizar una imagen completa, que ha sido escaneada a $40\times$, supone un alto coste computacional. Es por esto que, para analizar la imagen, se procede a realizar la anotación y la extracción únicamente de la región de interés (ROI), tal y como se observa en la Figura 3.2.

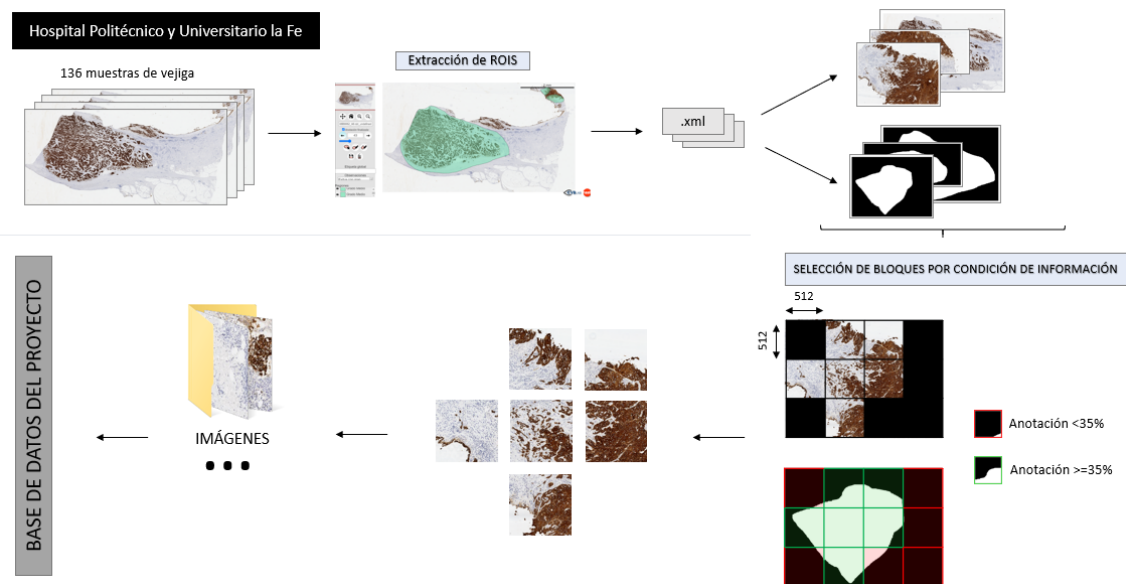


Figura 3.2: Preprocesado de la base de datos. Elaboración propia.

3.1.1 Anotación y extracción de la región de interés

Para extracción de la ROI, tal y como expone la Figura 3.3 se ha utilizado la aplicación “*MicroDraw*” para anotar las zonas de interés de las 136 muestras y extraer las coordenadas dentro de la imagen. Una vez se tienen las coordenadas, se extraen las imágenes y las máscaras correspondientes a la ROI anotada. La herramienta “*MicroDraw*” está creada sobre una librería de código abierto llamada *OpenSeadragon* que consigue, sin perder resolución, una representación de manera interactiva multi-resolución [20]. Un requisito importante para poder cargar las imágenes en la plataforma *MicroDraw*, es que las imágenes estén en formato *.dzi, ya que este formato mantiene en todo momento la calidad original de la imagen. Para ello, primero fue necesario realizar un cambio del formato *.svs a *.tiff mediante el programa gratuito “*ReadConverter*”. Posteriormente, se hizo uso de una librería de procesamiento de imágenes llamada “*vips*” para la conversión del formato *.tiff a *.dzi.

Para el tratamiento de estas imágenes se ha usado la librería “*OpenSlide*” que soluciona el problema de leer imágenes de gran resolución espacial, la lectura de una pequeña cantidad de datos de imagen con la resolución más cercana al nivel de zoom deseado y la lectura de imágenes en distintos formatos como en nuestro caso *.svs [21] [22].

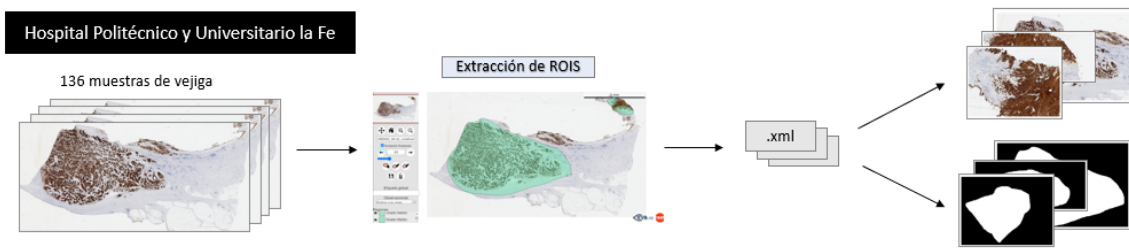


Figura 3.3: Esquema correspondiente a la anotación y extracción de las regiones de interés. Elaboración propia.

Para el presente trabajo se han obtenido unas muestras digitalizadas a un zoom de $40\times$, es decir, la resolución más alta posible, ya que es importante en la fase del preprocesado captar los detalles más pequeños de las imágenes como el TB. Es importante recalcar que, la resolución espacial se describe como la distancia mínima a la que es posible observar dos píxeles que están separado. Cuanto mayor es la resolución de una imagen, mayor es la cantidad de píxeles que contiene dicha imagen. Por lo tanto, trabajar con la resolución más alta ($40\times$), la región de interés contenida en la *boundingbox*¹ todavía sigue siendo muy pesada como para trabajar con ella. Entonces, por esa razón es necesario llevar a cabo un parcheado que permita manipular las imágenes y entrenar los modelos. En este TFM se ha considerado óptimo utilizar dimensiones de $512\times 512\times 3$, ya que este tamaño de imagen es fácilmente asumible por el *software* utilizado y es suficientemente grande como para observar las estructuras de interés relacionadas con el cáncer de vejiga (Figura 3.3). Para conseguir las anotaciones de las zonas de interés, primero se selecciona un *boundingBox* que contiene los píxeles de interés. Después, se elige un tamaño concreto de imagen óptimo del *boundingBox*.

En el mundo de la programación esto se traduce a la implementación de un algoritmo de ventana deslizante a resolución $40\times$ que permite conseguir, con las dimensiones deseadas, parches correspondientes al área de interés.

Cabe mencionar que se han descartado todas aquellas imágenes que no contenían el 35% de la información que se desea analizar, tal y como se ve en la Figura 3.4, dando lugar a una base de datos final que está compuesta por 3.070 bloques útiles para el proyecto.

¹*BoundingBox* es el rectángulo más pequeño capaz de albergar la información de interés

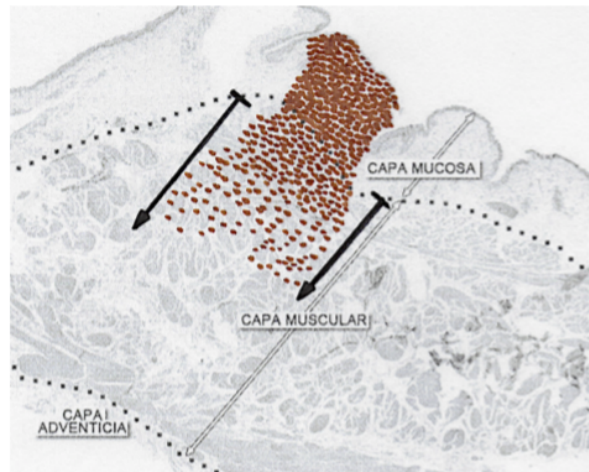


Figura 3.5: Esquema que reproduce la zona de estudio del TB. Fuente: [8]

Para llevar a cabo el *groundtruth* de la clasificación supervisada, la autora del presente proyecto se hizo cargo de la anotación de la región de interés y del etiquetado de cada imagen bajo supervisión de los patólogos del Hospital Universitari i Politècnic la Fe. Conseguir llevar a cabo esta tarea con éxito requiere mucho tiempo de aprendizaje de las propiedades de cada uno de los patrones. Las diferencias entre ellos pueden llegar a ser mínimas y dependientes del contexto histológico general. Además, el etiquetado se complica cuando dentro del parche se comprenden más de un patrón de crecimiento y normalmente el patólogo usa la experiencia de otros casos para su clasificación. Por tanto, para el diagnóstico se tiene en cuenta, primero, la identificación de los patrones que hay dentro de un parche, y, después, la experiencia del patólogo.

Como se aprecia en la Figura 1.17 del Apartado 1.1.2, se trata de una muestra ideal. Sin embargo, esto no ocurre siempre, ya que, en la base de datos empleada para este proyecto, cada parte del frente tumoral comprende diferentes patrones de crecimiento (véase la Figura 3.6).

La determinación del grado de malignidad es difícil de asignar con total seguridad, ya que, por ejemplo, el grado de malignidad II no solo depende del patrón trabecular. Si en un parche existen patrones de grado III y de grado II de manera equitativa, este dependería del contexto y de factores externos para su evolución. Además, los patrones explicados en la teoría del Apartado 1.1.2 no se expresan con tanta claridad en un caso real.

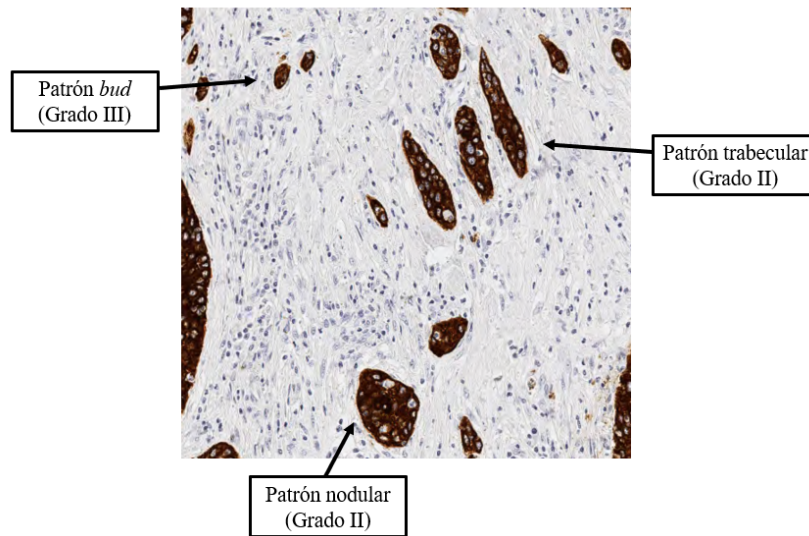


Figura 3.6: Frontera tumoral con diferentes patrones.

Es necesario recalcar que la alumna además de ser ingeniera, es biomédica y fue capaz de realizar un estudio y un aprendizaje exhaustivo de cada uno de los patrones y de su agresividad, bajo el aprendizaje previo transmitido por los expertos especializados en este tipo de tumor. Para resolver el problema de la falta de experiencia como experto en anatomía patológica, la autora tuvo que aprender cuál era el pronóstico para cada uno de ellos. Cabe destacar que, esta tarea se considera muy costosa para aquel que no es patólogo con experiencia en este tipo de cáncer. Se requieren conocimientos basados en la experiencia para realizarla con éxito en poco tiempo.

	Número de parches
Sano	772
Grado II	1.506
Grado III	792
Total	3.070

Cuadro 3.1: Tabla de número de imágenes por grado.

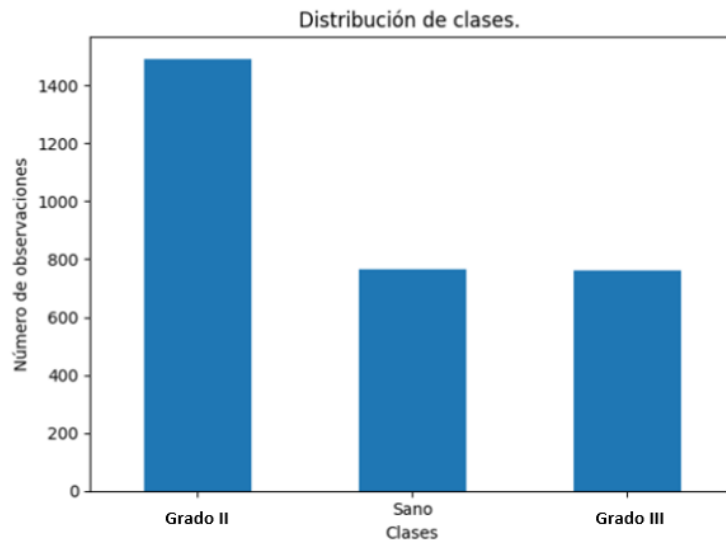


Figura 3.7: Distribución de clases en el *groundtruth*.

Tras realizar el etiquetado de las imágenes, se observa en el Cuadro 3.1 y en la Figura 3.7 que, los datos están desbalanceados para cada clase. Normalmente, los algoritmos de clasificación de imágenes supervisados suelen sobreajustarse a la clase con mayor número de observaciones, en este caso el “grado II”. Por lo tanto, para paliar el problema, es necesario aplicar al conjunto de entrenamiento una técnica de balanceo, pudiendo ser *undersampling* y *oversampling*. Ambas técnicas tienen como objetivo balancear el número de muestras de todas las clases. Por un lado, *undersampling* pretende eliminar algunas muestras de la clase mayoritaria, y por consiguiente, conseguir eliminar esa gran diferencia entre el número de observaciones de cada clase. Para este proyecto, puede ser una buena opción debido a que la base de datos utilizada es grande. Sin embargo, se debe tener en cuenta que al tratarse de imágenes muy complejas, la eliminación de información valiosa podría perjudicar al modelo. Por otro lado, *oversampling* es la agregación de copias a la clase minoritaria, como la duplicación de los datos en el conjunto de entrenamiento o la creación de nuevos ejemplos mediante transformaciones geométricas, rotaciones y traslaciones de la imagen.

Es importante recalcar que, para este proyecto se usa una base de datos relativamente grande por lo que no es necesario aplicar técnicas de *data augmentation*. Esta técnica crea artificialmente muestras sintéticas a partir de las originales, mediante transformaciones geométricas como son traslaciones, rotaciones o escalado, que permiten un aumento considerable de los datos sin ser necesaria más recopilación. Es una técnica muy buena cuando se utilizan bases de datos pequeñas.

A modo de resumen, se expone la Figura 3.8, que recoge los algoritmos principales de la etapa del preprocesado destinada a la confección de la base de datos definitiva a nivel de parche, a partir de la cual se lleva a cabo el desarrollo de los modelos predictivos.

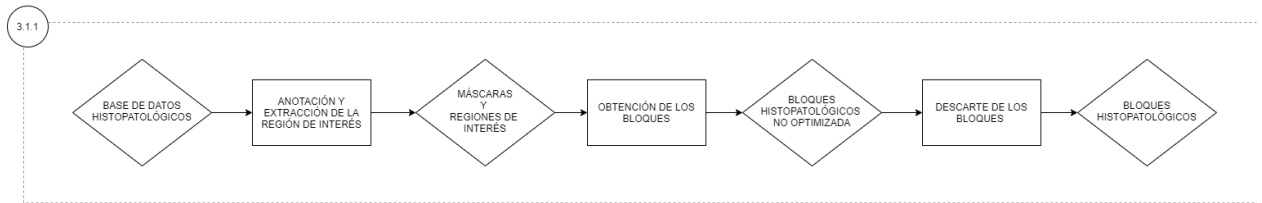


Figura 3.8: Diagrama de flujo del procedimiento realizado para la preparación de los datos.

3.3 Partición de los datos

La partición de los datos tiene como objetivo entrenar un modelo y validarlo. Esta técnica divide los datos en dos conjuntos, llamados test y entrenamiento. A su vez, el conjunto de entrenamiento se divide en dos subconjuntos, llamados entrenamiento y validación. Es muy importante saber que cada conjunto no debe contener las mismas imágenes, ya que, en ese caso, el modelo estaría siendo evaluado con las mismas imágenes con las que ha entrenado y, por tanto, los resultados no serían fiables ni tendrían la capacidad de predecir nuevas muestras. Por ello, lo que se pretende con esta técnica es que el modelo sea capaz de aprender a partir de un conjunto de imágenes, y que este conocimiento se pueda generalizar para futuros análisis de imágenes dotándole así de un carácter predictivo.

Para llevar a cabo la partición de datos en este proyecto, utiliza el método *k fold cross validation* que permite hacer una división de la base de datos en dos subconjuntos “entrenamiento y validación” un número de “*k*” iteraciones (Figura 3.9). De esta forma, en cada iteración las muestras de entrenamiento y validación son diferentes y nunca se repiten [23]. Este algoritmo se puede describir de la siguiente forma:

```

1: for  $i \leftarrow 1$  to  $K$  : do
2:    $train_i \leftarrow \sum_{j=1}^K F_j, \forall j \neq i$ 
3:    $val_i \leftarrow F_i$ 
4: end for
  
```

donde $F = F_1, F_2, \dots, F_K$ es el conjunto de datos de entrenamiento que se divide en K particiones.

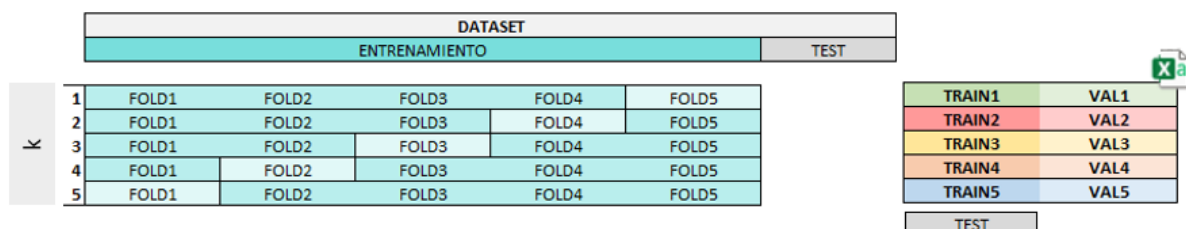


Figura 3.9: Dibujo explicativo de la metodología *kfold cross validation* con K valor igual a cinco. Elaboración propia.

Para la base de datos del presente proyecto se utiliza un valor para la $K = 5$, teniendo en cuenta el balanceo entre clases para que, en cada subconjunto (en inglés, *fold*), el número de muestras por clase sea similar y que todos los folds dispongan aproximadamente de la misma cantidad de muestras. Para

ello se desarrolla una función en entorno Matlab para dividir la base de datos etiquetada de manera equitativa por clases, generando, de esta manera, once hojas de cálculo Excel (Figura 3.10).

	TRAIN1	VAL1	TRAIN2	VAL2	TRAIN3	VAL3	TRAIN4	VAL4	TRAIN5	VAL5	TEST
SANO	556	62	556	62	556	62	556	62	556	62	154
GRADO BAJO	1.085	120	1.085	120	1.085	120	1.085	120	1.085	120	301
GRADO ALTO	571	63	571	63	571	63	571	63	571	63	158
	2.212	245	2.212	245	2212	245	2212	245	2212	245	613
TOTAL	2.457		2457		2457		2457		2457		613

Figura 3.10: Tabla que describe el número de imágenes por clase que hay en cada set elaborado por el método de *k fold cross validation*. Elaboración propia.

Este método se suele usar para modelos de clasificación supervisada donde no pueden usarse las mismas imágenes, tanto para entrenar como para validar. Para entender todo lo explicado con mayor detalle, a continuación se explican los conceptos teóricos más importante de la clasificación supervisada y la no supervisada.

3.4 Clasificación supervisada

Para llevar a cabo un modelo supervisado de clasificación se necesita una base de datos etiquetada. El modelo pretende encontrar relaciones entre las la entradas, en este caso las imágenes histopatológicas, y el resultado, es decir, a partir de las imágenes de entrada y conociendo la clase de cada una de ellas, el modelo predictivo ajusta una función que debe ser capaz de encontrar los patrones en las imágenes característicos de una clase determinada. Para ello, en este tipo de clasificación se necesita un conjunto de ejemplos que entrenen el modelo. Una vez el modelo haya aprendido y encontrado la relación entre la entrada y la salida deseada, éste estará listo realizar más predicciones.

3.4.1 Fundamentos de las redes neuronales artificiales: *Perceptrón multicapa (del inglés - MultiLayer Perceptron (MLP))*

La aparición de las redes neuronales artificiales (ANNs) fueron inspiradas por los sistemas neuronales biológicos debido a la capacidad de aprender y resolver problemas en base a la experiencia y extracción del conocimiento genérico. Como se puede apreciar en la Figura 3.11, las neuronas artificiales acogen características básicas del sistema neuronal real, como son las señales de entrada $X = [x_1, x_2, \dots, x_n]$ donde se les da un peso relativo $W = [w_1, w_2, \dots, w_n]$ que realza la intensidad de éstas. La señal de salida de la neurona viene descrita por el sumatorio de todas aquellas entradas ponderadas. Las ANN están distribuidas en paralelo como las conexiones del cerebro humano [24].

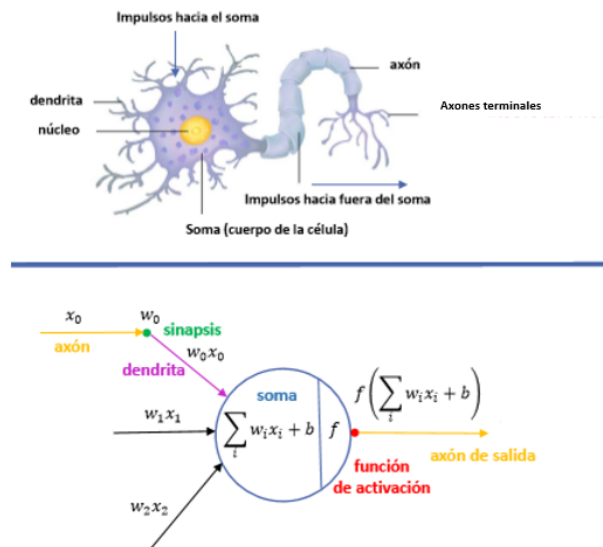


Figura 3.11: Similitud con las neuronas cerebrales. Fuente: [25]

Las redes neuronales se pueden clasificar dependiendo de su método de aprendizaje, como es el aprendizaje supervisado y el aprendizaje no supervisado. La diferencia entre ellos, es porque el supervisado hace uso de la información del groundtruth para asociar los patrones a una determinada clase. Por el contrario, los no supervisados intentan agrupar las muestras en su clase correspondiente sin conocer la etiqueta de las mismas.

Hoy en día, las ANNs se han establecido como una prometedora herramienta para el desarrollo de modelos de aprendizaje supervisados que aporta muy buenos resultados. Por este motivo, destaca tanto en el campo científico de la inteligencia artificial.

Para saber y entender que es una red neuronal, se van a describir los elementos básicos que componen los modelos computacionales. Se explican, también, las formalizaciones matemáticas correspondientes a la estructura del cerebro que se desea imitar. Además, se desarrollan con más profundidad los parámetros de ajuste necesarios para el desarrollo de una ANN.

Estructura básica de un MLP

Como se ha explicado anteriormente, las redes neuronales pretenden imitar la red neuronal biológica de la que disponen los humanos ya que, tiene la capacidad de aprender mediante ejemplos y, por consiguiente, llevar a cabo tareas como la identificación y detección de objetos, el reconocimiento de patrones, la clasificación de imágenes, etc. Por lo tanto, para conseguir una imitación de las redes neuronales cerebrales haciendo sinapsis las MLP deben seguir la misma estructura.

Como se aprecia en la Figura 3.12, la MLP se compone principalmente de N capas formadas, cada una de ellas, por un conjunto de nodos llamadas neuronas artificiales. Estas neuronas van a estar interconectadas entre las demás, de capa en capa, para poder así transmitir la información aprendida de manera óptima. En extremos de las ANNs se encuentran, por un lado, la capa de entrada, que es la primera capa de la estructura donde se toman los datos ponderados que se van a transmitir a las capas ocultas del centro, y por otro lado, la capa de salida donde la red neuronal llega a una conclusión (salida predicha, \hat{y}). En las capas del centro, es decir las capas ocultas, clasifica datos para cada clase.

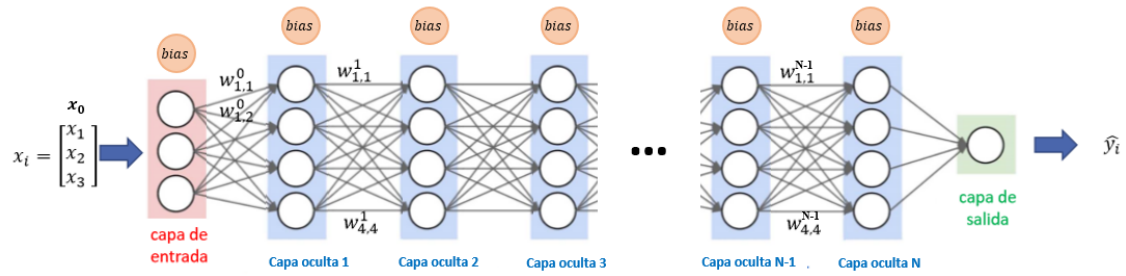


Figura 3.12: Esquema de un ANN de N capas, una capa de entrada, N-2 capa oculta y una de salida. Elaboración propia.

Funcionamiento del MLP

El funcionamiento de un algoritmo de retropropagación se compone de dos partes:

1. **Forward-propagating step:** El funcionamiento consiste en la transformación de los datos de entrada en una clasificación. Mediante sus pasos se aplican una serie de procesos matemáticos, como son la función de propagación, la función de activación y la función de transferencia, resolviendo de este modo la clasificación. [24].

La función de propagación se define como la suma ponderada de las entradas X multiplicadas por unos pesos W . El resultado de esta función, es interpretado por la siguiente neurona, también se conoce como valor neto de entrada (Ecuación 3.1). Los pesos son inicializados por números aleatorios pequeños o por transferencia de conocimiento de otras redes ya pre-entrenadas.

$$a_j = \sum_i x_i w_{j,i} + \theta_j \quad (3.1)$$

siendo i el índice de la entrada de la capa anterior, j es el número de neurona que le llegan las respuestas de la capa anterior, x los valores de las entradas, w los pesos ponderados de cada neurona y θ el término *bias*. Es importante aclarar la importancia de nodo *bias*, ya que permite obtener un mejor ajuste de los datos. Se puede modificar la función de activación sumando el término *bias*.

A continuación, se muestra la función de activación que se representa como x_j (Ecuación 3.2). Esta función activa o desactiva la salida de la neurona, propagando o no la información. De este modo, la red escoge la información más relevante en base a los pesos, que contienen la relación entre el problema y la solución, para transferir las características importantes de cada patrón a las futuras neuronas [24].

$$x_j = f\left(\sum_i x_i w_{j,i} + \theta_j\right) \quad (3.2)$$

2. **Backward-propagating step:** En esta segunda fase, se compara la salida predicha por el clasificador (\hat{y}) y el valor objetivo (y). La diferencia entre ellas, refleja la cantidad de error de aprendizaje (δ_k) (Ecuación 3.3) [24].

$$\delta_k = (y - \hat{y}) f'(a_k) \quad (3.3)$$

Donde k es la capa de salida, f' es la derivada de la función de activación aplicada en la primera etapa, a_k es la suma ponderada de la entrada para k y el error entre las salidas y y \hat{y} .

Cabe señalar que, δ permite a la red actualizar los pesos W que inicialmente se habían asignado aleatoriamente y orienta el aprendizaje de la red neuronal. Su finalidad es encontrar los pesos que describan mejor los datos de entrada.

El ajuste de los pesos o también conocido como el ajuste por regla generalizada se describe de la siguiente manera (Ecuación 3.4):

$$\Delta w_{k,j} = \eta \delta_k x_j \quad (3.4)$$

El ajuste del peso $w_{k,j}$, con una dirección desde j hacia k , depende de tres factores: δ_k , valor del error cometido entre la clase predicha y la real; x_j , valor de salida; η , el hiperparámetro tasa de aprendizaje (en inglés - *Learning rate*). [24].

Hiperparámetros

Resulta importante explicar los hiperparámetros de una ANN pues, de estos depende que se consiga una buena clasificación.

- Learning rate:** Controla la tasa de ajuste en la actualización de los pesos con respecto al gradiente de pérdida. En caso de asignar un valor alto, puede conducir a una inestabilidad en el aprendizaje de la red. En cambio, en caso de valores demasiado pequeños, se puede provocar una lentitud excesiva del aprendizaje y facilita el bloqueo en un mínimo local. En cuanto a las pérdidas que supone para el modelo, tal y como se aprecia en la Figura 3.13, un mal ajuste de la tasa de aprendizaje, supone un aumento en la pérdida de función de costes y, por consiguiente, un aumento en el error de predicción.

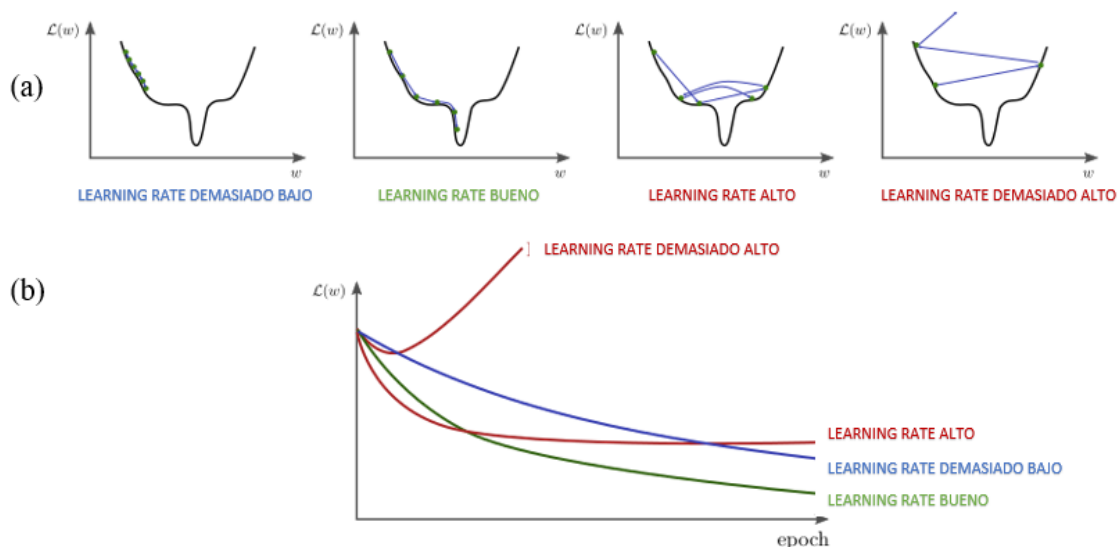


Figura 3.13: (a) Representaciones de cómo busca el mínimo global de la función de pérdidas con diferentes valores de *learning rate*. (b) Cómo responde la función de pérdidas con la elección del valor del *learning rate* con cada época. Fuente: [26]

- Épocas:** Son el número de iteraciones que el modelo efectúa el aprendizaje con el conjunto de datos de entrenamiento. Este hiperparámetro suele ser un número alto debido a que se usa en problemas complejos, por lo que empleando más iteraciones para la optimización de los pesos se consigue reducir del error. Sin embargo, asignar un número de épocas demasiado alto, podría conllevar a un sobreajuste (*overfitting*) de los datos de entrenamiento, es decir, el modelo aprendería de memoria todos los datos de entrenamiento y sus etiquetas, y, por tanto, perdería su carácter predictivo cuando tratara de clasificar las muestras de validación (Figura 3.14). Por otro lado, la asignación de pocas iteraciones conlleva a que el modelo no sea capaz de generalizar el conocimiento (*underfitting*) y, por consiguiente, a la falta de aprendizaje para una buena clasificación.

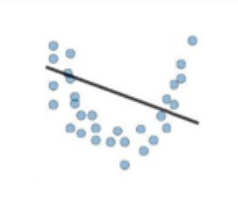


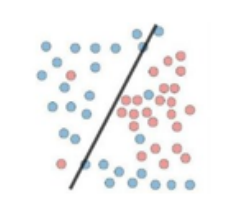
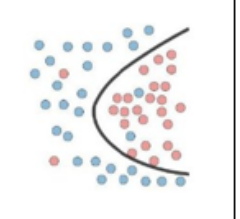
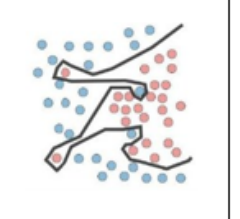



	Subajustado	Correcto	Sobreajustado
Síntomas	<ul style="list-style-type: none"> - Alto error de entrenamiento - Error de entrenamiento cerca de error de prueba - Sesgo alto 	<ul style="list-style-type: none"> - Error de entrenamiento ligeramente inferior al error de prueba 	<ul style="list-style-type: none"> - Bajo error de entrenamiento - Error de entrenamiento mucho menor que error de prueba - Alta varianza
Regresión			
Clasificación			
Deep Learning			
Remedios	<ul style="list-style-type: none"> - Modelo de complejidad - Añadir más funciones - Tren más largo 		<ul style="list-style-type: none"> - Regularizar - Obtener más datos

Figura 3.14: Tabla con los síntomas, remedios y efectos que provoca en problemas de regresión, clasificación y aprendizaje profundo cuando se entrenan pocas épocas *Underfitting*, número de épocas correcto y muchas épocas, *overfitting*. Fuente: [27]

- Batch size:** Se conoce como el número de muestras con las que se analiza el modelo en cada una de sus épocas.

El principal problema del aprendizaje en las ANNs es la minimización del valor de la función de pérdidas, Ecuación 3.3. Ésta está compuesta por dos partes, el término de error (δ_k) y el de regulación (η). El parámetro de regularización permite actualizar los pesos para minimizar la función de pérdidas en

las épocas posteriores. Por consiguiente, la actualización de los pesos termina siendo un problema de optimización (Figura 3.15). Los optimizadores que ayudan a encontrar el mínimo global de la función de error y su eficiencia dependen de la tasa de aprendizaje asignada. Los optimizadores más comunes son el Descenso por Gradiente Estocástico (SGD - del inglés, Stochastic Gradient Descent.), Descenso por Gradiente Adaptable (Adagrad), Descenso por Gradiente Adaptable (Adadelta), Momentum Adaptable (Adam) [28].

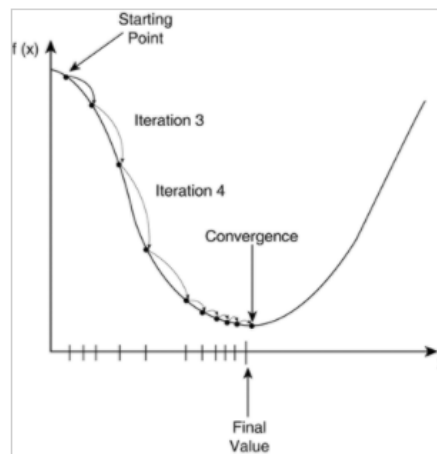


Figura 3.15: Esquema del funcionamiento del optimizador para encontrar el error mínimo de la función de pérdidas. Fuente: [24]

Una vez explicado, a grandes rasgos, el fundamento de las redes neuronales artificiales, a continuación se van a detallar en mayor profundidad las redes neuronales convolucionales (CNN) características del deep learning (aprendizaje profundo) en el que es enmarca el presente TFM. Las CNN han demostrado ser de gran validez para tareas de caracterización y reconocimiento de patrones a partir de una imagen.

3.4.2 Arquitectura red neuronal convolucional: Deep Learning

A diferencia de las MLP, las CNN son capaces de procesar sus capas imitando al comportamiento biológico del ser humano. Es decir, su capacidad de aprendizaje basada en ejemplos. Esa es una característica fundamental del *deep learning*. A diferencia del *machine learning* tradicional donde se extraían características de forma manual (es decir, desarrollando los descriptores manualmente) en el *deep learning* esas características se extraen de forma automática mediante las convoluciones entre capas. Por eso, mientras que el *machine learning* tradicional estaba basado en características, el deep learning está basado en ejemplos. Cuantos más ejemplos de una muestra, mejor es el aprendizaje. Su arquitectura se fundamenta principalmente en las capas convolucionales, salvo las últimas que están orientadas a la clasificación (Figura 3.18).

Una red neuronal convolucional es una variante de las (RNA) donde sus capas son convolucionales capaces de extraer automáticamente de una imagen características representativas para cada clase. Como se expone en la Figura 3.18, la arquitectura consta de dos partes esenciales:

- **Extractor de características o base model:** Una CNN está compuesta principalmente por capas convolucionales, que son las encargadas de extraer las características mediante funciones de activación, llamada *ReLU (Rectifier Lineal Unit)*, y capas de *pooling* que permiten reducir las dimen-

siones de los volúmenes de características de cada capa. Es importante detallar que, las primeras capas extraen características más básicas, generales, como contornos, mientras que las capas más profundas obtienen características de más alto nivel.

Convolución

La convolución consiste en aplicar un filtro, también conocido como kernel, a toda la imagen para extraer las características más relevantes en base a éste (Figura 3.16).

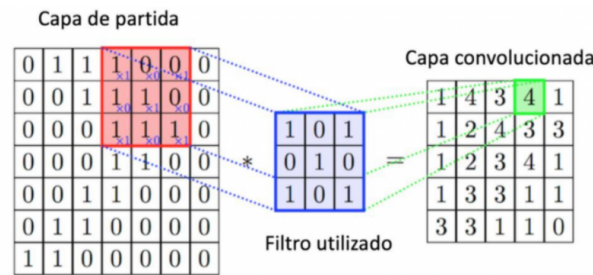


Figura 3.16: Representación del funcionamiento de la convolución. Fuente: [29]

Pooling

Es necesario llevar a cabo, para este tipo de arquitecturas, un muestreo ya que, tras aplicar una convolución el número de neuronas aumenta considerablemente. Por tanto, las capas de *pooling* se encargan de reducir el tamaño de las capas convolucionales filtradas para así disminuir las dimensiones de la salida de cara a la siguiente convolución.

Tal y como se puede observar en la Figura 3.17, hay dos tipos de muestreo. Por un lado, *Max pooling* que consiste en escoger el máximo de la matriz convolucionada dentro un bloque de unas dimensiones determinadas, y, por otro lado, *Average pooling* que aplica la media de todos los valores que están dentro del bloque.

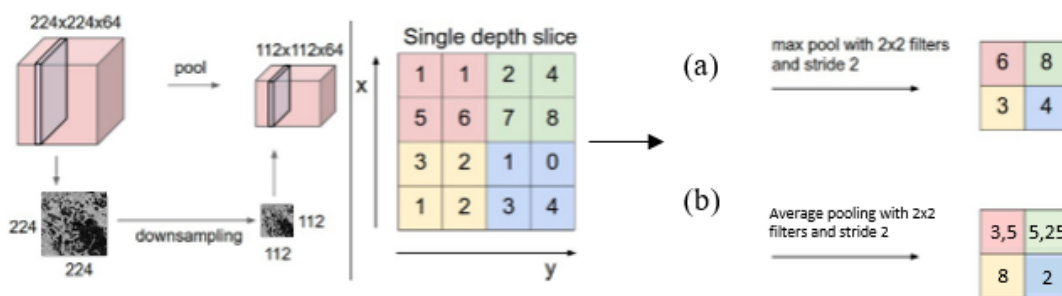


Figura 3.17: Tipos de muestreo. Elaboración propia

- **Clasificador o top model:** el principal objetivo de esta parte es llevar a cabo la clasificación de las imágenes a partir de las características representativas extraídas en la etapa anterior. Normalmente, esta arquitectura la conforman una capa que convierte los volúmenes de características en un vector 1-D, también conocida como capa *Flatten*, y capas continuas totalmente conectadas entre sí (técnica *fully connect*) y junto a una función de activación llamada *softmax* que permite asignar una probabilidad de clase.

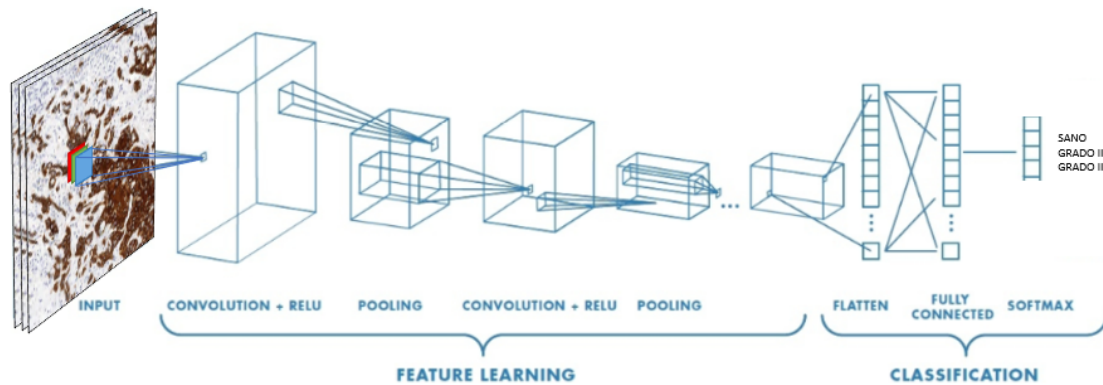


Figura 3.18: Arquitectura convolucional. Elaboración propia.

Las CNN tienen el mismo funcionamiento que el explicado para el MLP, siguiendo el *forward-backward propagation*, donde en la primera fase, se extraen las características con unos pesos aleatorios, se obtiene un error entre la clase predicha y la real mediante la función de pérdidas. Durante el backward, se actualizan los pesos mediante un optimizador con el objetivo de que en la época siguiente se reduzca el error minimizando la función de pérdidas y destacando poco a poco las características de cada patrón y atendiendo a la salida deseada y la predicha.

Normalmente una CNN requiere de una gran cantidad de muestras etiquetadas para que, durante el entrenamiento, sea capaz de extraer las características discriminativas entre las clases. Cabe detallar que, en el ámbito de la patología digital, en ocasiones no existen muchos datos a los que poder etiquetar por tratarse de alguna enfermedad rara, o también por ser una tarea tediosa que colleva mucho tiempo para su realización. Por esta misma razón, surgió la técnica conocida como *fine-tuning*. Ésta permite el uso de otras arquitecturas entrenadas con grandes bases de datos etiquetados para la transferencia de los pesos ajustados y optimizados. Dichos pesos, se transfieren en las capas del *base model*. La técnica permite congelar tantas capas como se desee, sin ningún tipo de actualización de pesos.

Arquitectura VGG (Visual Geometry Group)

Tras haber visto los conceptos principales para entender en mayor detalle el desarrollo de este trabajo final de máster, a continuación se va a detallar en profundidad la arquitectura utilizada para su consecución.

El presente trabajo final de máster se centra en las arquitecturas VGG16 y VGG19, que están formadas por 5 bloques convolucionales. El número de convoluciones que se realizan en los bloques es lo que las diferencia entre ellas. Tal y como se observa en la Tabla 3.19, el número de filtros aumenta (en potencia de 2) en cada bloque, es decir, al primer bloque se le aplica un conjunto de 64 filtros; al siguiente bloque se le adjudican 128 filtros; para el tercer bloque 256 filtros; y en los dos últimos bloques, 512 filtros. Durante todo el algoritmo, la dimensión de los kernels son de 3×3 y la reducción se realiza mediante la técnica *Maxpooling* de dimensión 2×2 . En la salida de las VGGs se aplica una agrupación aplanada y se conecta a una capa completamente conectada con 4096 neuronas, que a su vez, está conectada a otra capa con las mismas características y a una de 1.000 neuronas para llevar a cabo la clasificación. Cabe destacar que, todas estas capas están activadas por la función de activación ReLU y última capa por una *Softmax* (véase la Figura 3.20) [30] [31].

La arquitectura VGG19 fue diseñada para un desafío de reconocimiento visual a gran escala de ImageNet (ILSVRC). El objetivo de este desafío de clasificación de imágenes es entrenar un modelo que pueda clasificar correctamente una imagen de entrada en 1000 categorías de objetos independientes (por esa razón, en las últimas capas densas están compuestas por 1000 neuronas). Estas redes también demuestran una gran capacidad para generalizar a imágenes fuera del conjunto de datos de ImageNet mediante el aprendizaje por transferencia, técnica conocida como *fine-tuning*, como la extracción de características y el ajuste fino. Estas VGG reportaron muy buenos resultados en el *challenge* ImageNet [32].

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64	conv3-64	conv3-64	conv3-64
maxpool					
conv3-128	conv3-128	conv3-128	conv3-128	conv3-128	conv3-128
maxpool					
conv3-256	conv3-256	conv3-256	conv3-256	conv3-256	conv3-256
conv3-256	conv3-256	conv3-256	conv3-256	conv3-256	conv3-256
maxpool					
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
maxpool					
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Table 2: Number of parameters (in millions).

Network	A,A-LRN	B	C	D	E
Number of parameters	133	133	134	138	144

Figura 3.19: Familia de VGG. Fuente: [30]

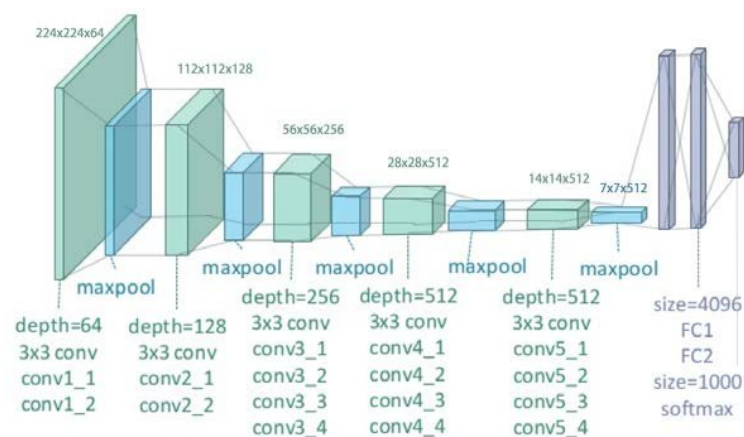


Figura 3.20: Arquitectura de VGG19. Fuente: [33]

En el presente trabajo final de máster, han implementado las arquitecturas VGG16 y VGG19, ya que ha demostrado ser útil en el campo de la patología digital para problemas de clasificación. Además, durante el análisis de estado del arte, las arquitecturas VGG destacan para el uso de técnicas de deep learning en la detección de los diferentes patrones para su clasificación en imágenes histológicas [34] [35] [36] [37] y [38]. Tal y como se ha comentado anteriormente, esta arquitectura ha sido pre-entrenada con los datos de ImageNet previamente para que los pesos iniciales adquieran un valor orientado a la base de datos usada en el presente proyecto (*fine-tuning*).

3.5 Clasificación no supervisada

A diferencia de la clasificación supervisada, son algoritmos iterativos que intentan llevar a cabo una clasificación sin tener en cuenta las etiquetas de las muestras. Dicha clasificación se basa en métodos de *clustering* que permiten llevar a cabo una agrupación de los píxeles en un número de clases determinadas. En este caso, puesto que se trata de imágenes, las técnicas más populares se centran en realizar las agrupaciones midiendo la distancia entre píxeles de manera que, aquellos píxeles cuyo valor de intensidad sea parecido, se agruparán dentro de la misma clase (*cluster*).

Hoy en día, están resurgiendo nuevos modelos no supervisados para el aprendizaje de la agrupación en los modelos no supervisados, ya que el etiquetado de las imágenes es una tarea muy tediosa y por lo tanto, una de las ventajas que presenta este tipo de modelo frente a los modelos supervisados es la reducción del tiempo para el desarrollo del modelo.

3.5.1 Clasificación no supervisada basada en autoencoders

En el presente proyecto, tal y como expone la Figura 3.21, para llevar a cabo la clasificación no supervisada he hecho uso de una arquitectura *autoencoder* para entrenar un modelo capaz de sacar características relevantes de las imágenes de entrada sin utilizar las etiquetas de dichas imágenes.

De esta forma, una vez se ha entrenado la arquitectura, se ha utilizado la parte del *encoder* para extraer las características del espacio latente de las imágenes. Y, a partir de esas características, se han implementado algoritmos de *clustering* para encontrar patrones subyacentes de las características asociados a las tres clases bajo estudio: sano, grado II y grado III.

Autoencoder

Los *autoencoders* son una variante de las redes neuronales que se usan para aprender, de manera no supervisada, las características de los datos. Hasta la fecha, este tipo de modelos suelen usarse para la generación de datos, la segmentación, reducción de ruido de una imagen y la reducción de dimensionalidad. Para este trabajo, se va a usar una arquitectura convolucional para la extracción de las características más representativas de cada clase.

Tal y como se muestra en la Figura 3.21 fase 1, el autoencoder consta de dos partes en la etapa de entrenamiento:

1. **Encoder:** Esta parte se encarga de extraer las características más representativas de los datos de entrada (X), en este caso las imágenes histopatológicas, y convertirlas en un espacio de menores dimensiones, también conocido como espacio latente (Z). Esta técnica realiza un mapeo no lineal para la extracción de las características de los datos, $Z_j = f_{\theta}(X_j)$. Para llevar a cabo la reducción,

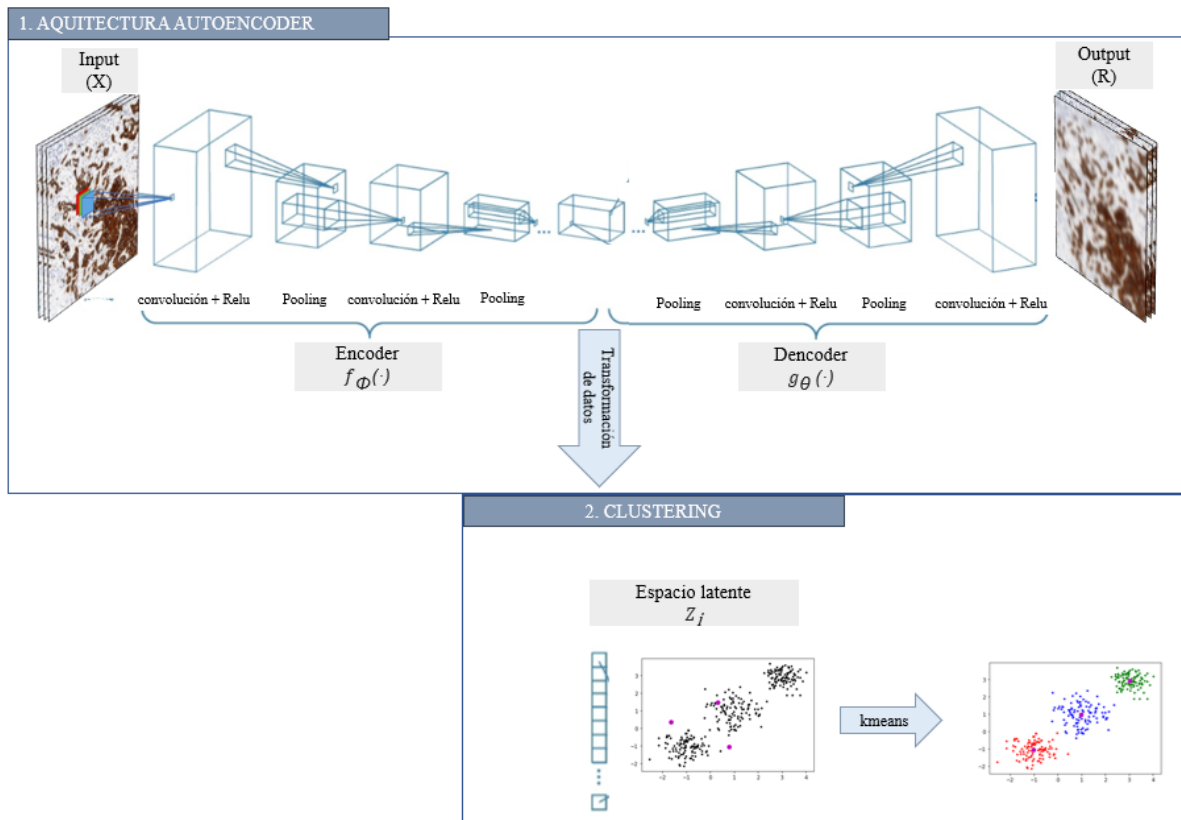


Figura 3.21: Esquema del algoritmo desarrollado para la clasificación no supervisada.

las dimensiones de las capas que forman el autoencoder se van disminuyendo hasta obtener las dimensiones deseadas. En este punto, se utilizan capas convolucionales y de pooling, explicadas en el Apartado 3.4.2 para llevar a cabo reducción de la dimensionalidad en la parte del encoder. A medida que se aplica una convolución, la dimensión va disminuyendo. Sin embargo, éste le aplica unos filtros de extracción de características que, dependiendo de los filtros asignados, extraerá más o menos características. Como se ha explicado anteriormente, las capas de pooling, permiten reducir la dimensión. Normalmente se aplica después de realizar una convolución.

2. **Decoder:** Permite la reconstrucción de la imagen en base al vector de características extraídas, $R_j = g_\theta(Z_j)$. Es necesario que las características incrustadas en el Z sea óptima debido a que, de ello dependerán, obtener una buena imagen reconstruida de tus datos de entrada. Una vez se ha reconstruido la imagen, el modelo calcula las pérdidas y las optimiza mediante el error cuadrático medio o *mean squared error* (MSE). Para poder llevar a cabo esta aproximación se hace uso de otro tipo de capas “Convolucionales Transpuestas” y “*Upsampling*” para desarrollar la parte del decoder. La convolucional transpuesta es la operación contraria a la que usa el típica convolución, ésta generaliza la información. En cuanto a la técnica *Upsampling*, permite aumentar la dimensión del bloque.

A continuación, en la Figura 3.22 se muestra la arquitectura del autoencoder convolucional empleada para el presente trabajo final de máster.

Layer (type)	Output Shape	Param #
Inputs (InputLayer)	(None, 128, 128, 3)	0
conv2d_1 (Conv2D)	(None, 128, 128, 128)	3584
conv2d_2 (Conv2D)	(None, 128, 128, 128)	147584
max_pooling2d_1 (MaxPooling2D)	(None, 64, 64, 128)	0
conv2d_3 (Conv2D)	(None, 64, 64, 256)	295168
conv2d_4 (Conv2D)	(None, 64, 64, 256)	590080
CODE (MaxPooling2D)	(None, 32, 32, 256)	0
up_sampling2d_1 (UpSampling2D)	(None, 64, 64, 256)	0
conv2d_transpose_1 (Conv2DTr)	(None, 64, 64, 256)	590080
conv2d_transpose_2 (Conv2DTr)	(None, 64, 64, 256)	590080
up_sampling2d_2 (UpSampling2D)	(None, 128, 128, 256)	0
conv2d_transpose_3 (Conv2DTr)	(None, 128, 128, 128)	295040
conv2d_transpose_4 (Conv2DTr)	(None, 128, 128, 128)	147584
OUTPUT (Conv2D)	(None, 128, 128, 3)	3459

Figura 3.22: Arquitectura autoencoder diseñada para llevar a cabo la reconstrucción de las imágenes histológicas Xj.

Como se aprecia en la figura anterior, la arquitectura empleada para el *encoder* la componen tres bloques convolucionales. Los dos primeros bloques están compuestos por dos convoluciones de 128 y 256 filtros, respectivamente, seguido de capas de *max pooling* con filtros tamaños de kernel (2,2) para reducir la dimensionalidad de los mapas de características previos a la mitad. El último bloque consta de una convolución de 512 filtros, que servirá para la adquisición del espacio latente. Para la fase de descomprensión, *decoder*, se aumentan las dimensiones aplicando la misma estructura que en los dos primeros bloques del *encoder*, pero a la inversa. Es necesario detallar que, para la base de datos empleada en este proyecto, ha sido necesario aplicar muchos filtros por ser imágenes con características muy complejas.

Clustering

En la segunda fase del algoritmo no supervisado, se implementan técnicas de *clustering*, que clasifican en base a las características de los datos en distintos grupos. Como se ha detallado anteriormente, la técnica de clasificación empleada para este trabajo final de máster es la conocida como “*kmeans*” que, consiste en un algoritmo iterativo capaz de agrupar las características extraídas de las muestras en k clases. En primer lugar, se definen k centroides aleatoriamente y, después, estos centroides se van actualizando en función del valor de las características. De esta manera, frente a una nueva muestra, se extraen las características, y se calcula la distancia entre dichas características y cada uno de los k centroides, en términos de distancia euclídea (aquí pon la fórmula de la distancia euclídea). Finalmente, a la muestra bajo estudio se le asigna la clase del centroide más cercano.

Por lo tanto, el método *kmeans* resuelve un problema de optimización, siendo la suma de las distancias cuadráticas la función a minimizar y se define como:

$$\min_C E(\mu_i) = \min_C \sum_{j=1}^k \sum_{i \in C_j} \|x_i - \mu_j\|^2 \quad (3.5)$$

donde x_i es cada observación del conjunto de datos, i el número de observaciones C_j el clúster, j sería el índice del cluster, μ_j la distancia media a ese centroide.

El algoritmo consta de tres pasos que se definen a continuación:

1. *Inicialización*: se escoge el número de grupos (k) y se establecen en el espacio de los datos (Figura 3.23 fase (a)).
2. *Asignación de la muestra a los centroides*: Se calculan las medias para cada muestra y se asigna un centroide en base a su media (Figura 3.23 fase (b)).
3. *Actualización de los centroides*: se actualiza la posición de los centroides y se reasignan los grupos en base al promedio (Figura 3.23 fase (c)).

$$\frac{\delta E}{\delta \mu_i} = 0 \implies \mu_j^{(t+1)} = \frac{1}{|C_j^{(t)}|} \sum_{x_i \in C_j^{(t)}} x_i \quad (3.6)$$

El modelo es iterativo por lo tanto los pasos 2 y 3 se repiten hasta que la posición de los centroides no cambia o cambia ligeramete sin superar un umbral.

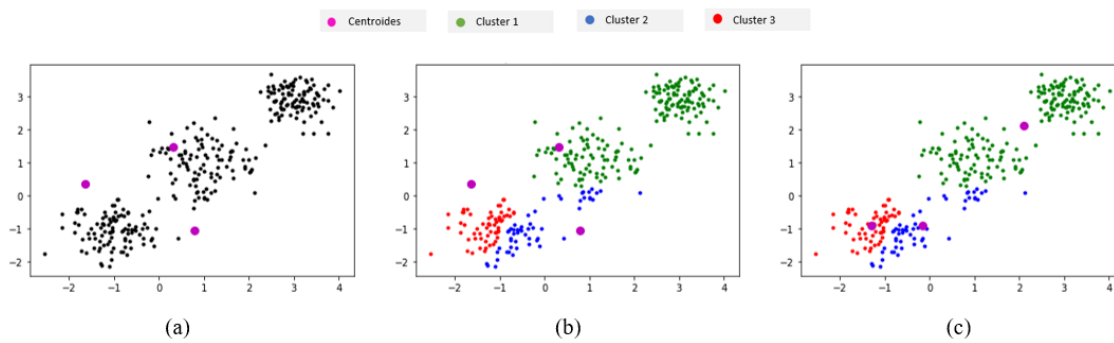


Figura 3.23: Esquema del funcionamiento de la técnica *kmeans*. (a) Inicialización de los centroides. (b) Paso de asignación de los centroides iniciales a los datos. (c) Paso de la actualización de la posición de los centroides.

Cabe destacar que, el *input* en la técnica *kmeans* debe ser un vector de una dimensión. En este caso, tal y como se muestra en la Figura 3.21, el espacio latente se adquiere fuera del *autoencoder* debido a que, las características de las imágenes son muy complejas y reducir mucho las dimensiones dentro de él conlleva a una mala reconstrucción de la imagen. Por lo tanto, una vez se ha entrenado el modelo *autoencoder*, para predecir las características del espacio latente, se extrae la parte del *encoder* y se le añade una capa que convierte los *outputs* del *encoder* en un vector. En particular, se ha añadido una capa de *Global Max Pooling*, con el objetivo de obtener un vector de características extrayendo el valor máximo del volumen de filtros en el eje de profundidad (ver Figura 3.22). El resultado de la predicción en la última capa del *encoder*, es el *input* de la etapa de clasificación.

Capítulo 4

Resultados y discusión

Se exponen los parámetros e hiperparámetros de los modelos desarrollados. Además, se presentan los resultados obtenidos por los modelos utilizados para la clasificación del cáncer de vejiga no músculo invasivo. Por último, se compararan los resultados entre los modelos desarrollados en este proyecto.

Índice de contenidos

4.1 Experimentos	50
4.1.1 Modelo supervisado	50
4.1.2 Modelo no supervisado	50
4.2 Métricas de evaluación	51
4.2.1 Análisis cualitativo: <i>Class Activation Mapping (CAM)</i>	51
4.2.2 Análisis cuantitativo	52
4.3 Resultados del modelo supervisado.	54
4.3.1 Etapa de validación	54
4.3.2 Predicción	57
4.4 Resultados del modelo no supervisado.	59
4.4.1 Etapa de validación	59
4.4.2 Etapa de predicción	60
4.5 Estado del arte	62
4.5.1 Contribución de este proyecto	63

4.1 Experimentos

A continuación, se muestra la arquitectura e hiperparámetros asignados para cada uno de los modelos desarrollados en este proyecto para la clasificación del cáncer de vejiga no músculo invasivo. Cabe detallar que, para obtener los parámetros e hiperparámetros óptimos para la gradación del cáncer de vejiga no músculo invasivo, se ha llevado a cabo una exploración rigurosa los mismos.

Cabe destacar que, para cada modelo desarrollado, se ha tenido en cuenta el acondicionamiento de los datos tal y como se ha explicado en el Apartado 3.2.

4.1.1 Modelo supervisado

Para la clasificación de cáncer de vejiga no músculo invasivo, se ha utilizado redes neuronales convolucionales, concretamente en la arquitectura VGG16 y VGG19. Entre ellas se ha escogido la VGG19 por ser la que mejores resultados se han obtenido. Durante la etapa de ajuste de hiperparámetros se ha utilizado constantemente como entrada al modelo una base de datos de imágenes histológicas con una dimensión $512 \times 512 \times 3$. En cuanto a los parámetros a optimizar del *base model* de la VGG19, se entrenaron todos los bloques y se utilizó la técnica de transferencia de conocimiento aplicando los pesos “imagenet”. Posteriormente, en el *top model* de la arquitectura, se eligió *GlobalAveragePooling2D* para convertir las características en un vector 1-D, seguido de dos densas de 1.000 neuronas cada una y en su última capa, formada por 3 neuronas (el número de clases), con un la función de activación *Softmax* para llevar a cabo la clasificación.

Los hiperparámetros escogidos para el entrenamiento del modelo final fueron 30 épocas, un *batch* de 16 muestras, el *Stochastic Gradient Descent* (SGD) como optimizador con un *learning rate* de 5×10^{-5} y una función de pérdidas *categorical cross entropy* por ser un problema de multiclase.

4.1.2 Modelo no supervisado

La entrada al modelo no supervisado ha sido una base de datos de imágenes histológicas de dimensiones $128 \times 128 \times 3$. Para el ajuste de los parámetros e hiperparámetros del modelo no supervisado se ha tenido en cuenta las dos técnicas usadas. Por un lado, la selección de la combiación óptima para la arquitectura propuesta del autoencoder, descrita en el Apartado 3.5, se entrenó durante 200 épocas de entrenamiento con un *batch-size* de 16, se utilizó “Adadelta” como optimizador con una tasa de aprendizaje de 0.75 para tratar de disminuir la función de pérdida de error cuadrático medio (MSE). Por otro lado, para el *clustering* se han asignado 300 iteraciones para cada entrenamiento del algoritmo *kmeans*, 3 centroides inicializados de manea aleatoria y 10 entrenamientos.

4.2 Métricas de evaluación

Una vez realizados todos los entrenamientos necesarios de los modelos desarrollados para este proyecto, se pretende analizar el rendimiento de cada uno de ellos desde dos perspectivas, una cualitativa y otra cuantitativa.

4.2.1 Análisis cualitativo: Class Activation Mapping (CAM)

CAM es una técnica propuesta por Zhou et al. [39] que muestra de manera visual de manera visual, las regiones de las muestras en las que se ha fijado el modelo para predecir una determinada clase. Además, permite determinar si dichas regiones están asociados con los patrones que reportan los expertos para determinar los diferentes grados del cáncer de vejiga. Su funcionamiento depende de una combinación lineal sobre la presencia de los diferentes patrones visuales que son determinados por los pesos de la predicción final y la clase contemplada (Figura 4.1).

Matemáticamente la importancia de la activación en el lugar (x,y) de una clasificación, se traduce como la siguiente Ecuación (4.1):

$$M_c(x,y) = \sum_k w_k^c f_k(x,y) \quad (4.1)$$

Donde $f_k(x,y)$ representa la activación de la neurona k en la última capa convolucional para localizarlo en los píxeles (x,y) . Para cierta clase c , la salida de la red viene definida como $\sum_k w_k^c f_k(x,y)$ suponiendo que w_k^c es el peso correspondiente a la clase c para cada neurona k . De modo que, la salida de la clase c será $\sum M_c(x,y)$ [39].

Este método es usado para que destaquen los patrones de crecimiento en diferentes localizaciones de la imagen para cada una de las clases predichas por el modelo [39].

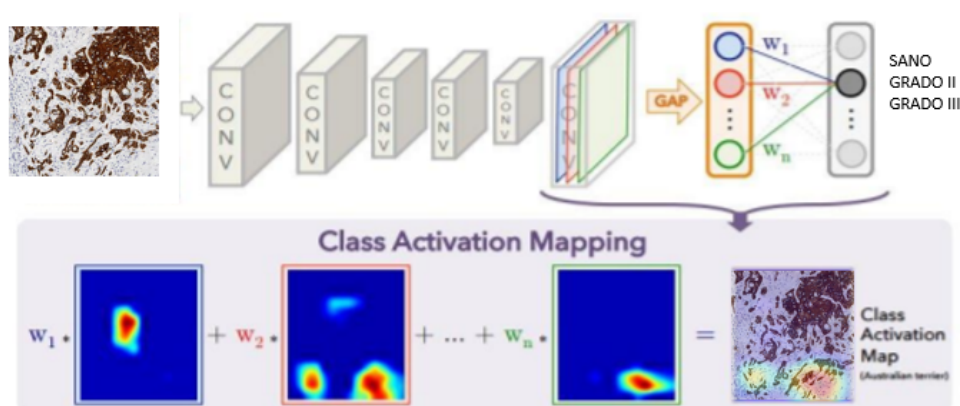


Figura 4.1: Esquema del funcionamiento de la técnica CAM. Elaboración propia

4.2.2 Análisis cuantitativo

Matriz de confusión

Dentro de las métricas cuantitativas, la matriz de confusión es una de las herramientas más intuitivas y sencillas, ya que mediante una tabla se puede observar el rendimiento del modelo de clasificación. En la Figura 4.2 se expone un ejemplo de clasificación binaria, donde se comparan las etiquetas predichas por el modelo con las etiquetas del ground truth. De esta forma, un verdadero positivo corresponde a una muestra que ha sido predicha con la clase “cáncer” y la etiqueta del ground truth era “cáncer”. Por el contrario, un falso positivo sería aquella muestra predicha como “cáncer” cuando en realidad era “sano”. La misma operación se aplica en el caso de los verdaderos negativos y los falsos negativos, pero para la clase contraria, es decir, clase “sano”. [40].

		Estimación del modelo	
		SANO	CÁNCER
Ground truth	SANO	Verdadero Negativo (TN)	Falso Positivo (FP)
	CÁNCER	Falso Negativo (FN)	Verdadero Positivo (TP)

Figura 4.2: Representación de la matriz de confusión. Elaboración propia

A partir de la matriz de confusión se pueden obtener diversidad de métricas de rendimiento, como las que se explican a continuación.

- **Sensibilidad (en inglés - *Recall (R) or sensitivity*):** Matemáticamente se usa la Ecuación 4.5 que expresa la efectividad para predecir las etiquetas verdaderas positivas frente a los valores acertados por el modelo [41].

$$\text{Sensibilidad} = \frac{VP}{VP + FN} \quad (4.2)$$

- **Especificidad (en inglés - *Specificity*):** Determina la efectividad para identificar las clases negativas que coinciden con la realidad, es decir, verdaderas negativas (Ecuación 4.3) [41].

$$\text{Especificidad} = \frac{TN}{TN + FP} \quad (4.3)$$

- **Valor Predictivo Positivo o Precisión (en inglés - *Positive predictive value (PPV)*):** Medida que expresa la proporción de pacientes con resultados TP que son correctamente diagnosticados [42] (Ecuación 4.4).

$$\text{PPV} = \frac{TP}{TP + FP} \quad (4.4)$$

- **Valor Predictivo Negativo (en inglés - Negative predictive value (NPV)):** Proporción de pacientes con resultados negativos que son correctamente diagnosticados (Ecuación 4.5) [42].

$$NPV = \frac{TN}{TN + FN} \quad (4.5)$$

- **Precisión (en inglés - Accuracy):** La precisión en modelos de clasificación se traduce como (Ecuación 4.6: Porcentaje de los valores verdaderos con respecto a todas las clasificaciones. Este valor es muy útil cuando las clases están balanceadas [41].

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (4.6)$$

- **F-Score:** Media armónica de la precisión [43] y viene definida como la siguiente Ecuación 4.7:

$$F - Score = \frac{2 \times PPV \times Sensibilidad}{PPV + Sensibilidad} \quad (4.7)$$

- **ACU (en inglés - Area Under Curve):** Mide el área bajo la curva ROC (en inglés - *Receiver Operating Characteristic*) poblacional que representa 1-Especificidad (Ecuación 4.3) frente a la sensibilidad (Ecuación 4.5) para cada posible valor umbral o punto de corte en la escala de resultados de la prueba en estudio (Ecuación 4.8) (Figura 4.3) [44].

$$\left. \begin{array}{l} y = Sensibilidad \\ x = 1 - Especificidad \end{array} \right\} \quad (4.8)$$

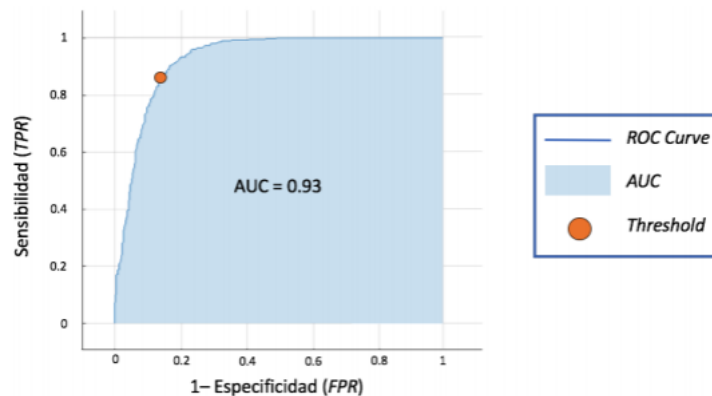


Figura 4.3: Ejemplo de Curva ROC y área bajo la curva (AUC). Fuente : [37]

- **Micro-Average:** Estima las contribuciones de todas las clases con la finalidad de calcular la métrica promedio. Se aplica para problemas de multclasificación con desbalanceo de muestras para cada clase.
- **Macro-Average:** Realiza el promedio sin tener en cuenta las muestras de cada clase, es decir, promediará atendiendo a todas las clases por igual.

4.3 Resultados del modelo supervisado

A continuación, se muestran los resultados obtenidos, durante la etapa de validación y la de test, del modelo supervisado desarrollado para la clasificación del cáncer de vejiga no músculo invasivo.

4.3.1 Etapa de validación

A continuación se van a mostrar las métricas tanto cuantitativas como cualitativas del funcionamiento de los modelos generados.

Métricas cuantitativas

En el Cuadro 4.1 se expone un promedio de la evaluación de los modelos obtenidos para cada subconjunto. Se muestran las diferentes métricas, explicadas anteriormente, para cada clase pudiendo así identificar la eficacia de diagnóstico. Destaca el valor de la sensibilidad en la métrica *micro-Avg*, que confirma la eficacia del modelo desarrollado con un 92 %.

Métrica	Sano	Grado II	Grado III		<i>micro-Avg</i>	<i>macro-Avg</i>
<i>Sensitivity</i>	1.00 ± 0.0072	0.93±0.0391	0.83±0.0847	-	0.92±0.0117	0.92±0.019
<i>Specificity</i>	1.00 ± 0.0049	0.91±0.0421	0.95±0.0258	-	0.96±0.0058	0.95±0.0081
<i>PPV</i>	0.99 ± 0.0142	0.91±0.0352	0.86±0.0609	-	0.92±0.0117	0.92±0.0154
<i>NPV</i>	1.00 ± 0.0025	0.93±0.0339	0.94±0.0255	-	0.96±0.0058	0.96±0.0067
<i>F1-score</i>	1.00±0.0107	0.92±0.0092	0.84±0.0275	-	0.92±0.0117	0.92±0.0137
<i>Accuracy</i>	1.00± 0.0055	0.92±0.0093	0.92±0.0104	-	0.95±0.0078	0.95±0.0078

<i>AUC</i>	0.93
------------	------

Cuadro 4.1: Promedio de las métricas de los modelos obtenidos por los subconjuntos en la etapa de validación para el modelo supervisado.

En la Figura 4.4, se muestra de forma visual el valor predictivo de cada modelo.

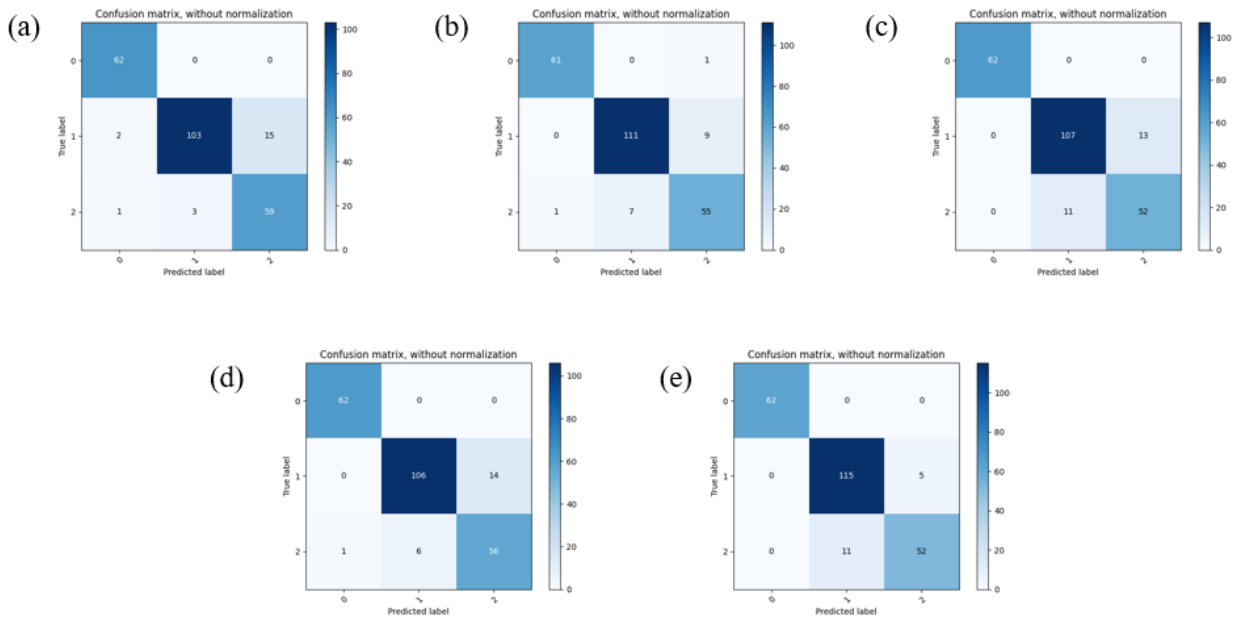


Figura 4.4: Matriz de confusión de los mejores modelos generados por los subconjuntos en la etapa de validación para el modelo supervisado. (a) $kfold=1$. (b) $kfold=2$. (c) $kfold=3$. (d) $kfold=4$. (e) $kfold=5$

A continuación, se observa en la Figura 4.5 las gráficas con respecto a la exactitud y la función de pérdidas durante las épocas del entrenamiento.

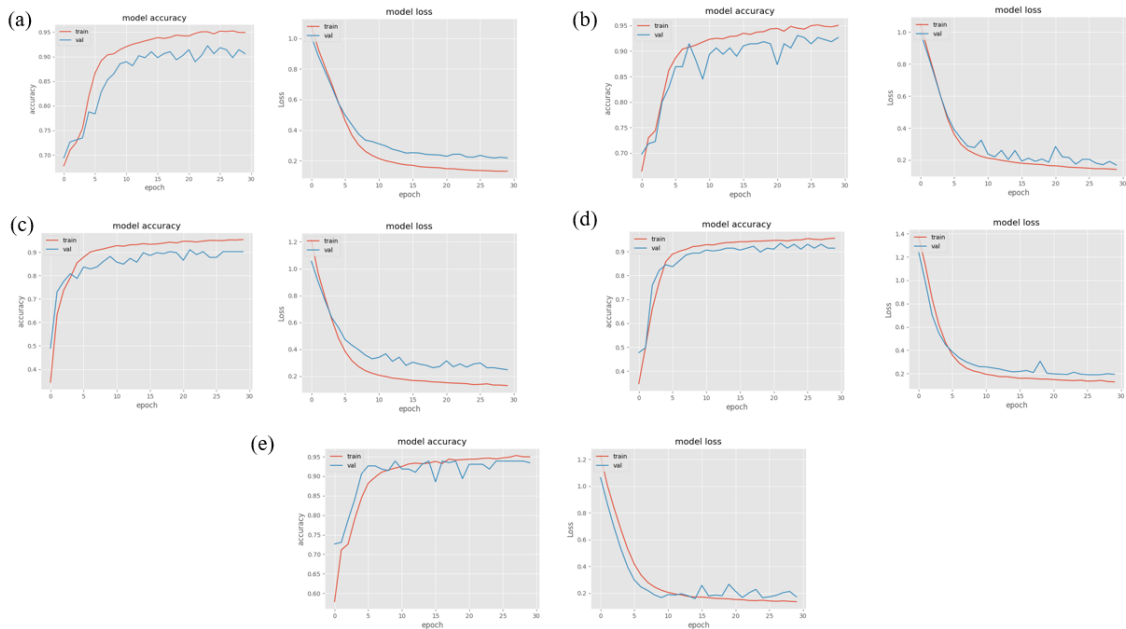


Figura 4.5: Representación de las curvas de aprendizaje durante el entrenamiento para el modelo supervisado. (a) Valor de la exactitud y de la función de pérdidas en cada época. (a) $kfold=1$. (b) $kfold=2$. (c) $kfold=3$. (d) $kfold=4$. (e) $kfold=5$.

Métricas cualitativas

En cuanto a las métricas cualitativas se determinan por la técnica CAM explicado en el Apartado 4.2.1.

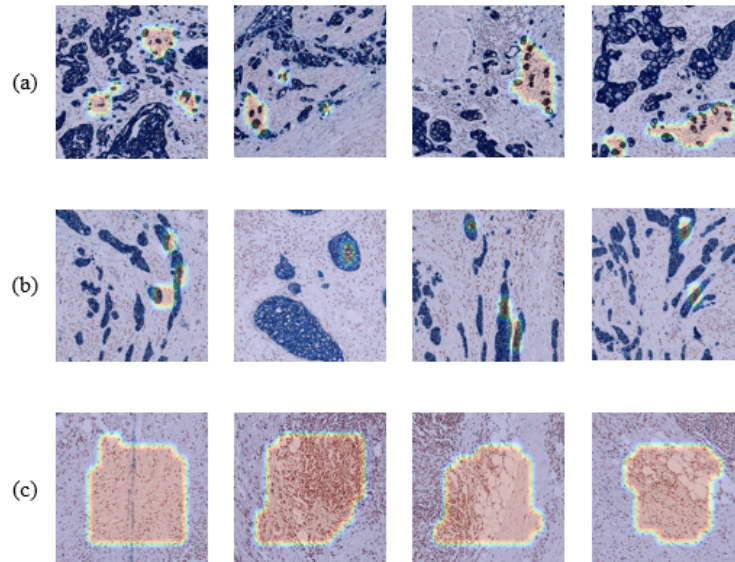


Figura 4.6: Resultados de los CAMs para los subconjunto de la etapa de validación. (a) Grado III. (b) Grado II. (c) Sano.

Una vez expuestos todos los resultados de la etapa de validación, el modelo es capaz de detectar con total seguridad la presencia del cáncer, y, en caso de haberlo, determinar la agresividad (90% de grado II y un 80% de grado III). Figura 4.6 sección (a) confirma que, el algoritmo es capaz de identificar los *buds* y descartar los demás patrones dentro de la clase “grado III”. En la sección (b) que corresponde al grado II, se aprecia como el algoritmo lo relaciona con patrones alargados dentro de la imagen. Es necesario recalcar que, los modelos tuvieron dificultad para determinar la clase de gradación “II”, visto que es un patrón que depende del umbral del tamaño específico. Por último, en la sección (c), el modelo descarta la presencia una muestra sana. Por lo tanto, podemos afirmar que, el modelo es capaz de detectar los patrones correspondientes a cada clase, y descartar lo otros con un 92% con un margen de error $\pm 1.2\%$ aproximadamente.

4.3.2 Predicción

Métricas cuantitativas

Métrica	Sano	Grado II	Grado III		micro-Avg	macro-Avg
Sensitivity	1	0.8362	0.8864	—	0.8901	0.9075
Specificity	0.9857	0.9639	0.8935	—	0.945	0.9477
PPV	0.9603	0.96	0.7178	—	0.8901	0.8794
NPV	1	0.8503	0.9626	—	0.945	0.9376
FI-score	0.9797	0.8939	0.7932	—	0.8901	0.8889
Accuracy	0.9894	0.8989	0.8918	—	0.9267	0.9267

AUC	0.927
------------	-------

Cuadro 4.2: Métricas del modelo definitivo en la etapa de predicción para el modelo supervisado.

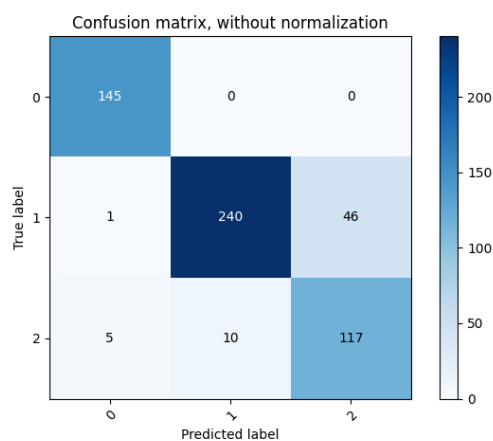


Figura 4.7: Matriz de confusión en la etapa de predicción para el modelo supervisado.

Métricas cualitativas

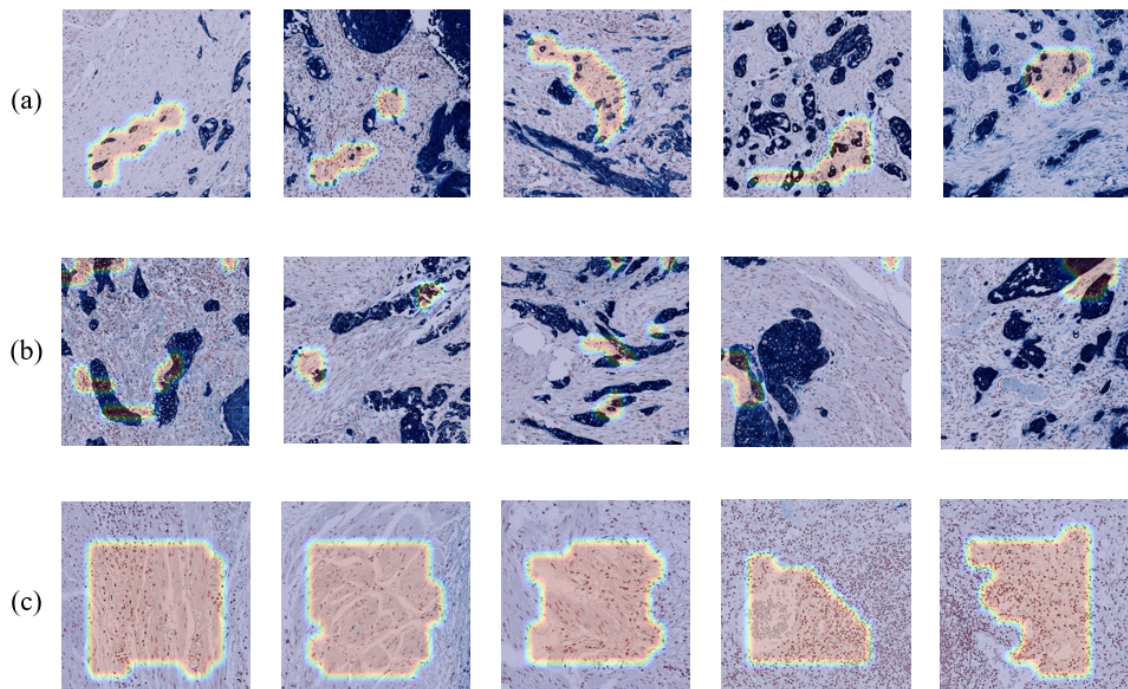


Figura 4.8: Resultados de los CAMs para los subconjunto de la etapa de validación. (a) Grado III. (b) Grado II. (d) Sano.

En la Figura 4.8, se exponen los resultados de CAMs para cada clase. La clasificación “Grado III” el modelo resalta zonas relevantes cuando hay presencia de un conjunto de *buds*. Para la clasificación del “Grado II” el modelo consigue identificar patron alargados como las trabéculas y nidos pero no es del todo preciso. Sin embargo, es importante detallar que el modelo no considera relevante para llevar a cabo la clasificación las zonas de la imagen donde hay presencia de tumor *budding*.

Discusión sobre la evaluación de predicción

En base a los resultado expuestos de la etapa de predicción, se aprecia que el modelo definitivo es capaz de identificar la presencia de células cancerígenas con total seguridad.

Es necesario destacar que, tal y como se expone en los resultado, existen predicciones erróneas. Esto se debido a que, para discernir entre los diferentes patrones, el experto se centra en la morfología que adquieren las células cáncergenicas en la frontera tumoral, el tamaño y la cantidad que existe en cada parche. La diferencia de tamaño entre el grado II, como es el patrón de nido, y grado III, como es el *bud*, puede ser mínima pudiendo confundir al modelo. Una solución frente a este problema sería, un aumento de imágenes para cada patrón, ya que a mayor cantidad de ejemplos, más conocimiento adquirirá para cada patrón. Otra solución sería la anotación por parte del experto de cada uno de los patrones que existen en la frontera tumoral, para así, una vez entrenado el modelo, éste será capaz de aprender el tamaño de cada patrón e identificar automáticamente los patrones en cada imagen que, por consiguiente, determinaría el grado de malignidad del tumor.

La intención de predecir con nuevos datos, es debido a que, pretende simular una práctica clínica real donde al algoritmo se le añaden muestras nuevas para clasificarlas en base a lo que ha aprendido.

A continuación, se procede a exponer los resultados obtenidos por el modelo no supervisado desarrollado en el presente trabajo final de máster.

4.4 Resultados del modelo no supervisado

4.4.1 Etapa de validación

Métricas cuantitativas

Métrica	Sano	Grado II	Grado III	<i>micro-Avg</i>	<i>macro-Avg</i>
<i>Sensitivity</i>	0.99±0.014	0.59±0.154	0.72±0.103	0.72±0.075	0.77±0.057
<i>Specificity</i>	1.00	0.85±0.058	0.73±0.102	0.86±0.038	0.86±0.035
<i>PPV</i>	1.00	0.79±0.077	0.49±0.083	0.72±0.075	0.76±0.051
<i>NPV</i>	1.00±0.005	0.69±0.075	0.88±0.038	0.86±0.038	0.86±0.033
<i>F1-score</i>	1.00±0.007	0.67±0.125	0.58±0.076	0.72±0.075	0.75±0.064
<i>Accuracy</i>	1.00±0.004	0.72±0.075	0.73±0.075	0.82±0.05	0.82±0.05

<i>AUC</i>	0.80±0.034
------------	------------

Cuadro 4.3: Promedio de las métricas de los modelos obtenidos por los subconjuntos en la etapa de validación para el modelo no supervisado.

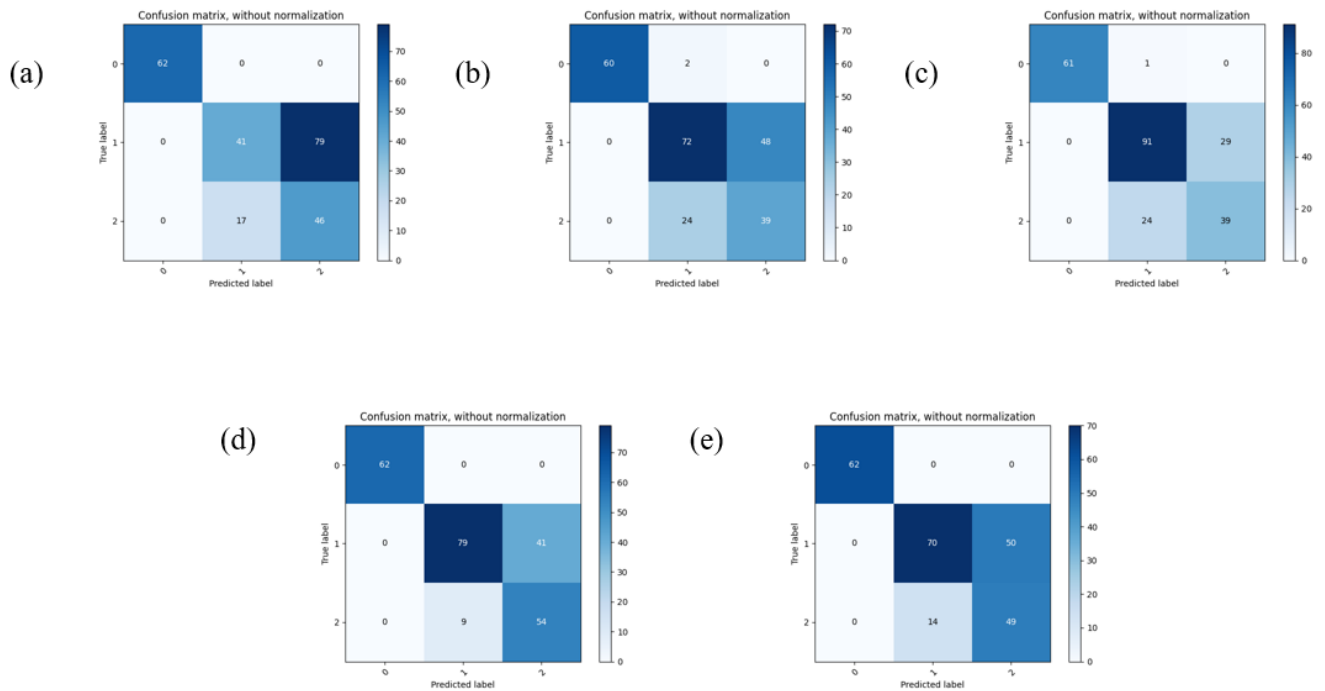


Figura 4.9: Matriz de confusión de los mejores modelos generados por los subconjuntos en la etapa de validación para el modelo no supervisado. (a) *kfold*=1. (b) *kfold*=2. (c) *kfold*=3. (d) *kfold*=4. (e) *kfold*=5

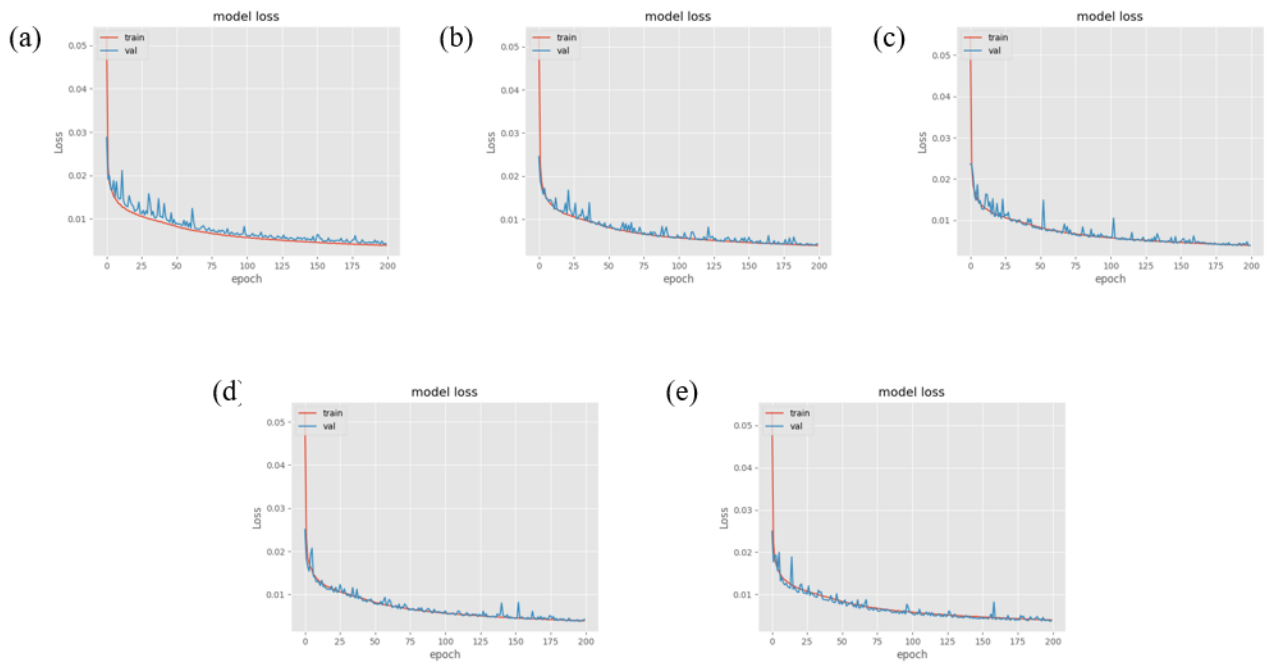


Figura 4.10: Representación de las curvas de aprendizaje durante el entrenamiento para el modelo no supervisado. (a) Valor de la exactitud y de la función de pérdidas en cada época. (a) kfold=1. (b) kfold=2. (c) kfold=3. (d) kfold=4. (e) kfold=5.

La eficacia de predicción del modelo no supervisado desarrollado de los distintos *kfold* es un 72% en términos de promedio. Se puede observar en el Cuadro 4.3 y la Figura 4.9 que, la sensibilidad para la predicción de los distintos grados es baja, siendo el grado II una predicción de 59% y grado III un 72%. El modelo no supervisado, es capaz de identificar con seguridad la presencia de cáncer en el parche. Por otro lado, la Figura 4.10, se observa como va disminuyendo el valor de la función de pérdidas del autoencoder de manera correcta.

4.4.2 Etapa de predicción

Métricas cuantitativas

Métrica	Sano	Grado II	Grado III		<i>micro-Avg</i>	<i>macro-Avg</i>
<i>Sensitivity</i>	0.99	0.96	0.75	—	0.92	0.90
<i>Specificity</i>	0.99	0.88	0.97	—	0.96	0.95
<i>PPV</i>	0.98	0.89	0.91	—	0.92	0.93
<i>NPV</i>	1.00	0.96	0.92	—	0.96	0.96
<i>F1-score</i>	0.99	0.93	0.82	—	0.92	0.91
<i>Accuracy</i>	0.99	0.92	0.92	—	0.94	0.94

<i>AUC</i>	0.92
------------	------

Cuadro 4.4: Métricas del modelo definitivo en la etapa de predicción para el modelo no supervisado.

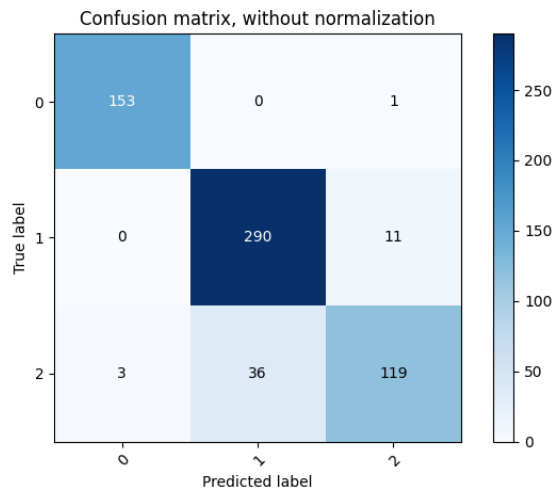


Figura 4.11: Matriz de confusión en la etapa de predicción para el modelo no supervisado.

Una vez presentados los resultados para cada etapa, se puede observar que el modelo implementado no predice correctamente para pequeñas cantidades en la base de datos, ya que la técnica de clasificación *kmeans* necesita de muchos valores para poder predecir de manera óptima. Una solución sería aumentar la cantidad de datos para cada subconjunto.

En la siguiente Tabla, se muestra la comparativa entre los modelos utilizados para el presente trabajo final de máster.

Métrica	VGG19	AE+kmeans
<i>Sensitivity</i>	0.89	0.92
<i>Specificity</i>	0.95	0.96
<i>PPV</i>	0.89	0.92
<i>NPV</i>	0.95	0.96
<i>F1-score</i>	0.89	0.92
<i>Accuracy</i>	0.93	0.94

Cuadro 4.5: Métricas del modelo definitivo en la etapa de predicción para el modelo no supervisado.

En base a lo expuesto en el Cuadro 4.5, el modelo supervisado tiene menos capacidad de predicción, siendo la *micro-Avg* un 89%. Sin embargo, se ha visto que, la VGG19 obtiene mejores resultados para la detección del tumor *budding* (88% para el modelo no supervisado y el 75% en el modelo no supervisado). Dentro del problema de gradación del cáncer de vejiga músculo invasivo y la detección del tumor *budding* el modelo supervisado sería el más indicado para usarse.

4.5 Estado del arte

Como se ha explicado en la introducción, hoy en día los patólogos gradifican el cáncer de vejiga músculo invasivo en base al patrón que hay en la frontera tumoral. Sabiendo estos patrones, los expertos pueden elegir un tratamiento apropiado para cada caso. Esta información hace posible desarrollar un modelo de clasificación automática.

Tras efectuar una recopilación de la información sobre el tema tratado en este trabajo final de máster, se puede observar que no existen trabajos orientados a la gradación del cáncer de vejiga músculo invasivo con tinción inmunohistoquímica CK AE1/3. Ciertamente existen varios estudios donde gradan el cáncer de vejiga en diferentes modalidades de imagen siendo las más comunes histológicas y cistoscópicas.

En cuando a los estudios de modalidad de imagen histológica es importante decir que fueron analizados desde parches de dimensiones 225x225 y tinción HE. Los artículos más destacados en esta modalidad son: *Binlin Wu et al.* [45], en su estudio, partieron de imágenes histológicas con tinción HE donde se y segmentaron los núcleos. Posteriormente, cuantificaron por agrupación local y alinearon los núcleos mediante un algoritmo de agrupación *k*-vecinos más próximos (del inglés - *k-nearest neighbours*), siendo *k* ejemplos más cercanos, que sirve para medir las distancias y el paralelismo del vecino más cercano. Finalmente, hicieron uso de un clasificador de machine learning conocido como “Support Vector Machine” (SVM), el cual clasifica el tumor en dos grupos, grado alto y bajo. En base a esta metodología, los autores reportaron resultados de sensibilidad y especificidad, alcanzando valores de 87.5% y 60%, respectivamente.

Por otra parte, *Ilaria Jansen et al.* [34] propusieron un estudio con una metodología similar al planteado en este trabajo fin de máster. En él, los autores estudiaron el tumor de vejiga no músculo invasivo para clasificar en los grados bajo y alto mediante, una segmentación con el modelo UNet que identifica la presencia de cáncer o no. Seguidamente, se aplica en aquellas imágenes un modelo de clasificación con arquitectura VGG16. Los resultados del estudio fueron “para la sensibilidad” un 71% y para la especificidad, un 76%.

Existen varios estudios que tratan sobre la clasificación en cáncer de vejiga por las características histológicas, aunque no para la gradación del tumor. Los temas dedicados más comunes son la presencia de biomarcadores para calcular el pronóstico del paciente [46], la distinción entre cáncer de vejiga no músculo invasivo y músculo invasivo [35],

Otros estudios que distan más de los objetivos del presente trabajo se basan en la identificación del cáncer de vejiga intentando discernir entre tejidos sanos y patológicos, es decir, no abordan una clasificación multiclase a fin de gradar el cáncer de vejiga. Por otra parte, también se pueden encontrar en la literatura estudios basados en la detección de biomarcadores para determinar el pronóstico del paciente [46], estudios para distinguir automáticamente si se trata de un cáncer de vejiga músculo invasivo o no músculo invasivo [35], la predicción de carga mutacional dependiendo de la estructura histológica [47], y la identificación de afectación a los ganglios linfáticos [36] [48].

Para la modalidad de imagen citoscopia existen varios estudios que estiman la presencia o no de cáncer [49], de afectación en los ganglios linfáticos [50] y detección de los distintos tipos de cáncer de vejiga y grados [51].

Además, existen otros artículos que miden la cantidad de células que hay en un campo 40X mediante modelos automáticos. El artículo [52], pretende determinar un pronóstico dependiendo de la cantidad de

número de buds, por lo tanto aplica un árbol de decisión como modelo de clasificación. Primero para las imágenes teñidas en inmunohistoquímica CK AE1/3 se segmentan las regiones de presencia de cáncer mediante la arquitectura *Inception* y a las imágenes inmunofluorescencia se les aplica *Random Forest* para detectar los buds. En otros artículos se analiza el tumor budding para otras estructuras anatómicas como es el cáncer de colon [38], [53] y cáncer de mama.

4.5.1 Contribución de este proyecto

En base a la literatura expuesta, la contribución del presente trabajo destacada por ser el primero estudio realizado con imágenes teñidas con inmunohistoquímica CK AE1/3, que permite la localización de los patrones de infiltración. Además, destacamos en ser los primeros en desarrollar dos modelos, uno supervisado y otro no supervisado, que es capaz en detectar la presencia de cáncer de vejiga músculo invasivo, y si lo hay, los patrones de infiltración en dos grados, grado II y grado III. También, ambos modelos son capaces de detectar el patrón infiltrativo, tumor *budding*, cuando hay presencia de otros patrones dentro del parche.

Conclusiones y líneas futuras

5.1 Conclusiones

De acuerdo con los objetivos principales de este trabajo, su principal aportación ha sido el diseño y desarrollo de un modelo predictivo basado en algoritmos de *deep learning* que sirvan de ayuda al diagnóstico a fin de facilitar la tediosa tarea de los patólogos en términos de tiempo y eficacia. Además, con este modelo también se pretende reducir el nivel de subjetividad que existe entre diferentes patólogos cuando tienen que analizar una determinada muestra. Se desarrolló dos algoritmos que consiguieron clasificar el patrón de agresividad del atendiendo a la arquitectura morfológica celular tanto en el frente tumoral como en su interior.

Primero realizó una exhaustiva investigación acerca de las técnicas de clasificación en el ámbito de la patología digital para el cáncer de vejiga y de la detección automática del tumor *budding*. Además, se estudiaron los patrones de crecimiento en el frente tumoral y la agresividad que le confieren al tumor. Se analizó con detalle el estudio de patrón tumor *budding*, la histología que éste adquiere, su comportamiento y cómo afecta, en términos de pronóstico, en la salud del paciente.

Las imágenes histológicas con las que se cuenta son demasiado grandes y tienen una carga computacional elevada. Por lo tanto, se desarrollaron distintas funciones que permitieron construir una base de datos de imágenes histopatológicas para poder trabajar sobre ellas y sean compatibles con los modelos predictivos.

En base al conocimiento adquirido de los patrones de agresividad y la supervisión de los expertos en el campo de anatomía patológica, se etiquetaron las imágenes histopatológicas atendiendo al grado de malignidad, y así obtener el *groundtruth* de la base de datos.

Se tuvo que acondicionar la base de datos para llevar a cabo el aprendizaje de los modelos. Para ello, se dividió mediante la técnica *kfold cross validation*, con la finalidad de reportar resultados fiables y objetivos. Se desarrollaron distintas funciones en torno el *software* Matlab para dividir las y preparar la base de datos para su posterior entrenamiento y validación.

Una vez se acondicionó la base de datos, se desarrollaron dos modelos a partir de *deep learning*, supervisado y no supervisado, para determinar la clasificación a nivel de parche del cáncer de vejiga. El modelo supervisado empleado para la clasificación fue la conocida como VGG19, y, por otro lado, para el modelo no supervisado se diseñó un autoencoder para extraer las características representativas de cada imagen y se clasificó con una técnica de *clustering* conocida *kmeans*.

Se obtuvieron métricas de evaluación de los modelos mediante un análisis cuantitativo y cualitativo, y así poder realizar una breve comparación entre ellos. Estos resultados demostraron ser los modelos valiosos para la predicción de la presencia de cáncer y su gradación.

Finalmente, a lo largo de la realización del trabajo final de máster se detectaron las limitaciones para su desarrollo y las líneas futuras del presente trabajo.

En definitiva, se puede concluir que, se ha logrado diseñar y desarrollar el primer sistema automático de clasificación, visto que en el estado del arte no existe ningún estudio que grade este tipo de tumor.

5.2 Líneas futuras

Dado que no existe ningún estudio en la literatura que lleve a cabo la clasificación del grado del cáncer de vejiga músculo invasivo, se proponen varias ideas que han ido surgiendo a lo largo de su realización. A continuación, se listan las distintas propuestas para mejorar y optimizar la solución a la clasificación:

- Realizar un modelo no supervisado con la técnica *Deep Clustering*. El modelo no supervisado generado en el presente proyecto, comprende la primera parte de la técnica *Deep Clustering*. El principal motivo por el que se propone esta técnica es debido a que, el método utilizado para la clasificación no supervisada *kmeans* no suele ser óptimo para imágenes tan complejas. Por lo tanto, se desea comprobar, con la técnica *deep clustering*, una solución más eficiente para la clasificación no supervisada del cáncer de vejiga.
- Diseñar un sistema *end-to-end* incorporando la identificación de ROIs y clasificación de las mismas en un solo algoritmo.
- Realizar una clasificación a nivel de biopsia, es decir, a imagen completa en vez de por parche, utilizando un mapa de calor que represente el grado de la biopsia.

Bibliografía

- [1] *Cáncer*. Inf. téc. Organización Mundial de la Salud. URL: <https://www.who.int/topics/cancer/es/>.
- [2] *Tipos de cáncer*. Inf. téc. Asociación Española Contra el Cáncer. URL: <https://www.aecc.es/es/todo-sobre-cancer/tipos-cancer>.
- [3] *Factores de riesgos*. Inf. téc. Asociación Española Contra el Cáncer. URL: <https://www.aecc.es/es/todo-sobre-cancer/que-es-cancer/factores-riesgo>.
- [4] *Global Cancer Observatory (GCO)*. Inf. téc. International Agency for Research on Cancer. URL: <https://gco.iarc.fr/>.
- [5] Sociedad Española de Oncología Médica (SEOM). *Las cifras del cáncer en España 2020*. 2020. URL: https://seom.org/seomcms/images/stories/recursos/Cifras_del_cancer_2020.pdf.
- [6] *Cáncer de vejiga*. Inf. téc. Asociación Española Contra el Cáncer. URL: <https://www.aecc.es/es/todo-sobre-cancer/tipos-cancer/cancer-vejiga/evolucion-cancer-vejiga>.
- [7] *Small cell bladder cancer and hematuria*. Inf. téc. Santa Clara Valley Medical Center Internal Medicine Residency. URL: <https://scvmcmed.com/2018/10/17/small-cell-bladder-cancer-and-hematuria-10-17-18/>.
- [8] Laura Lorenzo Soriano. “Valor pronóstico de la presencia de un componente tumoral indiferenciado *tumor budding* en pacientes con carcinoma vesical músculo-invasivo”. Doctorado en Medicina. Hospital Universitario i Politènic la Fe, jul. de 2018.
- [9] *Cáncer de vejiga no músculo invasivo: Guía para pacientes*. The Official Foundation of the American Urological Association. 1000 Corporate Boulevard, Linthicum, MD 21090, 2017.

- [10] *Pronóstico*. Inf. téc. Asociación Española Contra el Cáncer. URL: <https://www.aecc.es/es/todo-sobre-cancer/tipos-cancer/cancer-vejiga/pronostico>.
- [11] G. Feil y A. Stenzl. “Pruebas de marcadores tumores en el cáncer de vejiga”. es. En: *Actas Urológicas Españolas* 30 (ene. de 2006), págs. 38-45. ISSN: 0210-4806. URL: http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S0210-48062006000100006&nrm=iso.
- [12] *Pruebas para detectar el cáncer de vejiga*. Inf. téc. American Cancer Society. URL: <https://www.cancer.org/es/cancer/cancer-de-vejiga/deteccion-diagnostico-clasificacion-por-etapas/como-se-diagnostica.html>.
- [13] Gospodarowicz MK Brierley JD y Wittekind C. *TNM Classification of Malignant Tumours, 8th Edition*. eighth. UICC International Union Against Cancer: Oxford: Wiley-Blackwell, 2017.
- [14] Rafael Jimenez y col. “Grading the Invasive Component of Urothelial Carcinoma of the Bladder and Its Relationship With Progression-Free Survival”. En: *The American Journal of Surgical Pathology* 24 (jul. de 2000), págs. 980-987. DOI: 10.1097/00000478-200007000-00009.
- [15] *Guía clínica sobre el cáncer de vejiga con invasión muscular y metastásico*. European Association of Urology. 2010.
- [16] Imai Takeshi. “The growth of human carcinoma: a morphological analysis.” En: *Fukuoka Igaku Zasshi* (1954).
- [17] C Beaton y col. “Systematic review and meta-analysis of histopathological factors influencing the risk of lymph node metastasis in early colorectal cancer”. En: *Colorectal disease : the official journal of the Association of Coloproctology of Great Britain and Ireland* 15 (ene. de 2013). DOI: 10.1111/codi.12129.
- [18] S Coverlizza y col. “Colorectal adenomas containing invasive carcinoma. Pathologic assessment of lymph node metastatic potential”. En: *Cancer* 64 (dic. de 1989), págs. 1937-47. DOI: 10.1002/1097-0142(19891101)64:93.0.CO;2-X.
- [19] *Aperio CS2 Highly Reliable, Desktop Digital Pathology Scanner*. Inf. téc. San Diego, CA, USA: Leica Biosystems. URL: <https://www.leicabiosystems.com/es/patologia-digital/escanear/aperio-cs2/>.
- [20] *MicroDraw*. 2018. URL: <https://microdraw.pasteur.fr/>.
- [21] *OpenSlide*. URL: <https://openslide.org/>.
- [22] *OpenSlide: The Whole-Slide image processing library that stands behind most of Digital Pathology Projects*. Jul. de 2019. URL: <https://medevel.com/openslide-digital-pathology/>.

- [23] Lei Tang Payam Refaeilzadeh y Huan Liu. “Cross-Validation”. En: *Arizona State University* (nov. de 2008).
- [24] S. Lek e Y.S. Park. “Artificial Neural Networks”. En: *Encyclopedia of Ecology*. Ed. por Sven Erik Jørgensen y Brian D. Fath. Oxford: Academic Press, 2008, págs. 237-245. ISBN: 978-0-08-045405-4. DOI: <https://doi.org/10.1016/B978-008045405-4.00173-7>. URL: <http://www.sciencedirect.com/science/article/pii/B9780080454054001737>.
- [25] Valeriana Naranjo Ornedo y Adrian Colomer Granero. *FUNDAMENTOS DE REDES NEURONALES: Deep Learning aplicado al análisis de señales e imágenes*.
- [26] Hafidz Zulkifli. *Understanding Learning Rates and How It Improves Performance in Deep Learning*. Ene. de 2018. URL: <https://towardsdatascience.com/understanding-learning-rates-and-how-it-improves-performance-in-deep-learning-d0d4059c1c10>.
- [27] Paweł Cisło. *Red Neuronal Artificial. Underfitting and Overfitting*. Ago. de 2020. URL: <https://www.pinterest.es/pin/130393351699140243/>.
- [28] Sebastian Ruder. *An overview of gradient descent optimization algorithms*. Ene. de 2016. URL: <https://ruder.io/optimizing-gradient-descent/index.html#fn:6>.
- [29] *Lecciones de Deep Learning: Las capas convolucionales y el Max pooling*. Jul. de 2020. URL: <https://www.datasciencespain.com/en/lecciones-de-deep-learning-las-capas-convolucionales-y-el-max-pooling/>.
- [30] Andrew Zisserman Karen Simonyan. “VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION”. En: *International Conference on Learning Representations (ICLR)* (abr. de 2015).
- [31] Prashanth Kurumbudel. *What is the VGG-19 neural network?* Abr. de 2019. URL: <https://www.quora.com/What-is-the-VGG-19-neural-network>.
- [32] Adrian Rosebrock. *ImageNet: VGGNet, ResNet, Inception, and Xception with Keras*. 2017. URL: <https://www.pyimagesearch.com/2017/03/20/imagenet-vggnet-resnet-inception-xception-keras/>.
- [33] Yufeng Zheng, Clifford Yang y Aleksey Merkulov. “Breast cancer screening using convolutional neural network and follow-up digital mammography”. En: mayo de 2018, pág. 4. DOI: 10.1117/12.2304564.
- [34] Ilaria Jansen y col. “Automated Detection and Grading of Non-Muscle-Invasive Urothelial Cell Carcinoma of the Bladder”. En: *The American Journal of Pathology* 190 (abr. de 2020). DOI: 10.1016/j.ajpath.2020.03.013.

- [35] Peng-Nien Yin y col. “Histopathological distinction of non-invasive and invasive bladder cancers using machine learning approaches”. En: *BMC Medical Informatics and Decision Making* 20 (dic. de 2020). DOI: 10.1186/s12911-020-01185-z.
- [36] Le Hou Joel Saltz Rajarsi Gupta. “Spatial Organization and Molecular Correlation of Tumor Infiltrating Lymphocytes Using Deep Learning on Pathology Images”. En: *PMC* (2018).
- [37] Valery Naranjo Ornedo José Gabriel García Pardo y Adrián Colomer Granero. “DISEÑO Y DESARROLLO DE UN SISTEMA AUTOMÁTICO DE CLASIFICACIÓN DE ESTRUCTURAS GLANDULARES EN IMÁGENES HISTOLÓGICAS DE PRÓSTATA”. Tesis de mtría. Universidad Politécnica de Valencia, 2018.
- [38] John-Melle Bokhorst Lucia Rijstenberg Danny Goudkade Iris Nagtegaal Jeroen van der Laak Francesco Ciompi. “Automatic Detection of Tumor Budding in Colorectal Carcinoma with Deep Learning”. En: *LNCS* 11039 (2018).
- [39] Bolei Zhou y col. “Learning Deep Features for Discriminative Localization”. En: (dic. de 2015).
- [40] José Antonio Pérez Carrasco Álvaro Artola Moreno. “Clasificación de imágenes usando redes neuronales convolucionales en Python”. Tesis de mtría. Dpto. Teoría de la Señal y Comunicaciones Escuela Técnica Superior de Ingeniería Universidad de Sevilla, 2019.
- [41] Marina Sokolova y Guy Lapalme. “A systematic analysis of performance measures for classification tasks”. En: *Information Processing Management* 45.4 (2009), págs. 427-437. ISSN: 0306-4573. DOI: <https://doi.org/10.1016/j.ipm.2009.03.002>. URL: <http://www.sciencedirect.com/science/article/pii/S0306457309000259>.
- [42] Douglas G Altman y J Martin Bland. “Statistics Notes: Diagnostic tests 2: predictive values”. En: *BMJ* 309.6947 (1994), pág. 102. ISSN: 0959-8138. DOI: 10.1136/bmj.309.6947.102. eprint: <https://www.bmj.com/content/309/6947/102.1.full.pdf>. URL: <https://www.bmj.com/content/309/6947/102.1>.
- [43] Yutaka Sasaki. “The truth of the F-measure”. En: *Teach Tutor Mater* (ene. de 2007).
- [44] Juan Manuel Muñoz Pichardo Ana Rocío del Valle Benavides. “Curvas ROC (Receiver-Operating-Characteristic) y sus aplicaciones.” Tesis de mtría. Departamento de Estadística e Investigación Operativa de la Universidad de Sevilla, 2017.
- [45] Sushmita Mukherjee Binlin Wu Samantha V. Nebylitsa y Manu Jain. “Quantitative diagnosis of bladder cancer by morphometric analysis of H&E images”. En: *SPIE* (nov. de 2015).
- [46] Nicolas Brieu Katharina Nekolla Ansh Kapil. “Deep Learning Based Detection Of Tumor Tissue Compartments Improves Prognostic Immunoprofiling In Muscle-Invasive Bladder Cancer”. En: (2019).

- [47] Sung Hak Lee Hongming Xu Sunho Park. “Using transfer learning on whole slide images to predict tumor mutational burden in bladder cancer patients”. En: *bioRxiv* (2019).
- [48] G. Thomas Brown Stephanie A. Harmon Thomas H. “Multiresolution Application of Artificial Intelligence in Digital Pathology for Prediction of Positive Lymph Nodes From Primary Tumors in Bladder Cancer”. En: *JCO Clinical Cancer Informatics* (2020).
- [49] Josip Spanjol Ivan Lorencina Nikola Andelica y Zlatan Cara. “Using multi-layer perceptron with Laplacian edge detector for bladder cancer diagnosis”. En: *ScienceDirect* (2019).
- [50] Hirokazu Nosato Atsushi Ikeda y Yuta Kochi. “Support System of Cystoscopic Diagnosis for Bladder Cancer Based on Artificial Intelligence”. En: *Journal of Endourology* 34.3 (2020).
- [51] Axel Semjonow Okyaz Eminaga Nurettin Eminaga y Bernhard Breil. “Diagnostic Classification of Cystoscopic Images Using Deep Convolutional Neural Networks”. En: *JCO Clin Cancer Inform* (2018).
- [52] Nicolas Brieu y col. “Automated tumour budding quantification by machine learning augments TNM staging in muscle- invasive bladder cancer prognosis”. En: *Scientific Reports* 9 (mar. de 2019). DOI: 10.1038/s41598-019-41595-2.
- [53] Mohammad Faizal Ahmad Fauzi y col. “Tumor Budding Detection System in Whole Slide Pathology Images”. En: *Journal of Medical Systems* 44 (feb. de 2020). DOI: 10.1007/s10916-019-1515-y.

Parte II

Presupuesto

Capítulo 1

Presupuesto

En este capítulo del trabajo final de máster se consigue una valoración de la tasa económica, referida en euros. Se desglosan los diferentes costes como: el coste de mano de obra y los computacionales, dejando de lado los vinculados con la maquetación, desplazamientos o la electricidad.

Índice de contenidos

1.1 Presupuestos parciales	3
1.1.1 Coste mano de obra	3
1.1.2 Coste materiales.	4
1.1.3 Coste maquinaria	4
1.2 Presupuestos total	6

1.1 Presupuestos parciales

Se especifican de manera independiente los presupuestos parciales de los costes de personal, costes de los materiales, las del *software* y las del *hardware* para realizar este trabajo.

1.1.1 Coste mano de obra

Sin duda, la gran mayoría del valor económico del proyecto se debe a los costes de mano de obra. Este apartado es el más costoso visto que este proyecto es inevitablemente dependiente del personal para la realización de los algoritmos. En la Tabla 1.2 que se aprecia a continuación, se muestra una aproximación del salario que adquiere, determinada en tiempo, cada personal que ha contribuido en este trabajo; en este caso trata de D^a Valery Naranjo Ornedo, clasificada como cotutor del proyecto; D. David Ramos Soler, como tutor del proyecto; D. José Gabriel García Pardo, como cotutor. Todos aquellos mencionados anteriormente, están especializados en ámbitos diferentes en la temática, tienen como principal objetivo de revisar y orientar al alumno de una buena realización de la tesis y el autor del proyecto D^a. Anna Esteve Domínguez.

En el desglose el jornal muestra las horas aportadas por parte de los participantes considerándose jornada completa por parte del alumno (8 horas diarias en los días laborales) y 1 hora semanal por parte de los tutores. El coste por hora para cada uno (14.23€/h para el estudiante, 27.15€/hora para el co-tutor y

40€/hora para el tutor), que se calcula a partir de un sueldo base, siendo 1.300 €/mes el salario de estudiante y 2.600€/mes para el tutor. Cabe destacar que además del salario base, según el Ministerio de empleo y Seguridad Social para el 2017, se debe integrar en el jornal dos pagas.

Nº	Código	Descripción	Horas trabajadas (h)	Precio unitario (€/h)	Total (€)
1	MO.C	Directora del TFM (Catedrática)	20	40	800
2	MO.D1	Cotutor del TFM (Doctor)	70	27.15	1.900,50
3	MO.D2	Cotutor del TFM (Doctor)	10	27.15	271,50
4	MO.A	Autora del TFM: (Alumna del Máster de Ingeniería Biomédica)	640	14.23	9.107,2
TOTAL					12.079,2

Cuadro 1.1: Descomposición de los costes de mano de obra.

1.1.2 Coste materiales

Para la adquisición de las imágenes utilizadas en el presente trabajo final de máster, se ha llevado a cabo varias tareas previas tales como: la preparación de las biopsias de diferentes pacientes y el procedimiento de obtención de los portas con el tejido histológico.

Nº	Código	Descripción	Cantidad (u)	Precio unitario (€/h)	Total (€)
1	MM.B	Biopsias	120	600,00	7.200
2	MM.M	Muestras	140	10,00	1.400
TOTAL					8.600

Cuadro 1.2: Descomposición de los costes de los materiales.

1.1.3 Coste maquinaria

Para la realización del proyecto ha sido principalmente necesario los recursos tanto de *software* como de *hardware*. En este apartado se va a detallar las herramientas utilizadas teniendo en cuenta tanto la vida útil del material como el periodo como el intervalo de tiempo amortizado. Es cierto que, se han utilizado herramientas de *software* proporcionados por la Universitat Politècnica de València y es por eso que, dicha parte no habrá costes asociados. En cuanto a la escritura del proyecto se ha hecho uso del editor de texto online www.overleaf.com. En ambas tablas se ha valorado el tiempo de uso de cada equipo con respecto a los 9 meses de duración del trabajo; el período de amortización y el precio de cada una de las herramientas o equipos utilizados.

Nº	Código	Descripción del equipo/software	Precio (€)	Vida útil (meses)	Tiempo de uso (meses)	Total (€)
		DELL XPS 9370				
1	ME.OP	i7-8550U/IntelHDGraphics620/ 16GBDDR3-SDRAM/512GB SDD/13"	860	60	9	129
2	ME.CVB	Servidor grupo CVBLAB NVIDIA TITAN V	4.000	60	9	600
TOTAL						729

Cuadro 1.3: Descomposición de los costes del *hardware*.

Nº	Código	Descripción del equipo/software	Precio (€)	Vida útil (meses)	Tiempo de uso (meses)	Total (€)
1	ME.O	Overleaf	0	12	9	0
2	ME.SLM	Licencia Matlab Studen R2018a	0	12	9	0
3	ME.SLP	Licencia Python 3.7	0	12	9	0
4	ME.SLK	Librerías Keras	0	12	9	0
5	ME.MOF	Pack Office 365 Hogar y Estudiantes 2019	149	48	9	27,93
6	ME.LW	Licencia Windows 10 Home	175	24	9	32,81
TOTAL						60,74

Cuadro 1.4: Descomposición de los costes del *software*.

Descripción	Total (€)
Costes de <i>hardware</i>	729
Costes de <i>software</i>	60,74
TOTAL	789,74

Cuadro 1.5: Coste total de la maquinaria

1.2 Presupuestos total

El presupuesto total del trabajo (Cuadro 1.6) es la suma de los costes de los presupuestos parciales comentados anteriormente. Se tienen en cuenta los impuestos y otros gastos como son el 13% de los gastos generales, un 6% del beneficio industrial y el 21% correspondiente al Impuesto sobre el Valor Añadido (IVA).

Descripción	Coste (€)
Costes de mano de obra	12.079,2
Costes de material	8.600
Costes de maquinaria	789,74
Suma (PEM)	21.468,94
Gastos generales	2.790,96
Beneficio industrial	1.288,13
SUMA	25.548,03
IVA (21%)	5.365,08
PRESUPUESTO TOTAL	30.913,116

Cuadro 1.6: Presupuesto total del proyecto.

